Cheap Donuts and Expensive Broccoli: The Effect of Relative Prices on Obesity

Jonah B. Gelbach
University of Maryland, Department of Economics
Florida State University, College of Law

Jonathan Klick
Florida State University, College of Law

Thomas Stratmann
George Mason University, Department of Economics

March 13, 2007

Abstract: Americans have been getting fatter since at least the mid 1980s. To better understand this public health problem, much attention has been devoted to determining the underlying cause of increasing body weights in the U.S. We examine the role of relative food prices in determining an individual's body mass index, arguing that as healthful foods become more expensive relative to unhealthful foods, individuals substitute to a less healthful diet. Using data from the National Health Interview Survey (NHIS) for the period 1982-1996, we find that individual BMI measures, as well as the likelihood of being overweight or obese, exhibit a statistically significant positive correlation with the prices of healthful relative to unhealthful foods. These results are robust to endogenizing the relative price measure. While the magnitudes of our estimates suggest that relative price changes can only explain about 1 percent of the growth in BMI and the incidence of being overweight or obese over this period, they do provide some measure of how effective fat taxes would be in controlling the obesity epidemic. Our estimates imply, for example, that a 100 percent tax on unhealthful foods could reduce average BMI by about 1 percent, and the same tax could reduce the incidence of being overweight and the incidence of obesity by 2 percent and 1 percent respectively.

Cheap Donuts and Expensive Broccoli:
The Effect of Relative Prices on Obesity

1. INTRODUCTION

Obesity has grown at an epidemic rate in the U.S. since the mid 1980s, according to the Centers for Disease Control.[1] Hedley, *et. al.* (2004) report that during the period 1999-2002, more than 65 percent of adults ages 20 and above were overweight or obese, while 16 percent of children ages 6 to 19 were overweight. Given the large private and public costs that are associated with overweight and obese individuals, public health researchers and policymakers have devoted significant resources to address and reverse the upward trend in body weights. However, despite the increased attention this health issue has received, the causes of obesity are still relatively poorly understood, which limits the ability of public policy measures to affect the problem (Ogden, *et. al.*, 2003).

A number of economists have identified "natural" causes of the obesity epidemic, such as declining food prices and increasingly sedentary occupations. Continuing in this vein, we analyze the effect of changes in relative food prices on individual behavior. Specifically, we hypothesize that the price of unhealthful foods has declined more rapidly than the price of healthful foods. According to basic demand theory, such changes in relative food prices will cause a rational individual to change his consumption bundle in terms of which foods he chooses to consume. Specifically, on the margin, he will consume relatively more unhealthful foods and relatively fewer healthful foods. This change in diet will have the natural effect of increasing the individual's body weight as the nutritional content of his food consumption worsens. Examining this relationship

[1] http://www.cdc.gov/nccdphp/dnpa/obesity/

between relative food prices and the incidence of obesity is important for evaluating the efficacy of fat taxes in fighting the obesity epidemic.

To test this hypothesis, we examine the effect of changes in the relative prices between healthful and unhealthful foods on individual body mass indexes as reported in the National Health Interview Survey (NHIS) for the period 1982-1996, as well as the likelihood that an individual is overweight or obese, controlling for other relevant differences across individuals. We find that, as healthful foods become more expensive relative to unhealthful foods, individuals exhibit higher body mass indexes, as well as higher likelihoods of being overweight or obese, and these effects are statistically significant. We also examine the potential heterogeneity of this effect across the sex, race, and education dimensions. To rule out endogeneity of the price measure, we implement an instrumental variables strategy using the price of unleaded gasoline as our instrument (we discuss the justification of this instrument below).

One technical issue we face arises from the fact that our food price data are collected only at the regional level, which limits us to only four "clusters" and thus poses problems for the standard clustering method (since this method is justified only for large numbers of clusters). To remedy this problem, we use the wild cluster bootstrap-*T* method developed in Cameron, Gelbach, and Miller (2006, henceforth CGM). Using this approach, we find that our results remain statistically significant even after we use CGM's method account for within group dependence.

Our results suggest that individuals do substitute between healthful and unhealthful foods when relative prices change, as suggested by basic demand theory. However, this relative price effect accounts for only about 1 percent of the rise in average

BMI during our sample period. From a policy standpoint, our estimates provide some context for the proposal of using taxes on unhealthful foods as a means to curbing obesity. Our results suggest that the sensitivity of individuals to changes in relative food prices may not be sufficient to make fat taxes, within plausible ranges, a viable tool to lower obesity.

In section 2 of this paper, we review the economics literature regarding the determinants of the rise in obesity. In section 3, we lay out a very simple constrained optimization model that describes the effect of changes in relative prices on an individual's consumption decisions. In section 4, we describe how we constructed the food price indexes we use for our analysis and we present evidence that there was a substantial divergence between the price trends of healthful and unhealthful foods around the time the obesity epidemic is thought to have begun. We follow this with our empirical analysis of the effect of relative prices on individual level body mass index in section 5, and we address the problem regarding inference when the number of clusters is small in section 6. We consider the likely effects of food taxes and subsidies as public health measures aimed at fighting the obesity problem in section 7, after which we conclude.

2. ECONOMIC DETERMINANTS OF OBESITY

Finkelstein, Ruhm, and Kosa (2005) provide a comprehensive review of the literature on the economic causes of obesity as well as estimates of the costs of the obesity epidemic. For our purposes, the most closely related work to our own is Philipson and Posner (2003) and Lakdawalla and Philipson (2002). These papers provide the first systematic "price theory" explanation for the rise in obesity. Essentially, they argue that food prices have declined substantially over the last century, leading people to consume more food. Further, as the work of individuals became more sedentary in the latter half of the twentieth century, individuals burned fewer calories in general. The combination of consuming more calories and burning fewer calories at work naturally leads to increases in body weight. Lakdawalla and Philipson provide a micro analysis of the effect of food prices in total (i.e., they do not distinguish between healthful and unhealthful foods) on BMI, and find that declining food prices may account for as much as 40 percent of the change in BMI from 1976-1994.

Chou, Grossman, and Saffer (2004) suggest that not only did the nominal price of food fall in recent years, but so did preparation costs, as prepared food—which has generally lower nutritional content—became easier to purchase. They show that there is a positive correlation between the number of fast food restaurants in an area and the BMI of individuals living in that area, as well as a positive relation with the likelihood that an individual is obese. They also find that there is a large (negative) elasticity of BMI and the incidence of obesity with respect to the prevailing fast food prices an individual faces in his home area, though the effects are not statistically significant. They find smaller (though statistically significant) elasticities of BMI and incidence of obesity when they examine the prices available in full service restaurants and grocery prices.

Cutler, Glaeser, and Shapiro (2003) suggest that technological changes likely had an especially large effect on the price of mass produced, high calorie food, leading individuals to change the composition of their diet toward favoring unhealthful foods. They further suggest that this price sensitivity is most likely strongest among individuals lacking self-control, such as those individuals on the high end of the weight distribution. They provide some time series evidence to this effect.

While the bulk of the economic work on obesity focuses on the cost of food, Klick and Stratmann (2007) offer evidence, at least in the case of diabetics, that one factor leading to an increase in BMI has been the introduction of relatively low cost treatments for the eventual health effects of poor eating habits and exercise regimes. That is, as the privately borne health costs of obesity decline, individuals rationally substitute away from fastidious diet and exercise regulation. They speculate that this effect may be more general as individuals expect technology to bail them out of the long term adverse consequences of their behavior.

3. THE EFFECT OF RELATIVE PRICES ON DIET CHOICE

A rational individual will make his decisions regarding diet in the same manner in which he makes all consumption decisions. That is, his choice of what foods he will consume can be modeled as a constrained optimization problem. To simplify the exposition of this problem, without loss of generality, we assume that an individual can spend his disposable income (*I*) on healthful food (*H*) which costs p_H per unit, unhealthful food (*J*) which costs p_J per unit, and other goods (*X*), which cost p_X per unit.[2]

---

[2] We note that the cost of food consumption is not one dimensional. In principle, the cost parameters p_H and p_J subsume the nominal price of the food, the health costs of the food (which is presumably negative for healthful food and positive for unhealthful food, at least in the relevant ranges for the purposes of this paper), and the costs involved with preparing the food. Although the data we examine do not shed light on the food preparation component of relative food prices, work by Chou, Grossman, and Saffer (2004) suggests that this too is an important determinant of the effect of relative food costs on an individual's consumption decisions.

The usual demand function for healthful and unhealthful foods may be written

$$H^* = H^*(p_H, p_J, p_X, I) \tag{1}$$

and

$$J^* = J^*(p_H, p_J, p_X, I). \tag{2}$$

Standard comparative-statics results imply that consumption of healthful food is decreasing in its own price (i.e., $H_1^* < 0$), increasing in the price of unhealthful food (i.e., $H_2^* > 0$—since unhealthful food is a substitute for healthful food), and, if healthful food is normal, increasing in income (i.e., $H_4^* > 0$). Similarly, consumption of unhealthful food is decreasing in its own price (i.e., $J_1^* < 0$), increasing in the price of healthful food (i.e., $J_2^* > 0$), and, if unhealthful food is normal, increasing in income (i.e., $J_4^* > 0$).

Thus, basic demand theory implies that increases in the price of healthful food relative to the price of unhealthful food will lead consumers to shift consumption from healthful to unhealthful food. In fact, Lakdawalla and Philipson (2002) show that what might be called the overall price of food—a weighted average of both healthful and unhealthful foods—has fallen over time relative to the prices of other goods. To the extent that this reduction in food prices is due disproportionately to reductions in the prices of unhealthful foods, substitution toward unhealthful foods should occur from both healthful foods and overall goods. Because body weight is likely to increase more with increases in unhealthful food intake than with increases in healthful food intake,[3] this

[3] We note that in principle, all foods can be healthful or unhealthful under some circumstances. For example, if an individual is on the verge of starvation, even high fat foods with very little nutritional content beyond generic calories will likely improve the individual's health at least in the short term. Conversely, after an individual has already consumed enough nutrients such that the marginal value of

change in dietary composition should lead to an increase in body weight, all other things equal.[4] Below, we examine the reduced form relationship between relative food prices and BMI, taking the intermediate relationship between relative food prices and consumption patterns as given based on the law of demand.

A second hypothesis involves heterogeneity in how individuals evaluate the health costs of different foods. Specifically, we should observe that the sensitivity of an individual's consumption choices to relative nominal prices should be muted by the expected health costs associated with each food type. That is, people who are well aware of the health benefits of healthful foods and the health costs of unhealthful foods (or who are more sensitive to the health characteristics of the food groups because of low discount rates) may exhibit less sensitivity to prices in comparison to individuals who are either unaware of the health effects of foods (or who heavily discount the future health consequences of current consumption decisions). Effectively, for individuals who weigh the health aspects of the food itself particularly heavily, a (relative) price change will have a smaller relative effect on the total cost of consuming the food. To examine this implication, we estimate the relationship between BMI and relative food prices separately by education level, since those with more education may be more aware of the non-price costs entailed in the consumption of unhealthful foods.[5]

---

additional nutrients is zero, additional healthful foods may merely provide unnecessary calories that serve to increase the individual's body weight.

