Supersizing Supercenters? The Impact of Wal-Mart Supercenters on Body Mass Index and Obesity

January 1, 2010

Charles Courtemanche*
Department of Economics, University of North Carolina at Greensboro
P.O. Box 26165
Greensboro, NC 27402
cjcourte@uncg.edu
Office: 336-334-3910; Fax: 336-334-4089

Art Carden
Department of Economics and Business, Rhodes College
2000 N. Parkway
Memphis TN 38112
cardena@rhodes.edu
Office: 901-843-3829; Fax: 901-843-3736

Abstract
Research attributes much of the rise in obesity to technological progress reducing the cost of food consumption. We examine this hypothesis in the context of Walmart Supercenters, whose advancements in retail logistics have translated to substantial reductions in food prices. Using data from the Behavioral Risk Factor Surveillance System matched with Walmart Supercenter entry dates and locations, we examine the effects of Supercenters on body mass index (BMI) and obesity. We account for the endogeneity of Walmart Supercenter locations with an instrumental variables approach that exploits the unique geographical pattern of Supercenter expansion around Walmart's headquarters in Bentonville, Arkansas. An additional Supercenter per 100,000 residents increases average BMI by 0.25 units and the obesity rate by 2.4 percentage points. These results imply that the proliferation of Walmart Supercenters explains 11% of the rise in obesity since the late 1980s, but the resulting increase in medical expenditures offsets only a small portion of consumers' savings from shopping at Supercenters.

Keywords: Walmart, Wal-Mart, supercenter, obesity, body weight, body mass index
JEL Codes: I10

1. Introduction and Background

Rising obesity is both a public health and public finance concern. Lower food prices and easier access to food are plausible explanations for rising obesity, and the rise of "Big Box" chains is one of the most important retail trends of the post-World War II era. We present evidence that Walmart Supercenters' Every Day Low Prices for food help explain shoppers' larger waistlines since they first appeared in the late 1980s, but the cost of the increase in obesity attributable to Supercenters is small relative to the savings from Walmart-induced lower prices.

Between 1960 and 2006, the percentage of the population considered overweight rose from 43% to 67%, while the obesity rate grew from 13% to 34% (Flegal et al, 1998; National Center of Health Statistics, 2008).[1] Obesity has been linked to higher prevalence of diseases such as high blood pressure, diabetes, heart disease, and stroke (Strum, 2002). The consequences of obesity include an estimated 112,000 deaths and $117 billion in medical expenditures per year, with about half of these expenditures paid for by Medicare and Medicaid (Flegal et al, 2005; U.S. Department of Health and Human Services, 2001; Finkelstein et al, 2003).

Obesity is particularly amenable to economic analysis because it is the direct result of individual choices in the face of changing incentives (Philipson and Posner, 2003:S87-S88). Weight gain is caused by an imbalance between "calories in" and "calories out," making it the function of economic variables such as the price of calories and the opportunity cost of exercise. The emerging consensus is that technological progress has altered these variables in ways leading to weight gain.[2] For instance, Philipson and Posner (2003) and Lakdawalla and

---

[1] Someone is considered "overweight" if her body mass index (BMI = weight in kilograms divided by height in meters squared) is greater than or equal to 25 and obese if her BMI is greater than or equal to 30.
[2] The literature on the economic causes of obesity is too extensive to fully survey here. See Rosin (2008) for a more comprehensive review.

1

Philipson (2002) suggest that rising obesity in the United States is the result of falling real food prices due to increased agricultural productivity and a movement from jobs requiring physical activity to more sedentary employment. Cutler et al (2003) argue that rising obesity can be attributed to technological innovations that have reduced the time cost of food preparation. Chou et al (2004) link several factors reflective of technological progress – including the falling price of grocery food, the falling price of restaurant food, and the rising number of restaurants – to the increasing obesity rate.

We contribute to the obesity literature by suggesting an additional mechanism through which technological progress may have contributed to the rise in obesity. While the literature has emphasized the role of advancements in food *production* technology in lowering food prices and increasing obesity, by studying Walmart Supercenters we examine the role of advancements in food *distribution* technology. The Walmart revolution is fundamentally a revolution in retail logistics; as Karjanen (2006:152) characterizes it, "Walmart does not manufacture or source locally, but merely acts as a global commodity supply chain, distributing globally sourced goods to local markets." The company's major innovations have been in its distribution channels, with store locations, warehouse locations, and logistics practices planned in such a way as to minimize wasted motion. These innovations translate to lower prices – prices so much lower that Hausman and Leibtag (2004, 2005) suggest the Consumer Price Index is mis-measured because of not sufficiently accounting for Walmart.