[4] For an exposition of the relationship between food consumption and BMI, see Cutler, Glaeser, and Shapiro (2003), as well as Lakdawalla and Philipson (2002).

[5] We emphasize that there are other possible explanations besides this one for any observed differences in price sensitivity, since educational attainment is not exogenously assigned.

4. THE PATTERN OF FOOD PRICES IN THE U.S.

As suggested by Philipson and Posner (2003) and shown in Lakdawalla and Philipson (2002), the price of food in the U.S. generally rose less quickly than the overall price level during the second half of the twentieth century. During that time frame, then, people should have substituted away from non-food consumption toward food consumption, due to the law of demand. Lakdawalla and Philipson (2002) find support for this implication, finding that 40 percent of the increase in BMI during the period can be attributed to declining food prices.

However, while for much of this period the prices of healthful and unhealthful foods did not significantly differ, food price data suggest that there was a large divergence in the mid-1980s. As shown in Figure 1, the prices of many healthful foods, such as fresh fruits and vegetables and fish, started to grow much more rapidly than the prices of some unhealthful foods such as ground beef and carbonated beverages.

Cutler, Glaeser and Shapiro (2003) report that obesity nearly doubled between the 1971-75 and 1988-94 periods, while Lakdawalla, Philipson and Bhattacharya (2005) cite Costa and Steckel's (1995) finding that the increase in body weight dates back 150 years. Nevertheless, the mid-1980s divergence in prices of unhealthful and healthful foods that Figure 1 documents provides an interesting source of variation to test basic demand theory's predictions regarding *relative* food prices and individual weight.[6]

To analyze this hypothesis more rigorously, we collected data on individual food prices from the Bureau of Labor Statistics for the period 1980-2003. We then created

[6] Further, note that the mid 1980s is the period public health organizations, like the CDC, point to as the beginning of the obesity epidemic in the U.S. See, for example, http://www.cdc.gov/nccdphp/dnpa/obesity/trend/maps/index.htm .

food indexes categorizing foods as healthful and unhealthful as indicated in Table 1[7] for those foods for which BLS collected prices in each of the four Census regions for the entire period.[8] To net out the effect of changes in the general price level, we focus primarily on the ratio of the price of healthful foods to the price of unhealthful foods. The pattern of relative prices during the time period is represented in Figure 2.

As depicted in Figure 2, we find that the general trend in the ratio of prices of healthful and unhealthful foods has been upward. Further, there appears to have been an especially steep rise in relative prices from the late 1980s through the mid 1990s. It was during this time that obesity began to reach epidemic proportions in the U.S. (Flegal, Carroll, Ogden, and Johnson 2002).

5. RELATIVE PRICE EFFECTS ON OBESITY

We combine the food price data at the regional level with micro data from the National Health Interview Survey (NHIS) for the years 1982-1996[9] to examine the sensitivity of weight measures to relative food prices, focusing on functions of body mass

[7] We recognize that the availability of food price data is incomplete, but, if our instrumental variables approach implemented below is reliable, it will remedy any bias created by measurement error in the relative price index.

[8] Our construction of the indexes weights each food equally. While some more sophisticated weighting scheme might be optimal, it is unclear what weights should be used, so we stick with an equal-weighting approach.

[9] We used 1996 as the ending date in our sample because after this point the NHIS started using a different method of collecting data on height and weight. Prior to 1997, the individual answering the NHIS questionnaire provided information about height and weight for each person residing in the household. Beginning in 1997, only the respondent's height and weight were used, leading to a large decline in the usable sample. In preliminary analysis, we found that this change in reporting practice led to significant changes in reported height and weight (presumably because respondents are not representative of the population as a whole). Though our primary results are robust to including data from 1997-2003, we view the later period data as suspect and so choose not to report results from specifications including the later data.

index (BMI)[10] to measure weight. While BMI has been criticized for a variety of reasons, it is the only measure available from the NHIS and it is the measure typically used when developing public health policies regarding obesity. Figure 3 plots average BMI in our sample from 1982-97, while Figures 4 and 5 plot the share of the population who are overweight and obese, respectively, in each year of the same period.

We consider three outcome variables: BMI, a dummy variable indicating that a person is overweight, i.e., ($BMI \geq 25$), and a dummy variable indicating that a person is obese, i.e., ($BMI \geq 30$). For the overweight and obesity outcomes, we present results from linear probability models, though we find similar results if we use probit models. Our key regressor is the relative price of healthful foods, defined as the ratio of our healthful foods price index to our unhealthful foods price index. We also include the net consumer price index (net CPI), which we calculate after removing the prices that appear in our healthful and unhealthful food price indices as well as the price of unleaded gasoline, which we later use as an instrument for our relative food price measure. For each outcome variable, we report coefficient estimates from two sets of models. The first includes all price variables in levels, and the second includes all price variables in logs.

Before discussing our covariates, we note that one should expect BMI in year *t* to equal BMI in year *t-1* plus the net addition to BMI. The net addition to BMI will depend on many factors, most prominently the year-*t* level and composition of caloric intake and exercise. The body's metabolic state will also be important.[11] Unfortunately, we do not have longitudinal data, so we cannot include lagged BMI. Thus we rely on our

---

[10] $\text{BMI} = 703 \cdot \dfrac{\text{weight}\left(\text{in pounds}\right)}{\text{height}\left(\text{in inches}\right)^2}$

[11] For a discussion of these issues, see Cutler, Glaeser and Shapiro (2003).

demographic and region-level covariates to account for systematic variation in lagged BMI. One can thus regard our specifications as a reduced form for the true underlying model that relates contemporaneous BMI to contemporaneous prices and lagged BMI.[12]

We include data only on black and white individuals because other races were grouped together in early years of the survey, limiting our ability to control for racial and ethnic differences in body types. For covariates, we use a quadratic in age, since people generally gain weight as they grow older, with this process reversing itself after some point. In all specifications, we also include a dummy variable indicating whether a person's income exceeds $20,000 per year,[13] regional fixed effects, and region-specific linear trends.[14]

In pooled specifications, we also include several other covariates. First, we include two education dummies: one indicating exactly 12 years of education, the other indicating more than 12 years of education (thus high school dropouts are the reference group). These variables allow for the possibility that more highly educated people weigh the health effects of food consumption differently from poorly educated individuals, as well as any other source of heterogeneity across educational attainment. Second, we include race-by-sex dummies, allowing for separate intercepts for black women, black

---

[12] To our knowledge, other authors studying the economic determinants of growth in BMI also have focused only on contemporaneous economic variables.

[13] This coarse variable is the only consistent income measure that can be calculated throughout our period. While more precise measures are available in each of the NHIS years, inconsistencies in the year-to-year coding make it impossible to construct a finer, consistent income variable.

[14] We did try to estimate specifications with year dummies. Unfortunately, we ran into identification problems with these specifications. For the IV specifications, coefficient estimates typically could not be identified for both region dummies and year dummies. For the OLS specifications, the estimated cluster-robust covariance matrices were frequently plagued by singularities (even though the coefficients could be separately identified--this distinction arises because of the nature of the cluster-robust covariance estimator). As a result, we are not confident in inference based on these results (indeed, without ignoring dependence, we cannot even test appropriately for exclusion of year dummies). Instead, we simply rely on region-specific linear trends.

men, and white women (white men are the reference group). Summary statistics are provided in Table 2.

We present standard errors that allow for clustering at the region level as suggested by Bertrand, Duflo, and Mullainathan (2004). However, their results suggest the method performs poorly when the number of clusters is small (i.e., n < 30); in our case, we have only four clusters. To address this difficulty, we employ a modified version of the method proposed by Cameron, Gelbach, and Miller (2006), presenting the upper-tailed *p*-values (relative to the null of no effect) generated by that method.[15] We discuss the implementation of this method in section 6.

5.1 Pooled Results

In Table 3, we present the relative price coefficients[16] from our primary specifications, which pool the relative price effect for all people in the sample. Consider first the OLS regressions with price variables entered in levels. The coefficients in the first column suggest that a one-unit increase in the relative price of healthful foods is associated with greater average BMI by 0.764 points, a greater overweight share of 8.7 percentage points, and a greater obese share of 3.9 percentage points. To more precisely isolate our relative price effect from the Lakdawalla and Philipson total price effect, we provide specifications in Appendix Table 3 which further include a control for total food

[15] We use upper-tailed *p*-values because there is no plausible theoretical reason to believe that higher relative food prices could lead to a decrease in BMI. However, for readers interested in the two sided test, the p-values presented can simply be doubled (under the assumption that the null distribution of the test statistic is symmetric).

[16] Results for the other covariates are available in Appendix Table 1. We do not provide bootstrap *p*-values for the other covariates.

prices in a region. The results from these specifications are largely unchanged from those presented here, though our estimates are generally more precise.

Before we discuss the magnitudes of these effects, we first address their statistical significance. The numbers in parentheses are conventional standard errors, estimated with region-level clustering. Using these standard errors and the standard normal approximation to the *T*-ratio's distribution, each of the effects is statistically significant relative to the null of zero effect. However, it is unlikely that the normal approximation is appropriate with only four clusters. Thus, in brackets below the standard errors in Table 3, we report estimated upper-tailed *p*-values (relative to the null hypothesis of zero effect) from a more conservative approach. These *p*-values are 0.029 for BMI, 0.012 for overweight, and 0.04 for obese. We discuss the wild cluster bootstrap-*T* method we used to estimate these *p*-values in detail below; for now we simply conclude that our coefficient estimates are clearly statistically different from zero.

Next we consider the magnitude of the coefficient estimates. Over the period from 1982-96, BMI increased by 1.5 points, and the overweight and obese shares increased by 13.5 and 7.8 percentage points, respectively. By comparison to these changes, our estimates may appear enormous. However, it is important to recall from Table 2 that the standard deviation of the relative price variable is just 0.06. Thus, a one standard deviation increase in the relative price variable is associated with an increase in BMI of only about 0.05, and increases in the respective share of people who are overweight and obese of only about 0.52 and 0.23 percentage points, respectively.

Alternatively, consider the relative price variable's range, which is [0.34, 0.66]. Thus an increase from the least observed relative price level to the greatest would involve an increase of 0.32 in the relative price variable. Based on the results in column 1 of Table 3, this increase would be associated with an increase in BMI of about a quarter point and an increase in the respective overweight and obese shares of 2.8 and 1.2 percentage points. Thus even an increase in the relative price of healthful foods equal to its entire range over our period of observation could account for at most about one-fifth of the increase in our measures of weight. In sum, we do not regard these results as particularly large—a point to which we return in our policy discussion below.

One potential concern with our results is that the relative price measure is jointly determined with BMI. Essentially, we are regressing a measure of observed quantity on price. To examine this endogeneity, we report results from instrumental variables (IV) specifications in which we use the regional price of one gallon of unleaded gasoline to instrument for the relative price variable. We hypothesize that fuel prices comprise a relatively large portion of the cost of healthful foods (such as fruits and vegetables) and are largely governed by world commodity markets, state and local taxes, and other cost shifters. Further, it seems unlikely that gasoline prices have a direct effect on individual health metrics. One can imagine stories in which gasoline prices directly affect health, such as the possibility that high gasoline prices induce individuals to substitute toward more physically strenuous modes of transportation (e.g., walking), but the fact that gasoline expenditures make up a relatively small portion of most people's expenditures coupled with the fact that gas prices in our sample differ by only $0.37 per gallon from the lowest point to the highest point suggests that this effect would seem to be trivial.

We find that our relative price measure is positively correlated with the price of unleaded gasoline and the effect is statistically significant.[17] The IV coefficient estimates in column 2 of Table 3 are about two-thirds the size of the OLS estimates, and their standard errors are (predictably) greater than the OLS estimates. However, Hausman tests based on the region-clustered standard errors clearly do not reject the null of exogeneity.[18] In any case, since the column 2 IV estimates are smaller than the column 1 OLS estimates, we can regard the—small—estimates in column 1 as upper bounds.

In column 3 of Table 3, we report estimated coefficients from specifications in which all price variables appear in logs rather than levels. Since the mean relative price level over our sample is 0.48, dividing the column 3 log-specification coefficients by 0.48 yields a basis for comparison to the estimates in column 1. Since the estimates in column 3 are roughly one-half those in column 1, we see that there is little difference across level-log specifications in the substantive results. We note also that the wild cluster bootstrap-*T* *p*-values in column 3 generally are similar to those in column 1. Finally, we note that the IV estimates in column 4—in which we enter all price variables in logs, including the instrumental variable (the price of unleaded gasoline)—are generally similar to those in column 3.

In sum, based on Table 3's results, we conclude that increases in the relative price of healthful food cause a statistical significant increase in BMI and the share of the population who are overweight and obese. These estimates do not appear to be seriously plagued by endogeneity. However, statistical significance is obviously a different concept from economic significance, and we find that changes in the relative price of foods can

[17] First stage results for the instrument are available in Appendix Table 2.
[18] Lakdawalla and Philipson (2002) argue that their results suggest that it is appropriate to treat food prices as exogenous, an approach taken by Lakdawalla, Philipson and Bhattacharya (2005).

explain at most a very small amount of the population-wide increase in BMI, overweight and obesity.

5.2 Effects by Education Level

As we discussed above, people who are better informed about the health costs of unhealthful food may view food prices as a less important part of the total food cost Seemingly, more educated individuals will be more aware of the health costs of unhealthful foods. Further, as suggested by Fuchs (1982), education could proxy for discount rates with individuals with low discount rates pursuing more education. Since the health costs of poor eating habits will generally come to bear only in the longer term, individuals with low discount rates will weigh them more heavily in their evaluation of relative costs. All of these hypotheses imply sensitivity to our relative food price measure declines with education. To examine these hypotheses, and heterogeneity more generally, we estimate our models separately by education level.

In Table 4, we present the same specifications used in Table 3, dividing the sample into three groups: people who did not complete high school, people with only a high school education, and people with more than a high school education. In all OLS specifications, we find that individuals with less than a high school education exhibit the greatest sensitivity to changes in relative prices. In the OLS log specifications, people with more than a high school education are less sensitive to relative price changes than those with just a high school education for all three outcome variables. This relationship does not hold in the OLS specifications which include the price variables in levels for

BMI and for the likelihood of being overweight. In general, the magnitudes of the OLS estimates are not substantially different from the population-wide estimates in Table 3.

The IV estimates in Table 4 show that high school dropouts are more sensitive to relative price changes than are people in the other two groups. For most coefficients, the log-specification coefficients in column 4 suggest substantially greater effects than the level-form coefficients in column 2. Nonetheless, the magnitudes of these estimates are not so great as to reverse our conclusion, that variation in relative food prices explains relatively little of the increase in weight outcomes.

5.3 Heterogeneity of Effects Along Other Dimensions

We also examined whether sensitivity to relative price changes differed by sex and race. We present estimates for males and females separately in Table 5. By and large, men and women appear not to differ in terms of price sensitivity.

We report race-specific results in Table 6. Based on the level specifications (whether OLS or IV), blacks' weight does not appear to respond to price variation. By contrast, the log specification results suggest that both BMI and obesity increase with the relative price of healthful foods. Among whites, we find consistent evidence that all three outcomes increase with the relative price variable.

5.4 Summary of Results

As we noted above, the estimated effects are generally small in relation to the increases in BMI, overweight and obesity over the period we examine. Between 1982 and 1996, the average value of our relative price variable increased from 0.473 to

0.544—an increase of 0.071. Using a coefficient of 0.764 for the effect of the relative price on BMI would thus imply an increase in BMI of 0.0568—just 3.6% of the 1.5-unit increase in our sample's average BMI over the same period. Using a coefficient of 0.087 for the effect of relative price on overweight implies an increase of overweight equal to just 0.6 percentage points—or 4.6% of the 13.5 percentage-point increase in overweight incidence. Finally, a coefficient of 0.039 implies an increase of 0.3 percentage points in the incidence of obesity—which is only 3.6% of the 7.8 percentage-point increase in obesity that occurred between 1982-96. Results do vary somewhat by educational attainment and race, but not so much as to change the basic tenor of our findings. We conclude that our estimates suggest statistical significant but economically insubstantial effects.

6. INFERENCE WITH RELATIVELY FEW CLUSTERS

An important issue we need to address involves dependence of residual components of BMI within Census regions. In section 6.1, we discuss well known approaches to robust covariance estimation in situations like ours. In particular, we explain why the small number of Census regions R is problematic for standard inference methods. In section 6.2, we discuss a conventional pairs bootstrap approach for inference based on bootstrapped estimates of the standard error and the Wald statistic, noting the relative advantage of bootstrapping the latter. In section 6.3, we outline the wild cluster bootstrap method that we use to try to address this small R problem. In section 6.4, we provide details of Monte Carlo simulations that indicate our approach is successful in accounting for within region dependence. Our discussion is informal and omits details for the sake of simplicity. Interested readers should consult Cameron, Gelbach, and Miller (2006) and the references cited therein for details.

6.1 Robust Covariance Estimation Methods

Write the relative price model as

$$BMI_{irt} = P_{rt}\gamma + W_{irt}\beta + \varepsilon_{irt} \tag{1}$$

where BMI_{irt} is the body mass index for person *i* living in region *r* at time *t*; P_{rt} is the scalar relative price, in region *r* at time *t*, of healthful foods in terms of the price of unhealthful foods; W_{irt} is a $1 \times K$ vector of covariates that may vary with person, region, and/or year; γ and β are parameters to be estimated; and ε_{irt} is the residual component of BMI.

Define

$$X_{irt} \equiv \left(P_{rt}\ W_{irt}\right) \tag{2}$$

$$\Theta \equiv \begin{pmatrix} \gamma \\ \beta \end{pmatrix} \tag{3}$$

so that we may write the model as

$$BMI_{irt} = X_{irt}\Theta + \varepsilon_{irt} \tag{4}$$

The variance of the centered-scaled OLS estimator of Θ converges to

$$V\left(\sqrt{n}\left[\widehat{\Theta} - \Theta\right]\right) = Q^{-1}E\left(\frac{X'\varepsilon\varepsilon'X}{n}\right)Q^{-1} \tag{5}$$

where $Q \equiv En^{-1}\left(X'X\right)$ and X and ε are the obvious stacked matrix and vector, respectively. Under a standard law of large numbers (LLN), we can assume that

$\left(X'X / n\right)^{-1}$ is a consistent estimator for Q^{-1}. If all residuals were iid, then we could use the "default" estimator, $\sigma^2 (X'X / n)^{-1}$, for the entire matrix in (5), where $\hat{\sigma}^2 = n^{-1}\Sigma_{irt}\hat{\varepsilon}_{irt}^2$, with $\hat{\varepsilon}_{irt}$ being the OLS estimated residual. This estimator works because with iid residuals, $\hat{\sigma}$ is consistent for $\left(E\left[\varepsilon_{irt}^2\right]\right)^{1/2}$, the standard deviation of true residuals, and since the "meat" of the sandwich matrix on the right hand side equals $\sigma^2 Q$ under iid residuals. Supposing instead that there is heteroskedasticity, but not dependence, we instead use

$$\widehat{M}^{HW} \equiv \frac{1}{N-1}\sum_{irt} X'_{irt} X_{irt} \hat{\varepsilon}_{irt} \hat{\varepsilon}_{irt} \tag{6}$$

where N is the total sample size, to estimate the middle matrix. This is the usual Huber-White estimator; plugging it in to (5) yields a consistent estimator for the variance provided that there is no dependence. Dependence occurs when some residuals are correlated. In this case, $\widehat{M}^{HW}$ is biased since it omits consideration of non-zero expectation terms of the form

$$m_{irtjr't'} \equiv X'_{irt} X_{jr't'} \varepsilon_{irt} \varepsilon_{jr't'} \tag{7}$$

for which at least one of $i \neq j$, $r \neq r'$, or $t \neq t'$ holds. As has been discussed extensively in the literature (e.g., see Moulton (1990), Bertrand, Duflo, and Mullainathan (2004), and CGM), such terms enter when there is correlation across such observations in <u>both</u> the regressors and the residuals within clusters. Roughly speaking, the magnitude of bias due to using $\widehat{M}^{HW}$ when there is dependence is increasing in the magnitude of both regressor and residual correlation. Because our price data vary only at the region-by-time level, P_{rt}

is perfectly correlated within region-year cells. Moreover, ours is a parsimonious empirical model, and tastes and other determinants of eating behavior differ regionally, so we do not think it appropriate to ignore the possibility of dependence. The usual "cluster" robust middle matrix estimator that allows for dependence within region-by-year cells is

$$\widehat{M}^{RT} \equiv \frac{1}{RT-1}\sum_{rt} X'_{rt}\hat{\varepsilon}_{rt}\hat{\varepsilon}'_{rt}X_{rt} \tag{8}$$

where there are R regions and T years of data, and X_{rt} is the matrix of stacked regressor observations on all individuals in region r at time t; letting n_{rt} be the number of individuals in this cell, X_{rt} has dimension $n_{rt} \times K + \dim P$ (whereas each X_{irt} has just one row). The column vector ε_{rt} has n_{rt} rows and contains all residual values for the region-year cell. If we assume that there is no other source of dependence, then we can treat $X'_{irt}\varepsilon_{rt}$ as random vectors that are independent across region-year. The same arguments that justify $\widehat{M}^{HW}$ in the heteroskedasticity-only case also justify $\widehat{M}^{RT}$ in the case of dependence occurring only within region-year cells. Use of this matrix is justified when we can correctly assume that terms in (7) have expectation zero whenever either $r \neq r'$ or $t \neq t'$.