Since 1962, Walmart has grown from a single store in Rogers, Arkansas into the world's largest retailer. The discount store, Walmart's original business model, focuses primarily on consumer goods like clothing and electronics. The Supercenter, which debuted in 1988 in

2

Washington, Missouri, adds a full-service grocery store to create "one-stop shopping."[3] The number of Supercenters has risen rapidly since the early 1990s, coinciding with an increase in the growth rate of obesity that is particularly evident in the 1990s (Figure 1). As of 2009, Walmart operated 833 discount stores, 2,705 Supercenters, 605 members-only Sam's Club warehouses, and 151 smaller grocery stores called "Neighborhood Markets" in the U.S. Together these stores accounted for over $400 billion in annual sales.[4] Thanks mostly to the Supercenter, Walmart has become a force in the grocery market. Walmart's grocery market share averaged 14.9% in 2004 across 68 metropolitan areas, and its grocery volume was over twice the sales volume of the largest supermarket chain Kroger (Lord, 2006, pp. 55-62).

Walmart has attracted attention in the scholarly literature and in the popular press.[5] The most obvious manifestation of what Fishman (2006) calls *The Wal-Mart Effect* is the company's traditional policy of "Every Day Low Prices." Numerous studies have documented Walmart's negative impact on prices (Hausman and Leibtag, 2004, 2005; Basker 2005b; Basker and Noel 2007). Hausman and Leibtag (2004) cite data from studies showing that, even after accounting for discount cards and sales, Walmart maintains a price advantage of 8-27% on various food items. Basker and Noel (2007) estimate that grocery stores reduce their prices by 1-1.2% after the entry of a Walmart Supercenter but Supercenters still hold a price advantage of 10%. Other research has studied the labor market consequences of Walmart entry, with Basker (2005a) and Hicks (2007) finding positive effects on employment and/or wages but Neumark et al (2008) and

---

[3] The dates and locations in the preceding sentences were obtained from http://walmartstores.com/AboutUs/7603.aspx.
[4] This information was obtained from http://walmartstores.com/AboutUs/7606.aspx, http://walmartstores.com/AboutUs/7605.aspx, and http://walmartstores.com/sites/AnnualReport/2009/financial_highlights.html.
[5] More extensive reviews of the literature on Walmart can be found in Basker (2006b), Hicks (2007), and Carden et al. (2009a).

Dube el at (2007) finding negative effects.[6] Recent work has also examined Walmart's impacts

on a diverse set of other outcomes, ranging from economic indicators such as the poverty rate

(Goetz and Swaminathan, 2006) and small business activity (Sobel and Dean, 2008) to cultural

indicators such as social capital (Goetz and Rupasingha, 2006; Carden et al, 2009a), leisure

activities (Carden and Courtemanche, 2009), and traditional values (Carden et al, 2009b).

This paper contributes to the Walmart literature by introducing another possible

consequence of Walmart entry: Supercenters' low food prices may have increased the incidence

of obesity. We examine the impacts of Supercenters on body mass index (BMI) and obesity

using data on the county-level prevalence of these stores matched with individual survey

responses from the Behavioral Risk Factor Surveillance System (BRFSS). We address the

endogeneity of Supercenter entry with an instrumental variables approach that exploits the

geographic pattern of Supercenter expansion around Walmart's headquarters in Bentonville,

Arkansas in the 1990s and 2000s. We also perform a number of robustness checks that account

for potential confounders such as Sam's Clubs, Walmart discount stores, and differential trends

in weight on the basis of population, economic development, and geographic location.