Next we observe that price data are likely to be correlated over time within region, as the empirical autocorrelations of our price data suggest. If residual determinants of BMI also vary systematically over time within region, then we have dependence at the region level, not just the region-year level. That is, we cannot assume zero expectation for terms in (7) unless $r \neq r'$; terms from the same region that involve

$t \neq t'$ may be nonzero. The cluster-robust middle matrix estimate that accounts for this within-region dependence is

$$\widehat{M}^R \equiv \frac{1}{R-1} \sum_r X_r' \hat{\varepsilon}_r \widehat{\varepsilon}_r' X_r \tag{9}$$

Observe that the data matrix X_r is $n_r \times K + \dim P$, while ε_r is $n_r \times 1$.

The key difference between $\widehat{M}^R$ and $\widehat{M}^{RT}$ is that $\widehat{M}^r$ includes the observed values of all terms in (7) that have $r = r'$, so that the covariance of $X_{irt}\varepsilon_{irt}$ terms from the same region but different years are no longer restricted to be 0.

Let $\bar{n}$ be the average number of observations per region-year cell, let T be the number of years of data, and let R be the number of regions. Consistency of the Huber-White estimator $\widehat{M}^{HW}$ is based on asymptotic arguments involving the total number of observations, $N = \bar{n}RT$ which is more than one million in our case. Consistency of the region-year cluster robust estimator $\widehat{M}^{RT}$ is based on the product *RT*; at 60, this product is likely sufficient to justify use of inference based on first-order asymptotic theory. Consistency of the region cluster robust estimator $\widehat{M}^R$ is based on the growth of *R*. Unfortunately, in our case *R* is only four, since our price variable varies only at the Census region level. Thus there is little reason to trust inference based on using $\widehat{V}^R \equiv (X'X/n)^{-1} \widehat{M}^R (X'X/n)^{-1}$ to construct a statistic that will then be compared to critical values of the standard normal distribution. We now turn to consideration of bootstrap-based alternatives.

6.2 Bootstrapping: Standard Error or Wald Statistic?

The standard approach to inference is to (*i*) determine the null hypothesis, (*ii*) calculate a test statistic whose distribution is known under the null hypothesis, and (*iii*) find critical values of this distribution that determine a region such that the null hypothesis should be rejected when the test statistic falls outside this region. The corresponding first-order asymptotic theory approach typically used in a regression setting with regional clustering is to specify $H_0 : \gamma = \gamma_0$ for some interesting value γ_0 and then define the test statistic

$$T_N \equiv \frac{\hat{\gamma} - \gamma_0}{\hat{\sigma}^R} \tag{10}$$

where $\hat{\gamma}$ is the OLS estimate of γ, γ_0 is the value (e.g., 0) that γ is assumed to take under H_0, and $\hat{\sigma}^R$ is the estimated standard error of $\hat{\gamma}$ found by taking the square root of the appropriate diagonal element of $\hat{V}^R$. While common parlance is to refer to T_N as the t-statistic, it is more precise to call it a Wald statistic, since its distribution is *t* only when the residuals are truly normally distributed. As usual, though, under weak conditions on the data generating process, T_N's distribution converges to the standard normal as *R* grows. As such, an upper-tailed, asymptotically consistent level-α test of H_0 can be based on whether $|T_n| > z_{1-\alpha}$, the $1-\alpha$ quantile of the standard normal distribution; if so, then we reject H_0. However, as noted above, consistency of $\hat{\sigma}^R$ relies on asymptotics with respect to *R*. Since *R* is small here, higher order terms in the Edgeworth expansion of T_N's distribution will be more important, and thus there is no guarantee that the first-order test just described is accurate.

In a recent paper CGM argue that previously known theoretical results imply that appropriately constructed bootstrap-based tests can greatly improve inferential accuracy in the small cluster case (see the discussion in CGM and citations to Horowitz (2001) and Hall (1992)). In their Monte Carlo results using the wild cluster bootstrap-*T,* CGM's tests have correct coverage rates with as few as five clusters. That is, their Monte Carlo bootstraps reject a true null (within simulation error of) five percent of the time when using level 0.05 tests. Here we give a very brief discussion of a similar approach to theirs (the key differences are that (i) we use upper-tailed rather than two-sided tests, (ii) we use different wild-bootstrap weights than they do, something we discuss below, and (iii) we use the method with both OLS and IV estimators).

The basic idea of the bootstrap is that, since the sample in hand – the "real" data – is representative of the underlying population, we can treat the sample distribution as if it were the population distribution. Thus we consider the sampling distribution of parameters of interest, here T_N, based on the set of all possible re-samples from the real sample. Typically in bootstrap applications, one draws *B* re-samples from the real sample, with each re-sample having the same number of observations as the real sample. Using a random number generator to draw these re-samples makes this type of bootstrap a Monte Carlo procedure. A common approach to obtaining the b^{th} re-sample is to re-sample pairs of (Y, X) observations, with replacement from the real sample, placing probability $1/N$ on drawing any given observation for the size *N* re-sample; the i^{th} re-sampled observation from the b^{th} re-sample may be called (Y_i^{b*}, X_i^{b*}). Once one has drawn N observations in this fashion, one re-estimates the model of interest using the size N re-sample.

At this stage of the procedure, many researchers simply calculate the regression coefficient of interest, in this case γ_b^*, based on the regression of Y^{b*} on X^{b*}, which are respectively the vector of N re-sampled observations on the dependent variable and the $N \times \dim(\theta)$ matrix of re-sampled observations on the price variable and the K controls, drawn for the b^{th} re-sample. Repeating the procedure B times yields the empirical bootstrap distribution $\{\gamma_b^*\}_{b=1}^{B}$ for γ^*. A common approach is the following procedure.

Procedure 1 (Bootstrapping the standard error)

1. Calculate $\widetilde{\sigma}_B$, the standard deviation of γ^* based on the empirical bootstrap distribution $\{\gamma_b^*\}_{b=1}^{B}$, i.e., $\widetilde{\sigma}_B = \sqrt{\frac{1}{B}\Sigma_{b=1}^{B}\left(\gamma_b^* - \hat{\gamma}\right)^2}$
2. Calculate the null Wald statistic $\widetilde{T}_B = \left(\hat{\gamma} - \gamma_0\right) / \widetilde{\sigma}_B$, where the null is $H_0 : \gamma = \gamma_0$.
3. Reject the null if $\left|\widetilde{T}_B\right| > z_{1-\alpha}$.

For reasons too technical to articulate here, the accuracy of this procedure does not improve on the standard first-order asymptotic theory approach discussed above.[19]

A bootstrap-based approach that *can* improve on the standard one is to bootstrap the Wald statistic directly. That is, instead of simply calculating γ_b^* on each iteration of the bootstrap procedure, one calculates

[19] The issue here is that $\widetilde{\sigma}_B$ is not asymptotically pivotal; an asymptotically pivotal statistic is one whose distribution does not depend on unknown parameters. See CGM for a detailed discussion of this point and why it matters.

$$T_b^* \equiv \frac{\gamma_b^* - \gamma_0}{\sigma_b^*} \tag{11}$$

where σ_b^* is a consistent estimate of the standard error of γ_b^* based on estimation results from the b^{th} re-sample. Notice that σ_b^* changes each time we generate a new bootstrap re-sample, so that T_b^* does as well. Unlike the first-order asymptotic theory and bootstrap standard error approaches described above, with this bootstrap *T* approach we do not use critical values based on the standard normal distribution. Rather, we sort the realized values observed in the empirical distribution of Wald statistics, $\{T_b^*\}_{b=1}^{B}$, so that

$$T_{[1]}^* \le T_{[2]}^* \le \ldots \le T_{[B]}^*, \tag{12}$$

where $T_{[c]}^*$ is the c^{th} order statistic of the bootstrap empirical distribution. We then reject the null hypothesis against an upper-tailed alternative if the absolute value of the real-data Wald statistic $\hat{T}$ exceeds the $1-\alpha$ quantile of the empirical bootstrap distribution. That is, we reject if

$$\hat{T} > T_{[(B+1)\alpha]}^* \tag{13}$$

Observe that for $B = 999$, the relevant order statistic of the bootstrap distribution of Wald statistics is the $1000\alpha^{th}$ order statistic, so that for $\alpha = 0.05$, we compare $\hat{T}$ to the 950^{th} order statistics of the empirical bootstrap distribution of Wald statistics. Because the Wald statistic's asymptotic distribution is known to be the standard normal distribution, it is an asymptotically pivotal statistic (again, see CGM for more). As such, when the sample size is small, tests based on critical values estimated from its empirical bootstrap distribution can be much more accurate than tests using critical values from the standard normal distribution. Thus, our focus will be on using a bootstrap of the Wald statistic to determine critical values rather than on a method that uses standard normal distribution critical values.

6.3 Using the Wild Cluster Bootstrap to Address Small R

We need to distinguish our bootstrap further, however. The re-sampling procedure discussed above, which is called the pairs bootstrap, does not deal with the dependence problem that we discuss above. In forming the bootstrap distribution of a statistic, one must replicate the dependence properties of the underlying data generating process. Thus, when there is within-region dependence of the form we suspect here, we cannot simply re-sample observations with equal probability $1/N$. Rather, when we sample observation i from region r and year t, we must also sample all other observations in region r, including not just those present for year t, but also those present in other years. Intuitively, always including all observations within a region preserves the structure of any dependence within region. Thus we have independent data at the regional level, which is to say that we have independent observations on four random vectors in this case.

One approach to using the bootstrap would be to randomly and with replacement draw four clusters in each of B re-samples, with each region having probability one fourth of selection on each re-sample draw. One could then use the procedure described above. It can be shown that there are only 35 possible samples that can be generated by with replacement sampling. Of these, the ones that involve drawing a single cluster four times cannot be used (since the regressor matrix would then be singular), so there are really only 31 useable samples. In this situation, there would be no need to use Monte Carlo methods; we could simply enumerate these samples and calculate γ^* for each of them. Its sampling distribution based on this non-parametric cluster bootstrap is then determined by these 31 values.

In practice, there are surely more than 12 possible values that $\hat{\gamma}$ may take on, so the nonparametric cluster bootstrap does not provide an appropriate basis for inference. Thus we turn to an alternative approach, which CGM call the wild cluster bootstrap. The following is a general procedure for using the wild cluster bootstrap, in which the null hypothesis $H_0 : \gamma = \gamma_0$ is imposed at each step.

Procedure 2 (Wild Cluster Bootstrap with H_0 Imposed)

1. Estimate by OLS the regression of Y_{irt} on $X_{irt} = (P_{rt}, W_{irt})$. Let $\hat{\gamma}$ be the resulting estimate of the coefficient on the price variable, P_{rt}, and form the real-data Wald statistic $\hat{T} = \left(\hat{\gamma} - \gamma_0\right) / \hat{\sigma}^R$

2. Estimate by OLS the regression of $Y_{irt} - P_{rt}\gamma_0$ on the controls W_{irt}; call the resulting coefficient estimate $\hat{\beta}^H$ (where *H* signifies that we have imposed the null hypothesis). Let $\hat{\varepsilon}_{irt}^H$ be the estimated residual from this regression for person i in region *r* at time *t*.

3. Do the following *B* times:

    a. Generate *R* = 4 independent realizations of a random variable *U*, $\{U_1, U_2, U_3, U_4\}$ with the distribution of *U* having the properties that (i) its expectation is zero, (ii) its variance is one, and (iii) *U* is independent of all variables in the real data. We refer to *U* as "the auxiliary random variable". Conventional choices for *U's* distribution are the Rademacher and Mammen distributions (see CGM for details). Each of these is a four-point distribution, so that the problem with the nonparametric bootstrap discussed above remains. Instead of these distributions, we follow a suggestion by Liu (1988) and use a product of normal random variables whose mean is zero and variance is one.[20] This distributional choice for U results in a continuous distribution of the bootstrap coefficient estimator γ_b^*, solving the finiteness problem described above.

b. For each observation i in the r^{th} region and the t^{th} year, generate the following re-sampled realization of the dependent variable:

$$Y_{irt}^* = P_r \gamma_0 + W_{irt} \hat{\beta}^H + U_r \hat{\varepsilon}_{irt}^H \tag{15}$$

We note that Y_{irt}^* has three components. The first component is due to the price variable and is zero under the null given that $\gamma_0=0$. Thus our synthetic random variable is independent of the relative price variable by construction. The second component is due to the covariates W_{irt}, and the third is due to the product of the fitted null residual, $\hat{\varepsilon}_{irt}^H$ and the auxiliary random variable U_r. Since the auxiliary random variable has mean 0 and variance 1 and is independent of all other components of Y_{irt}^*, the law of iterated expectations implies that the expected value of the composite residual $U_r \hat{\varepsilon}_{irt}^H$ is 0. Moreover, independence of the auxiliary random variables ensures that for all i' and t', we have $E\left(U_r^2 \hat{\varepsilon}_{irt}^H \hat{\varepsilon}_{i'rt'}^H\right) = E\left(\hat{\varepsilon}_{irt}^H \hat{\varepsilon}_{i'rt'}^H\right)$, while $E\left(U_r U_{r'} \hat{\varepsilon}_{irt}^H \hat{\varepsilon}_{i'r't'}^H\right) = E\left(\hat{\varepsilon}_{irt}^H \hat{\varepsilon}_{i'r't'}^H\right) = 0$, since U_r and $U_{r'}$ are independent for $r \neq r'$ and residuals from different regions are assumed independent. Thus the distribution of the wild bootstrap auxiliary random variable guarantees that the covariance properties of the underlying data are preserved on each bootstrap iteration.

c. For OLS specifications, estimate the simple linear regression of re-sampled realizations of the dependent variable on all observations of $X_{irt} = \left(P_{rt}, W_{irt}\right)$. On the b^{th} iteration, call the resulting coefficient on the price variable γ_b^*. Calculate the cluster robust variance estimate for γ_b^* (as in step 1 above, where we used the real rather than the re-sampled data), σ_b^*, and calculate $T_b^* \equiv \left(\gamma_b^* - \gamma_0\right) / \sigma_b^*$ (as in (11)).

d. For IV specifications, estimate the instrumental variables regression of the vector of re-sampled realizations of the dependent variable on all observations of $X_{irt} = \left(P_{rt}, W_{irt}\right)$, instrumenting for the variable price variable using the price of unleaded gasoline (in levels or logs, as appropriate). On the b^{th} iteration, call the resulting coefficient on the price variable γ_b^*. Calculate the cluster robust variance estimate for γ_b^* (as in step 1 above, where we used the real rather than the re-sampled data), σ_b^*,

[20] We thank Colin Cameron for suggesting this approach.

and calculate $T_b^* \equiv \left(\gamma_b^* - \gamma_0\right) / \sigma_b^*$ (as in (11)). We note that the null value Y_{irt}^* has no endogeneity problems, since we impose $\gamma_0 = 0$ in generating it.

4. For an upper-tailed test, find the (B+1)(1-α)th order statistic of the empirical bootstrap distribution from step 3c or 3d above. Reject the null hypothesis when $\hat{T}$ exceeds this order statistic. Since we use *B=999*, a level-.05 test involves the 950th order statistic of the bootstrap distribution of $\left\{T_b^*\right\}_{b=1}^{B}$, a level-.10 test involves the 900th order statistic, and so on. An estimated *p*-value may be found as the number of replications *b* such that $T_b^* > \hat{T}$; these are the values that we report in our tables.

6.4 Does This Method Work When *R*=4?

It is reasonable to worry that 4 clusters might be insufficient to yield correct inference. Even in the presence of an asymptotic refinement, 4 may be so far from "large" as to render CGM's wild cluster bootstrap-*T* method inappropriate in our application. While CGM's Monte Carlos show excellent performance of this approach with 5 clusters, their data differ from ours. Thus it is important to provide evidence that the method works with data like ours.

To address this concern, we ran several Monte Carlo experiments using our actual data. Each Monte Carlo involves randomly generating synthetic outcome-variable data that (i) share key moment properties with our data and (ii) satisfy our null hypothesis that the price variable has no effect on the synthetic outcome variable. On each Monte Carlo replication, we (i) estimate the replication-specific OLS or IV price coefficient and (ii) use CGM's wild cluster bootstrap-*T* routine to estimate the *p*-value for that Monte Carlo replication. We then count the number of Monte Carlo replications for which the estimated *p*-value is less than or equal to 0.05, which we call the "number of Monte Carlo rejections". Up to simulation error, this number should equal 0.05 times the number of Monte Carlo repetitions—significant deviations from this level would suggest the wild cluster bootstrap-*T* method does not work well with data like ours.

The only part of the above procedure that bears elaboration is the step in which we generate the synthetic outcome-variable data. To do this, we create

$$Y_{irt}^{m} = P_r \gamma_0 + W_{irt} \hat{\beta}^H + L_r^m \hat{\varepsilon}_{irt}^H , \tag{16}$$

where L_r^m is a random variable drawn for region-*r* on the m^{th} Monte Carlo iteration. We use the same distribution for each L_r^m that we use for U_r. For the same reason that this distribution "works" for the wild cluster bootstrap, it also works for the Monte Carlo experiment: it generates synthetic data whose residual components have the same first- and second-moment properties as our actual data.[21] This discussion shows that Y_{irt}^{m} is created in a manner exactly analogous to Y_{irt}^{*}.

We carried out this procedure for the OLS level specification of the BMI outcome variable in level form (using 1,000 Monte Carlo replications), and also for the IV level specifications of all three outcome variables (using 250 Monte Carlo replications in each case). We investigate the wild cluster bootstrap-*T*'s performance for only the full population specifications (i.e., those in Table 3).[22]

[21] The product-of-normals distribution that we use has skewness 1 (see Liu 1988). As a consequence, our choice of auxiliary random variable distribution ensures that the third-moment properties of our composite random variables are the same as the underlying data's (this statement holds for both the wild bootstrap routines and the Monte Carlo experiments). This distribution does have considerable kurtosis, however, and our composite random variables will have very fat tails—especially when we use *two* auxiliary random variables, as is the case for each wild cluster bootstrap replication of each Monte Carlo experiment. This approach nonetheless works for a simple intuitive reason: the second-order properties of our estimators, which govern the performance of the wild cluster bootstrap-*T*, depend only on the first through third moments of the synthetic data.

[22] Each Monte Carlo replication involves a "synthetic real-data" regression plus 999 replications, so that we ran 1.75 million regressions to carry out this experiment (750,000 of which involved the more time-consuming IV estimator). With our current computational resources, this process took several weeks to complete.

For the OLS Monte Carlo, a Monte Carlo rejection rate of 0.05 would imply a number of Monte Carlo rejections equal to 50. A 95% confidence interval for the observed number of rejections under the null $H_0 : p = 0.05$ is [38, 64].[23] In fact, our experimental number of Monte Carlo rejections is 59—so that we do not reject this null against a two-sided alternative.[24]

For the IV specifications, a Monte Carlo rejection rate of 0.05 would imply a number of Monte Carlo rejections equal to 12.5 (i.e., 12 or 13). A 95% confidence interval for the observed number of rejections under the null $H_0 : p = 0.05$ is [7, 20]. For the BMI and overweight Monte Carlos, the observed number of rejections was 10 in each case; for the obese Monte Carlo, it was 9.[25] Thus we conclude that the wild cluster bootstrap performs remarkably well with our data, despite the fact that we have just 4 clusters. We regard this demonstration—that CGM's method works well in an application of interest with so few clusters—to be an additional contribution of the present paper.

7. TWINKIE TAXES AND SALAD SUBSIDIES

[23] Let *B(M,k,p)* be the probability of observing *k* successes in *M* Bernoulli trials given a true probability of *p*. Then *B(1000,64,0.05)= 0.028 and B(1000,65,0.05)= 0.021, while 1-B(1000,38,0.05)= .031* and *1-B(1000,37,0.05)= .021.* Thus a (conservative) 95% confidence interval is [38,64].

[24] We ran the OLS Monte Carlo using a somewhat different specification from the one reported in Table 3. The key differences are: (i) we entered the educational attainment variable linearly even though it is a categorical variable, and (ii) we included data after 1996. We discovered the educational attainment and post-1996 data issues discussed above after having completed the very time-consuming Monte Carlo process; given the generally similar results in the Monte Carlo specification and those reported in Table 3, we have not re-run the OLS Monte Carlo experiment. The IV Monte Carlos were all run with the same specifications used in Table 3.

[25] We used the same sequence of random number seeds for the Monte Carlos involving the three IV specifications. This choice ensures that the four auxiliary residuals in the m^{th} Monte Carlo replication are the same for all three dependent variables, so that the three experiments should not be regarded as independent of each other. Of course, for each outcome variable, the 250 Monte Carlo replications are still mutually independent.

Obesity has garnered significant attention from the public health community, and substantial resources have been devoted to developing solutions to the obesity problem. In some regards, as pointed out by Philipson and Posner (2003), obesity may not be as much of a problem as many seem to think, at least in the narrow economic sense. People optimize over many dimensions when making their consumption decisions, and it is presumably reasonable for some individuals to trade off the health benefits of a rigorous diet and exercise regime in order to gain the (repeated) short term satisfaction of eating unhealthful food and sitting idle.

Further, from a public policy standpoint, obesity does not generate the kind of harm with which economists generally concern themselves. That is, the costs of obesity are generally private costs. Obesity does not tend to generate negative externalities, except so-called fiscal externalities, which arise since obese people disproportionately use medical resources which are often subsidized through tax revenues. However, from a policy standpoint, the public may find the fiscal externality problem sufficient to justify public intervention in the obesity epidemic.

Perhaps more controversially, public intervention may be justified on "behavioral economics" grounds. Specifically, the public may decide that paternalistic intervention is required if either (i) it is the case that individuals suffer from "internality" problems, or (ii) the current self naively affects the future self adversely due to weakness-of-will or other non-rational behavioral phenomena. Our results provide context for recent proposal to tax junk food, subsidize healthful foods, or both. Our results show that such taxes and subsidies would change the relative prices people face, leading to a decline in BMI (presumably by inducing people to shift dietary composition toward more healthful food).