The evidence suggests that the entry of an additional Supercenter per 100,000 residents

increases average BMI by 0.25 units and individuals' probability of being obese by 2.4

percentage points. The effects are strongest for women, low-income married individuals, and

those living in the least populated counties. These estimates imply that the proliferation of

Walmart Supercenters explains 11% of the rise in obesity since the late 1980s. Importantly,

though, the resulting increase in medical expenditures offsets only 6% of consumers' savings

from shopping at Walmart. The obesity effect alone is therefore not sufficient to conclude that

Walmart entry is bad for communities. Instead, our results simply provide information about one

---

[6] These different conclusions hinge on a methodological debate that we detail in Section 4.

of many factors community planners and policymakers should consider when evaluating costs and benefits of proposed Walmart stores.

2. Theory

The initial reasoning for suspecting a causal relationship between Walmart Supercenters and obesity is simple: Supercenters lead to cheap food (Hausman and Liebtag, 2004; Basker and Noel, 2007) and cheap food leads to weight gain (Philipson and Posner, 2003; Lakdawalla and Philipson, 2002; Chou et al., 2004). However, a more thorough analysis reveals that Supercenters could affect weight through a complex combination of substitution and income effects, and that the theoretical effect is ambiguous. The theoretical ambiguity motivates the need for empirical analysis.

Consider a representative agent who divides her food budget between healthy grocery food such as fresh fruits and vegetables (H), unhealthy grocery food such as processed snacks (U), and restaurant food (R). The prices of the three food types are P_H, P_U, and P_R. Body mass index (B) is increasing in consumption of each of the three food types, but assume that unhealthy food and restaurant food increase BMI more strongly than healthy food because of their greater caloric density.[7] Also, BMI is decreasing in calories expended (E), which we treat as exogenous for simplicity. Formally,

$$B = B(H, U, R, E) \tag{1}$$
$$H = H(P_H, P_U, P_R) \tag{2}$$
$$U = U(P_H, P_U, P_R) \tag{3}$$
$$R = R(P_H, P_U, P_R) \tag{4}$$

where

---

[7] Whether eating at restaurants leads to obesity is the subject of debate. Studies show that a higher frequency of eating fast food is associated with increased intake of calories, fat, and saturated fat consumed (i.e. Satia et al, 2004). Accordingly, Chou et al (2004) and Rashad (2006) find evidence that a drop in prices at either fast-food or full-service restaurants increases obesity. Chou et al (2004) and Dunn (2008) estimate a positive association between restaurant prevalence and obesity, but Anderson and Matsa (2007) find no evidence of a causal effect.

$$\frac{dB}{dH} > 0; \ \frac{dB}{dU} > 0; \ \frac{dB}{dR} > 0; \ \frac{dB}{dH} < \frac{dB}{dU}; \ \frac{dB}{dH} < \frac{dB}{dR}; \ \frac{dB}{dE} < 0. \tag{5}$$

Assume the three types of foods are substitutes, H and R are normal goods, and U could be either normal or inferior. We allow for the possibility that junk food is an inferior good because research documents a negative relationship between income and weight throughout most of the U.S. income distribution (i.e. Lakdawalla and Philipson, 2002; Chou et al., 2004), with one explanation being that additional income allows individuals to switch from cheap processed food to more expensive but healthier fresh food (Basiotis and Lino, 2002; Ranney and McNamara, 2002; Drewnowski and Specter, 2004). As shown in Table 1, the overall effects of P_H, P_U, and P_R on H, U, and R are

$$\frac{dH}{dP_H} < 0; \ \frac{dU}{dP_H} ? \ 0; \ \frac{dR}{dP_H} ? \ 0; \ \frac{dH}{dP_U} ? \ 0; \ \frac{dU}{dP_U} ? \ 0; \ \frac{dR}{dP_U} ? \ 0; \ \frac{dH}{dP_R} ? \ 0; \ \frac{dU}{dP_R} ? \ 0; \ \frac{dR}{dP_R} < 0. \tag{6}$$

[INSERT TABLE 1 HERE.]