However, the small magnitudes of our estimates cast doubt on the efficacy of such taxes and subsidies, at least in reasonable policy ranges. For example, consider a 100 percent tax on all unhealthful foods in our price index—surely an upper bound on what is politically viable. Estimated at the sample averages, the effect on our relative price variable would be to lower it from 0.48 to 0.24, generating a reduction of 0.24. Multiplying this difference by our estimated effect of 0.764 leads to an implied reduction of BMI of less than 0.2 points—which is less than 1 percent of average BMI (and just 13% of the increase in BMI that occurred from 1982-96 alone). Similar calculations for the likelihood of being overweight and obese are 2 percent and 1 percent respectively. These seem to us like small effects, and they were generated using a tax rate that is likely to be politically infeasible.[26]

8. CONCLUSION

Obesity is a growing concern. By the mid-1980s, public health officials noted a substantial increase in the incidence of obesity and warned of the attendant health consequences that would likely follow. Recently, economists have begun studying the phenomenon, noting that a substantial amount of the growth in body weights can be attributed to simple price effects. That is, as the price of food declines, people consume more of it.

We investigate a related price effect, arguing that changes in the price of healthful food relative to unhealthful food may explain some of the short term and regional

[26] Moreover, Lakdawalla, Philipson & Bhattacharya (2005) have suggested that Pigouvian taxes on specific foods that are aimed at counteracting the rise in weight could have negative unintendend consequences. Given results in Lakdawalla & Philipson (2002), an alternative approach that might have a bigger effect on weight would be to tax *all* food. Such a policy would likely be regressive, hitting lower-income people the hardest, and it, too, could have important unintended consequences.

variation in body weights. As healthful foods become more expensive relative to unhealthful foods, we should expect individuals to substitute toward eating relatively more unhealthful foods, leading to increases in body weight.

We test this hypothesis using data from the NHIS from 1982 to 1996, showing that BMI and the price of healthful food relative to unhealthful food are positively correlated, with effect being statistically significant. Using instrumental variables specifications, we show that simultaneity between relative food prices and BMI is unlikely to be driving our results, suggesting that the positive relationship between the ratio of healthful food prices to unhealthful food prices and BMI is causal. Ours is also the first applied paper to use Cameron, Gelbach and Miller's (2006) wild cluster bootstrap-*T* method for inference with clustered price data and a small number of clusters. We present Monte Carlo evidence suggesting that that method works quite well with our data. This is potentially very important, since with so few clusters, many researchers likely would either have ignored clustering or abandoned this research. Our findings suggest that neither approach is desirable.

There is substantial heterogeneity in the sensitivity of individuals to relative price changes by education level, with more-educated individuals exhibiting less sensitivity to changes in relative prices. This is consistent with the notion that more informed individuals will weigh non price elements of the cost of food more heavily in their decisions. We also find some racial heterogeneity of the effect, but we do not find that sensitivity differs by sex.

The effect we identify is relevant to the policy debate over using fat taxes (or subsidies for healthful foods) to counteract trends in obesity. Our results suggest that the sensitivity of individuals to relative food prices is too small for fat taxes to have much of an effect, at least in reasonable ranges of tax rates. For example, a 100 percent tax on unhealthful foods would reduce average BMI by less than 1 percent, according to our results.

References

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan (2004). "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119(1): 249-275.

Costa, Dora and Steckel, Richard (1995). "Long-Term Trends in Health, Welfare, and Economic Growth in the United States." National Bureau of Economic Research (Cambridge, MA) Historical Working Paper No. 76.

Cameron, Colin, Jonah Gelbach, and Douglas Miller (2006). "Bootstrap-Based Improvements for Inference with Clustered Errors." *Unpublished Working Paper*, available at http://www.glue.umd.edu/~gelbach/papers/under_review.html

Chou, Shin-Yi, Michael Grossman, and Henry Saffer (2004). "An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System." *Journal of Health Economics*, 23(3): 565-587.

Cutler, David, Edward Glaeser, and Jesse Shapiro (2003). "Why Have Americans Become More Obese?" *Journal of Economic Perspectives*, 17(3): 93-118.

Finkelstein, Eric, Christopher Ruhm, and Katherine Kosa (2005). "Economic Causes and Consequences of Obesity." *Annual Review of Public Health*,26: 239-257.

Flegal Katherine, Margaret Carroll, Cynthia Ogden, and Clifford Johnson (2002). Prevalence and Trends in Obesity Among U.S. Adults: 1999-2000." *JAMA*, 288(14): 1723-1727.

Fuchs, Victor (1982). "Time Preference and Health: An Exploratory Study." Economic *Aspects of Health*, edited by Victor Fuchs, pp. 93-120. Chicago: University of Chicago Press.

Hall, Peter (1992). *The Bootstrap and Edgeworth Expansion* (New York: Springer-Verlag).

Hedley, Allison, Cynthia Ogden, Clifford Johnson, Margaret Carroll, Lester Curtin, and Katherine Flegal (2004). "Prevalence of Overweight and Obesity Among US Children, Adolescents, and Adults, 1999-2002." *JAMA*: 291(23): 2847-2850.

Horowitz, Joel (2001). "The Bootstrap in Econometrics," in *Handbook of Econometrics, Volume 5, eds.* James J. Heckman and Edward E. Leamer, Elsevier Science, pp. 3159-3228.

Klick, Jonathan, and Thomas Stratmann (2007). "Diabetes Treatment and Moral Hazard." *Journal of Law and Economics*, forthcoming.

Lakdawalla, Darius and Tomas Philipson (2002). "The Growth of Obesity and Technological Change: A Theoretical and Empirical Examination." *NBER Working Paper*: No. W8946.

Lakdawalla, Darius, Tomas Philipson and Jay Bhattacharya (2005). "Welfare-Enhancing Technological Change and the Growth of Obesity," *American Economic Review,* 95(2): 253-257.

Liu, R. Y. (1988). "Bootstrap procedure under some non-I.I.D. models," *Annals of Statistics* 16, 1696–1708.

Ogden, Cynthia, Margaret Carroll, and Katherine Flegal (2003). "Epidemiologic Trends in Overweight and Obesity." *Endocrinology & Metabolism Clinics of North America*, 32(4): 741-760.

Philipson, Tomas and Richard Posner (2003). "The Long-Run Growth in Obesity as a Function of Technological Change." *Perspectives in Biology and Medicine*, 46(3S): S87-S107.

U.S. Department of Health and Human Services (2000). *Healthy People 2010*, vol. II available at http://www.healthypeople.gov/Publications .