Walmart Supercenters (S) sell healthy food and unhealthy food, but they do not sell restaurant food.[8] The effects of Walmart Supercenters on prices are

$$\frac{dP_H}{dS} < 0; \ \frac{dP_U}{dS} < 0; \ \frac{dP_R}{S} = 0. \tag{7}$$

BMI can therefore be expressed as

$$B = B(H(P_H(S), P_U(S), P_R), U(P_H(S), P_U(S), P_R), R(P_H(S), P_U(S), P_R), E) \tag{8}$$

and the overall effect of Supercenters on BMI is

$$\frac{dB}{dS} = \frac{\partial B}{\partial H}\left(\frac{\partial H}{\partial P_H}\frac{\partial P_H}{\partial S} + \frac{\partial H}{\partial P_U}\frac{\partial P_U}{\partial S}\right) + \frac{\partial B}{\partial U}\left(\frac{\partial U}{\partial P_H}\frac{\partial P_H}{\partial S} + \frac{\partial U}{\partial P_U}\frac{\partial P_U}{\partial S}\right) + \frac{\partial B}{\partial R}\left(\frac{\partial R}{\partial P_H}\frac{\partial P_H}{\partial S} + \frac{\partial R}{\partial P_U}\frac{\partial P_U}{\partial S}\right). \tag{9}$$

Since the magnitudes of the partial effects are ambiguous, the overall effect is ambiguous. A clear theoretical prediction would emerge if we assumed that unhealthy food is

---

[8] Some Supercenters do contain fast food restaurants, but there is little reason to suspect that Supercenters reduce the overall market price of restaurant food.

normal and that the cross-price elasticities of demand are zero—the first two terms would be positive and the third term would be zero, making the overall effect positive—however, these assumptions are implausibly strong. Relaxing them leads to several possible scenarios in which Supercenters would reduce *BMI*. For instance, if income effects dominate substitution effects and unhealthy food is inferior, Walmart shoppers could eat less junk food and lose weight. Additionally, if the magnitude of $\frac{\partial R}{\partial P_H}\frac{\partial P_H}{\partial S}$ is large, individuals could lose weight by substituting away from restaurant food and toward healthy home-cooked meals. Of particular interest is the possibility that, if Walmart leads to more substantial price reductions for healthy food than unhealthy food and $\frac{\partial U}{\partial P_H}$ is large, weight could fall as consumers substitute toward the former from the latter – an effect that is plausible given Gelbach et al.'s (2007) finding that the relative food prices impact obesity. Supercenters' effect on the relative price of healthy versus unhealthy food is unclear. On one hand, Walmart's efficiency along the supply chain is especially important for rapidly-depreciating perishable items like fresh fruits and vegetables. Also, Walmart's advances in transportation logistics and forecasting may have made it profitable for retailers to stock a wider range of fresh fruits and vegetables than was previously available locally, and the relative prices of goods that were previously unavailable necessarily fall. On the other hand, a business model built around a global distribution network and extended warehouse storage may be more amenable to processed foods than fresh produce.[9]

---

[9] The empirical literature does not reveal a clear relationship between the health quality of foods and the extent to which Walmarts affect their prices. Hausman and Leibtag (2004) find that the effects of "SMCs" (supercenters, mass merchandisers, and club stores) on the prices of apples, apple juice, and bananas are smaller than the effects on cookies, eggs, ham, and ice cream, but some of the largest effects they estimate are on the prices of lettuce and tomatoes. Basker and Noel (2007) estimate Walmart's effect on a number of food prices. The largest price effect is for margarine, but the second-largest is for lettuce and the smallest price decreases are for soda, shortening, ground beef, sugar, and eggs. The effects on the prices of milk and parmesan cheese are actually positive.

Food prices are the most obvious mechanism through which Supercenters could affect weight, but there are several other less obvious possibilities that reinforce the theoretical ambiguity. Cheap cigarettes and alcohol from Walmart could theoretically affect weight, as nicotine stimulates the metabolism and curbs appetite while alcohol contains empty calories.[10] Walmart's low prices for non-food items could impact eating habits to the extent that these items and food are substitutes or complements. "Sprawl-Mart" may reduce calorie expenditure by changing the nature of shopping from a walk between several downtown retailers to a one-stop drive to the suburbs.[11] Alternatively, Walmart's cheap sporting goods could encourage physical activity. Given the theoretical complexities involved with analyzing Walmart's effects on body weight, we next turn to empirical analysis.

3. Data

We focus on the effect of Walmart Supercenters on obesity, but we also utilize data on Walmart discount stores and Sam's Clubs in some regressions. We use Walmart Supercenter and discount store entry dates and locations generously made available online by Thomas J. Holmes.[12] Sam's Club locations through May of 2003 were collected by Austan Goolsbee and Chad Syverson and were generously provided by Chad Syverson via email. We updated the Sam's Club data using information on store opening dates since 2003 provided by Walmart Stores, Inc. In our main regressions we adjust for market size by calculating the number of stores per 100,000 residents in the county, using population data from the U.S. Census Bureau.

---

[10] The causal effect of cigarette prices on body weight is the subject of debate in the literature; see Chou et al (2004), Gruber and Frakes (2006), Baum (2008), Finkelstein et al (2008), and Courtemanche (2009).

[11] Recent studies have attempted to link obesity to sprawl, as the spreading out of a city to suburban areas reduces the viability of walking, bicycling, and mass transit as transportation options. Ewing et al. (2003), Giles-Corti et al. (2003), and Frank et al. (2004) show that increases in automobile-oriented infrastructure are associated with higher obesity, while Eid et al (2008) and Plantinga and Bernell (2007) find that at least some of this association can be attributed to reverse causality.

[12] http://www.econ.umn.edu/~holmes/data/WalMart/index.html. The data were used in Holmes (2008).

Scaling by population is common in the Walmart literature since presumably the effect of an additional Walmart store should be stronger in smaller counties with fewer shopping alternatives than larger counties.[13]

To construct the set of instruments discussed in Section 3.1, we utilize distance from the center of each county to the location of Walmart's headquarters in Bentonville, Arkansas, computed using the great circle distance formula. As a robustness check, we also use distance from Walmart's nearest food distribution center. Food distribution center locations and entry dates were obtained from Holmes' website.

Our individual-level data come from the 1996-2005 waves of the restricted version of the Behavioral Risk Factor Surveillance System (BRFSS), a telephone survey of health conditions and risky health behaviors conducted by state health departments and the Center for Disease Control. The BRFSS consists of repeated annual cross sections of randomly-selected individuals starting in 1984. In 1984, only 15 states participated, but this number grew rapidly and all states were participating by 1996. The number of respondents also grew rapidly, from 12,258 in 1984 to 355,710 in 2006. We utilize the restricted version of the BRFSS, which at the time of this paper included only the years 1994 to 2005, because this version includes the county identifiers for all individuals.[14] We use only the years starting in 1996 because, as noted by Neumark et al. (2008), distance from Bentonville predicted discount store entry until 1995, after which point Walmart spanned the entire country and "there was only filling in of stores in areas that already had them" (p. 412). By focusing only on years after 1995, we help to ensure that our distance-based instruments predict only Supercenter entry and not discount store entry as well.

---

[13] For examples, see Basker (2005a), Goetz and Rupasingha (2006), Goetz and Swaninathan (2006), Sobel and Dean (2008), and Carden et al. (2009a. 2009b).

[14] We thank the Center for Disease Control for graciously making the restricted data available. The publically available data contains some county identifiers in all years after 1988, but they are hidden for many states.

We utilize the following BRFSS variables. The BRFSS includes self-reported height and weight, which allows us to construct our dependent variables: BMI and an indicator variable for whether or not the respondent is obese (BMI≥30).[15] We also construct a set of individual-level controls containing the following variables. The data include household income categories; we calculate continuous measures of real incomes (in 2005 dollars) by assigning individuals incomes equal to the midpoint of their category and then adjusting for inflation using the consumer price index for all urban consumers from the Bureau of Labor Statistics. We convert information on educational attainment to a set of dummy variables for some high school, high school graduate, some college, and college graduate. We also include age and binary variables for whether or not the individual is married, female, black, and a race other than black or white.

We match these individual-level data to the number of Supercenters, discount stores, and Sam's Clubs in the county at the end of the year preceding the respondent's interview.[16] This allows some time for changes in store presence to affect eating habits, and for changes in eating habits to affect weight. In unreported regressions, we also estimated models using average store presence over the twelve months preceding the respondent's interview, and the results were similar.