Figure 1:
Price of Selected Foods

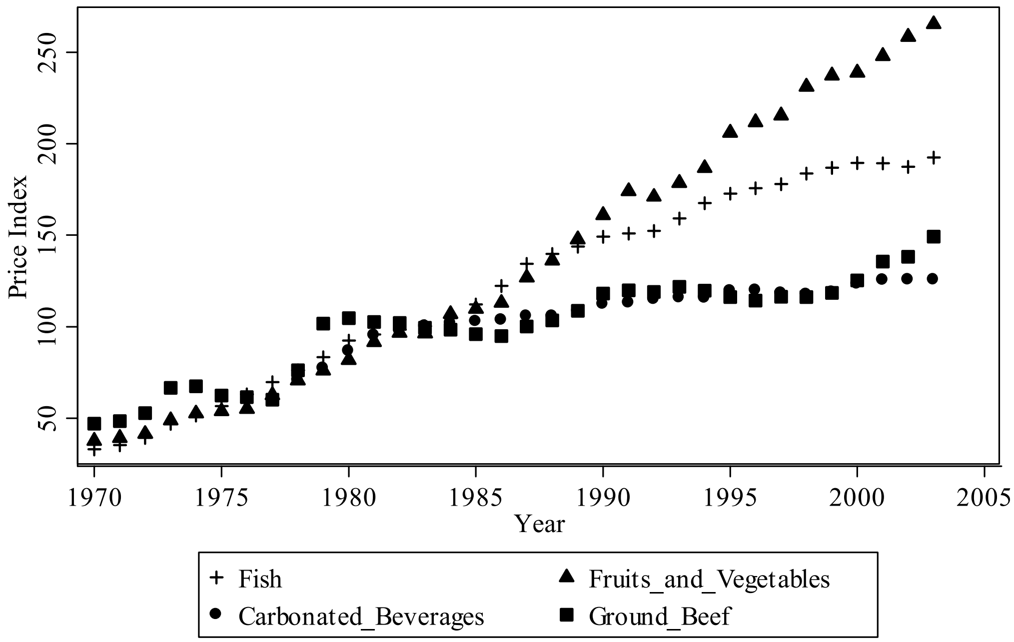

Figure 2:
Price Healthful/Price Unhealthful

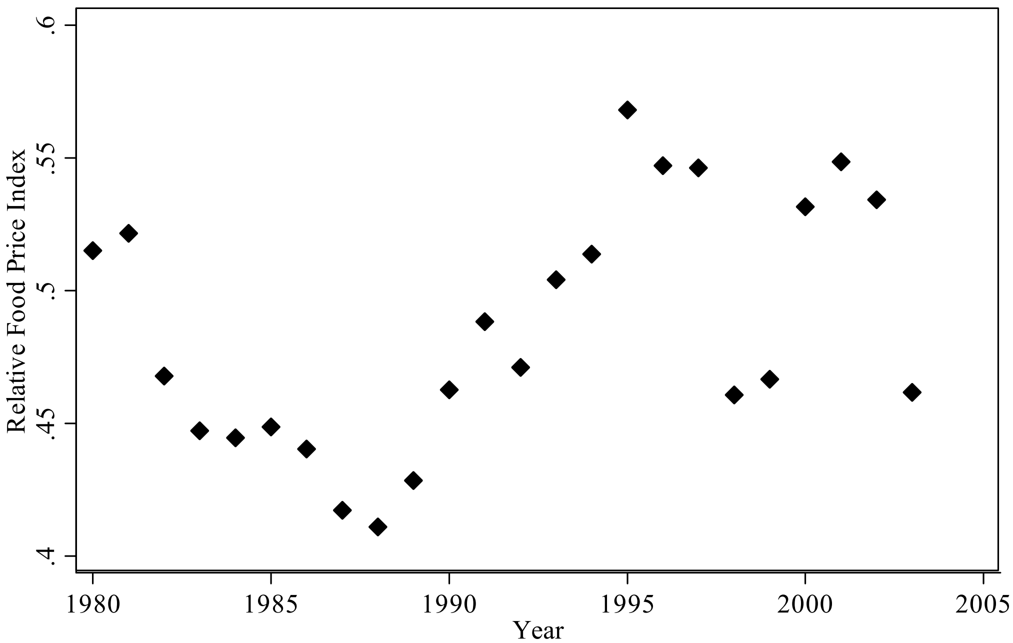

Figure 3:
Average BMI by Year

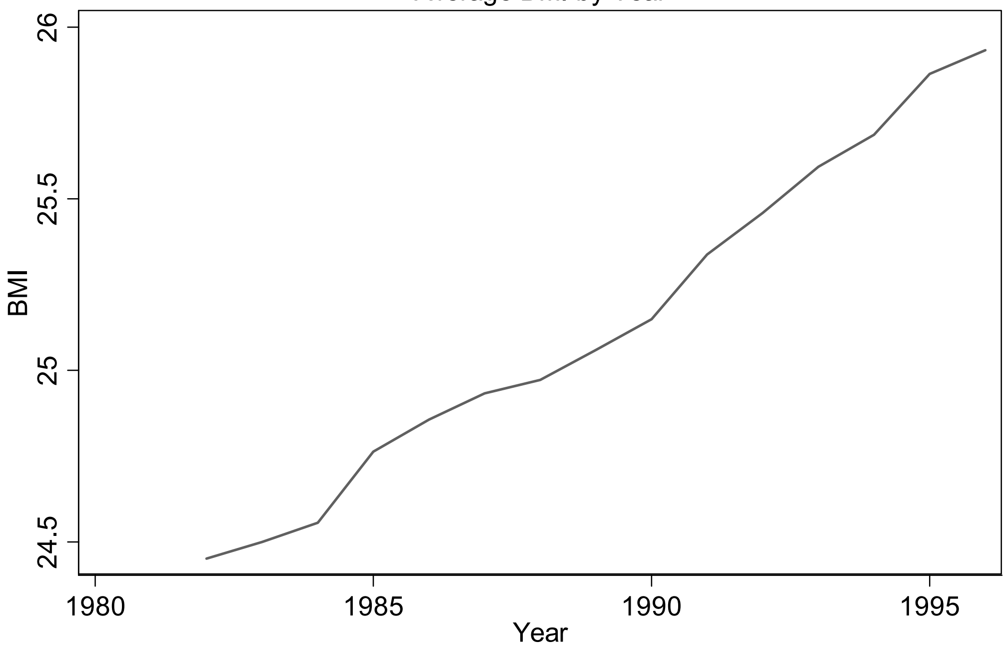

Figure 4:
Percent Overweight by Year

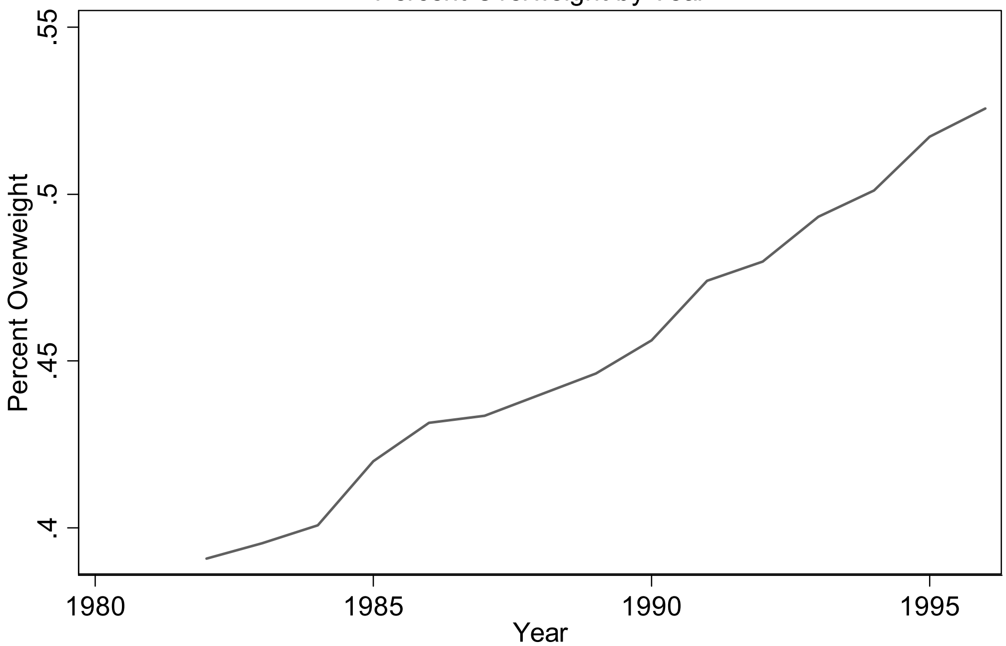

Figure 5:
Percent Obese by Year

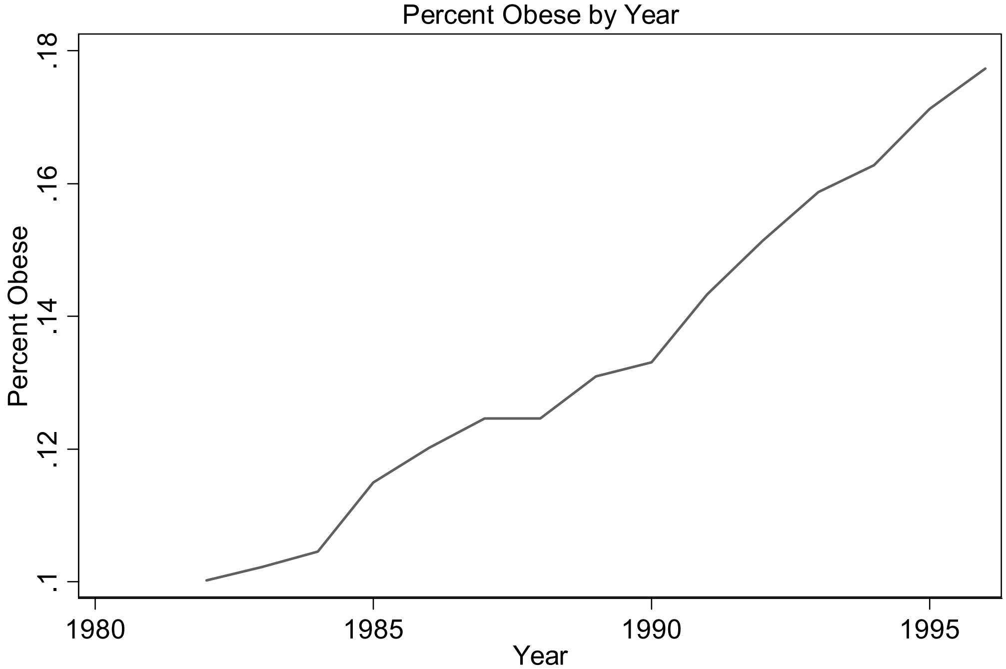

Table 1

Foods Used to Construct Price Indexes	
Healthful Foods	Unhealthful Foods
Apples, Red Delicious, per lb. (453.6 gm)	American processed cheese, per lb. (453.6 gm)
Bananas, per lb. (453.6 gm)	Bologna, all beef or mixed, per lb. (453.6 gm)
Beans, dried, any type, all sizes, per lb. (453.6 gm)	Butter, salted, grade AA, stick, per lb. (453.6 gm)
Beans, green, snap (cost per pound/453.6 grams)	Cola, nondiet, cans, 72 oz. 6 pk., per 16 oz. (473.2 ml)
Broccoli, per lb. (453.6 gm)	Cupcakes, chocolate (cost per pound/453.6 grams)
Cabbage, per lb. (453.6 gm)	Ice cream, prepackaged, bulk, regular, per ½ gal.
Carrots, short trimmed and topped, per lb. (453.6 gm)	Malt beverages, all types, all sizes, per 16 oz.
Celery, per lb. (453.6 gm)	Margarine, stick, per lb. (453.6 gm)
Cherries, per lb. (453.6 gm)	Peanut butter, creamy, all sizes, per lb. (453.6 gm)
Corn on the cob, per lb. (453.6 gm)	Pork and beans, canned (cost per 16 ounces)
Crackers, soda, salted, per lb. (453.6 gm)	Potato chips, per 16 oz.
Cucumbers, per lb. (453.6 gm)	Potatoes, frozen, French fried, per lb. (453.6 gm)
Grapefruit, per lb. (453.6 gm)	Shortening, vegetable oil blends, per lb. (453.6 gm)
Grapes, Emperor or Tokay (cost per 453.6 grams)	
Grapes, Thompson Seedless, per lb. (453.6 gm)	
Jelly (cost per pound/453.6 grams)	
Lemons, per lb. (453.6 gm)	
Lettuce, iceberg, per lb. (453.6 gm)	
Milk, fresh, whole, fortified, per ½ gal. (1.9 lit)	
Mushrooms (cost per pound/453.6 grams)	
Onions, dry yellow, per lb. (453.6 gm)	
Onions, green scallions (cost per pound/453.6 grams)	
Oranges, Navel, per lb. (453.6 gm)	
Peaches, any variety, all sizes, per lb. (453.6 gm)	
Pears, Anjou, per lb. (453.6 gm)	
Peppers, sweet, per lb. (453.6 gm)	
Radishes (cost per pound/453.6 grams)	
Strawberries, dry pint, per 12 oz. (340.2 gm)	
Tomatoes, field grown, per lb. (453.6 gm)	
Tuna, light, chunk, per lb. (453.6 gm)	
Yogurt, natural, fruit flavored, per 8 oz. (226.8 gm)	

Note: All price data come from Bureau of Labor Statistics

Table 2
Summary Statistics

Variable	Description	Mean	SD	Source
BMI	Body Mass Index	25.09	4.76	NHIS
Overweight	= 1 if BMI ≥ 25	0.45	0.50	NHIS
Obese	= 1 if BMI ≥ 30	0.13	0.34	NHIS
Relative Price	Price Healthful/Price Unhealthful	0.48	0.06	BLS
Price Healthful	Average Price of Foods in Column 1 of Table 1	0.82	0.11	BLS
Price Unhealthful	Average Price of Foods in Column 2 of Table 1	1.73	0.18	BLS
CPI	Consumer price index net of food and energy	127.22	23.17	BLS
White Male	= 1 if individual is white and male	0.41	0.49	NHIS
Black Male	= 1 if individual is black and male	0.05	0.23	NHIS
White Female	= 1 if individual is white and female	0.46	0.50	NHIS
Black Female	= 1 if individual is black and female	0.08	0.27	NHIS
High Income	= 1 if annual household income is above $20,000	0.63	0.48	NHIS
Age	Age in years	44.26	17.86	NHIS
High School	= 1 if individual is a high school graduate and has no additional education	0.38	0.49	NHIS
High School Plus	= 1 if individual has more than a high school education	0.38	0.49	NHIS
Unleaded	Price of one gallon of unleaded gasoline	1.13	0.13	BLS
Total Food Price	Index of all food prices deflated by CPI	0.961	0.030	BLS

NHIS: National Health Interview Survey available at http://www.cdc.gov/nchs/nhis.htm
BLS: Bureau of Labor Statistics

Table 3
Effect of Relative Price of Healthful Foods on BMI, Overweight, and Obesity
Pooled Sample
(standard errors clustered by region in parentheses)
[upper-tailed p-values generated by wild cluster bootstrap in brackets]

	Specification of Price Variables			
	Levels		Logs	
	OLS	IV	OLS	IV
BMI	0.764	0.531	0.345	0.372
	(0.277)	(0.473)	(0.138)	(0.282)
	[p = 0.029]	[p = 0.133]	[p = 0.018]	[p = 0.053]
Overweight	0.087	0.052	0.039	0.029
	(0.024)	(0.044)	(0.010)	(0.017)
	[p = 0.020]	[p =0.144]	[p = 0.008]	[p = 0.035]
Obese	0.039	0.025	0.017	0.019
	(0.022)	(0.029)	(0.009)	(0.017)
	[p = 0.040]	[p = 0.192]	[p = 0.038]	[p = 0.098]

Note: All specifications include the following additional controls: CPI (in same form, i.e., level or log, as relative price variable); black male; white female; black female; high income; high school; high school plus; age and its square; three region dummies; and region-specific linear time trends. Results from these coefficients for the price level specifications are available in Appendix Table 1. IV specifications use unleaded to instrument for the relative price variable (in same form, i.e., level or log, as relative price variable) and include all of the covariates in both stages. Results for unleaded from the first stage regression are available in Appendix Table 2.

Table 4
Effect of Relative Price of Healthful Foods on BMI, Overweight, and Obesity
By Education Level
(standard errors clustered by region in parentheses)
[upper-tailed p-values generated by wild cluster bootstrap in brackets]

	Specification of Price Variables			
	Levels		Logs	
	OLS	IV	OLS	IV
Not HS Graduate				
BMI	0.905	0.862	0.576	0.857
	(0.483)	(0.763)	(0.212)	(0.515)
	[p = 0.038]	[p = 0.065]	[p = 0.016]	[p = 0.016]
Overweight	0.108	0.095	0.059	0.065
	(0.059)	(0.080)	(0.027)	(0.046)
	[p = 0.031]	[p = 0.068]	[p = 0.027]	[p = 0.060]
Obese	0.027	0.036	0.027	0.056
	(0.030)	(0.075)	(0.014)	(0.043)
	[p = 0.224]	[p = 0.296]	[p = 0.033]	[p = 0.048]
HS Graduate				
BMI	0.599	0.470	0.278	0.347
	(0.273)	(0.572)	(0.090)	(0.151)
	[p = 0.045]	[p = 0.207]	[p = 0.019]	[p = 0.031]
Overweight	0.058	0.021	0.033	0.027
	(0.017)	(0.069)	(0.006)	(0.019)
	[p = 0.016]	[p = 0.356]	[p = 0.003]	[p = 0.072]
Obese	0.045	0.037	0.019	0.022
	(0.022)	(0.029)	(0.007)	(0.008)
	[p = 0.044]	[p = 0.146]	[p = 0.015]	[p = 0.007]
More than HS				
BMI	0.707	0.223	0.236	0.064
	(0.267)	(0.483)	(0.186)	(0.319)
	[p = 0.022]	[p = 0.326]	[p = 0.105]	[p = 0.398]
Overweight	0.083	0.028	0.027	-0.000
	(0.013)	(0.027)	(0.008)	(0.012)
	[p = 0.002]	[p = 0.191]	[p = 0.013]	[p = 0.487]
Obese	0.038	0.007	0.010	-0.003
	(0.028)	(0.029)	(0.012)	(0.013)
	[p = 0.074]	[p = 0.390]	[p = 0.216]	[p = 0.561]

Note: All specifications include the following additional controls: CPI (in same form, i.e., level or log, as relative price variable); black male; white female; black female; high income; high school; high school plus; age and its square; three region dummies; and region-specific linear time trends. Results from these coefficients are available upon request. IV specifications use unleaded to instrument for the relative price variable (in same form, i.e., level or log, as relative price variable) and include all of the covariates in both stages. Results from the first stage regression are available upon request.

Table 5
Effect of Relative Price of Healthful Foods on BMI, Overweight, and Obesity
By Sex
(standard errors clustered by region in parentheses)
[upper-tailed p-values generated by wild cluster bootstrap in brackets]

	Specification of Price Variables			
	Levels		Logs	
	OLS	IV	OLS	IV
Women				
BMI	0.844	0.608	0.365	0.419
	(0.302)	(0.500)	(0.165)	(0.341)
	[p = 0.029]	[p = 0.124]	[p = 0.027]	[p = 0.090]
Overweight	0.086	0.039	0.040	0.033
	(0.025)	(0.054)	(0.014)	(0.026)
	[p = 0.021]	[p = 0.241]	[p = 0.013]	[p = 0.098]
Obese	0.034	0.018	0.018	0.023
	(0.023)	(0.034)	(0.011)	(0.020)
	[p = 0.075]	[p = 0.292]	[p = 0.047]	[p = 0.114]
Men				
BMI	0.709	0.560	0.336	0.383
	(0.270)	(0.446)	(0.101)	(0.211)
	[p = 0.033]	[p = 0.081]	[p = 0.008]	[p = 0.017]
Overweight	0.090	0.075	0.039	0.033
	(0.022)	(0.041)	(0.006)	(0.008)
	[p = 0.011]	[p = 0.053]	[p = 0.000]	[p = 0.004]
Obese	0.045	0.038	0.016	0.018
	(0.020)	(0.027)	(0.007)	(0.015)
	[p = 0.021]	[p = 0.047]	[p = 0.020]	[p = 0.091]

Note: All specifications include the following additional controls: CPI (in same form, i.e., level or log, as relative price variable); black male; white female; black female; high income; high school; high school plus; age and its square; three region dummies; and region-specific linear time trends. Results from these coefficients are available upon request. IV specifications use unleaded to instrument for the relative price variable (in same form, i.e., level or log, as relative price variable) and include all of the covariates in both stages. Results from the first stage regression are available upon request.

Table 6
Effect of Relative Price of Healthful Foods on BMI, Overweight, and Obesity
By Race
(standard errors clustered by region in parentheses)
[upper-tailed p-values generated by wild cluster bootstrap in brackets]

	Specification of Price Variables			
	Levels		Logs	
	OLS	IV	OLS	IV
Blacks				
BMI	-0.299	-1.661	0.242	0.409
	(0.392)	(1.650)	(0.134)	(0.200)
	[p = 0.792]	[p = 0.975]	[p = 0.085]	[p = 0.084]
Overweight	-0.048	-0.187	0.011	0.008
	(0.037)	(0.106)	(0.012)	(0.011)
	[p = 0. 938]	[p = 0.950]	[p = 0. 215]	[p = 0.219]
Obese	0.013	-0.020	0.022	0.049
	(0.048)	(0.148)	(0.011)	(0.008)
	[p = 0. 407]	[p = 0.501]	[p = 0. 072]	[p = 0.000]
Whites				
BMI	0.862	0.689	0.335	0.327
	(0.322)	(0.479)	(0.162)	(0.289)
	[p = 0.015]	[p = 0.080]	[p = 0.033]	[p = 0.126]
Overweight	0.101	0.070	0.041	0.028
	(0.027)	(0.045)	(0.013)	(0.020)
	[p = 0. 008]	[p = 0.091]	[p = 0. 014]	[p = 0.090]
Obese	0.040	0.026	0.015	0.014
	(0.029)	(0.036)	(0.011)	(0.017)
	[p = 0. 074]	[p = 0.237]	[p = 0. 094]	[p = 0.287]

Note: All specifications include the following additional controls: CPI (in same form, i.e., level or log, as relative price variable); black male; white female; black female; high income; high school; high school plus; age and its square; three region dummies; and region-specific linear time trends. Results from these coefficients are available upon request. IV specifications use unleaded to instrument for the relative price variable (in same form, i.e., level or log, as relative price variable) and include all of the covariates in both stages. Results from the first stage regression are available upon request.

Appendix Table 1
Effect of Relative Price of Healthful Foods on BMI, Overweight, and Obesity
Covariates for OLS Specifications from Table 3
(standard errors clustered by region in parentheses)

	Price Variables in Levels		
	BMI	Overweight	Obese
Relative Price	0.764	0.087	0.039
	(0.277)	(0.024)	(0.022)
CPI	0.008	0.000	0.001
	(0.009)	(0.000)	(0.000)
Black Male	0.007	-0.009	0.010
	(0.099)	(0.008)	(0.007)
White Female	-1.479	-0.198	-0.004
	(0.037)	(0.005)	(0.002)
Black Female	0.827	-0.005	0.105
	(0.168)	(0.013)	(0.009)
High Income	-0.464	-0.026	-0.032
	(0.036)	(0.002)	(0.002)
Age	0.310	0.028	0.013
	(0.006)	(0.000)	(0.000)
Age^2	-0.003	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
High School	-0.583	-0.048	-0.038
	(0.070)	(0.009)	(0.003)
High School Plus	-1.182	-0.106	-0.070
	(0.085)	(0.011)	(0.004)
Midwest	-72.056	-4.389	-4.726
	(21.777)	(1.078)	(1.209)
South	-98.986	-7.489	-5.359
	(22.894)	(1.134)	(1.264)
West	-76.626	-7.482	-2.737
	(16.929)	(0.885)	(0.928)
Northeast	0.031	0.005	0.001
Trend	(0.055)	(0.003)	(0.003)
Midwest	0.067	0.007	0.004
Trend	(0.045)	(0.002)	(0.002)
South	0.081	0.009	0.004
Trend	(0.044)	(0.002)	(0.002)
West	0.069	0.009	0.003
Trend	(0.047)	(0.002)	(0.003)
Observations	1,039,457	1,039,457	1,039,457
R^2	0.409	0.413	0.196

Appendix Table 2
Coefficients on Instrument from First Stage of Table 3 IV Specifications
Pooled Sample
(standard errors clustered by region in parentheses)

	Specification of Price Variables	
	Levels	Logs
Unleaded	0.235	0.483
	(0.074)	(0.118)
Incremental R^2	0.115	0.143

Note: All specifications include the following additional controls: CPI (in same form, i.e., level or log, as relative price variable); black male; white female; black female; high income; high school; high school plus; age and its square; three region dummies; and region-specific linear time trends. IV specifications use unleaded to instrument for the relative price variable (in same form, i.e., level or log, as relative price variable) and include all of the covariates in both stages. Incremental R^2 represents the difference between the R^2 yielded by the first stage regression with the instrument included and the R^2 from a similar regression with the instrument omitted. Results for the additional covariates are available upon request.

Appendix Table 3
Effect of Relative Price of Healthful Foods on BMI, Overweight, and Obesity
Including Total Food Price as a Control
Pooled Sample
(standard errors clustered by region in parentheses)

	Specification of Price Variables			
	Levels		Logs	
	OLS	IV	OLS	IV
BMI	0.844	1.063	0.380	0.551
	(0.222)	(0.320)	(0.080)	(0.264)
Overweight	0.092	0.087	0.041	0.040
	(0.019)	(0.027)	(0.006)	(0.013)
Obese	0.043	0.063	0.019	0.032
	(0.019)	(0.021)	(0.006)	(0.013)

Note: All specifications include the following additional controls: CPI (in same form, i.e., level or log, as relative price variable); black male; white female; black female; high income; high school; high school plus; age and its square; three region dummies; and region-specific linear time trends. These specifications also include the total food price variable as a control. IV specifications use unleaded to instrument for the relative price variable (in same form, i.e., level or log, as relative price variable) and include all of the covariates in both stages. Results for the other covariates are available upon request.