We also add county-level characteristics as controls in some of the robustness checks. First, we include county unemployment rates from the Bureau of Labor Statistics (BLS). Additionally, we use number of general merchandise stores and grocery stores from the

---

[15] Self-reported weight and height may be problematic since people tend to underreport their weight and, to a lesser extent, exaggerate their height. Some economists have employed a correction for self-reported BMI developed by Cawley (1999). This correction the National Health and Nutrition Examination Survey, which includes both actual and self-reported weight and height, to estimate actual BMI as a function of self-reported BMI and a variety of demographic characteristics. Researchers have generally found that the correlation between actual and self-reported BMI is very high, and that correcting for measurement error does not substantially alter the coefficient estimates in regressions (Cawley, 1999 and Lakdawalla and Philipson, 2002). We therefore elect not to employ the correction in this paper.

[16] For instance, if a respondent is interviewed in 2000, we match her individual information to the number of stores in her county at the end of December, 1999.

Economic Census. These variables are only available every five years; we use the 1992, 1997, and 2002 censuses and impute the other years using linear interpolations and extrapolations.[17] Since general merchandise stores include Walmart Supercenters, we subtract them out to obtain the number of other general merchandise stores in the county. We again calculate the number of stores per 100,000 residents.

Our matched sample consists of 1,644,094 observations. We report variable names, sources, descriptions, and summary statistics in Table 2. The sample obesity rate is 21%. The average respondent has a BMI of 26.6 and lives in a county with 0.33 Supercenters per 100,000 residents.

[INSERT TABLE 2 HERE.]

4. Identification Strategy

A widely acknowledged problem with identifying the causal effects of Walmart is endogeneity bias from the non-random nature of Walmart's locations. Neumark et al. (2008) offer a detailed discussion of Walmart's expansion strategy using quotes from Sam's Walton's autobiography and maps of store locations and openings over time. To summarize, Walmart slowly expanded outward from Benton County, Arkansas, saturating a market before moving on to a new market further away. This process continued from 1962 to 1995, at which point Walmart stores had reached all corners of the U.S. Walmart's store location decisions until 1995 can therefore loosely be summarized as the result of a two-step process. First, the company decided on the distance ring (i.e. less than 100 miles from Bentonville, 500-600 miles from Bentonville, etc.) at which to focus its attention in a given year. Then, within that distance ring it

---

[17] The 2007 Economic Census was not publically available at the time of this paper, so we use only 1992, 1997, and 2002. In 1992, both variables are available in the Economic Census' *Census of Retail Trade.* In 1997 and 2002, they are in *Retail Trade.*

chose locations that maximized the expected sum of future profits. This process creates the potential for endogeneity bias in studies examining Walmart's effects. In particular, the areas with the highest expected future profits – presumably those where economic development is projected to occur the fastest, those which offer the most generous tax advantages, and those which are the most receptive to Walmart entry – are likely trending differently than other areas along numerous unobservable dimensions.

Different methods of addressing this endogeneity concern have been used in the literature, but an increasingly popular approach is to instrument for Walmart presence with interactions of year fixed effects and variables representing distance from Walmart's headquarters in Benton County. Recent papers by three different groups of authors have utilized variations of this approach to identify Walmart's impacts: Neumark et al. (2008) found a negative effect on retail employment, Dube et al. (2007) found negative effects on retail earnings and health benefits, and Carden et al. (2009a) found no effect on social capital. Identifying Walmart's effect using the interaction of time and distance from Benton County is appealing for several reasons. First, it is a natural approach given Walmart's aforementioned expansion path, meaning that the instruments are sufficiently powerful predictors of Walmart presence. Second, distance from Bentonville is plausibly exogenous since it was determined long before Walmart's founding. Additionally, by identifying off the *interaction* of time and distance as opposed to simply distance, these papers retain the ability to account for time-invariant confounders using fixed effects or differences.

This approach, however, is not without controversy. Replying to Neumark et al. (2008), Basker (2006) argue that the distance-year interactions are correlated with unobserved determinants of employment, making them invalid instruments. Specifically, she claims that

large population centers – which may experience different labor market shocks than other areas – are located disproportionately near the coasts, making them relatively far from Bentonville. She shows that Neumark et al.'s approach predicts an implausibly large effect of Walmart entry on local manufacturing employment. Neumark et al. claim, though, that even if a distance-based identification strategy is inappropriate to examine the impacts of Walmart on employment in geographically-concentrated industries such as manufacturing, this does not mean it is inappropriate to examine impacts on employment in less geographically-concentrated industries such as retail. They also show that the implausible effect on manufacturing employment disappears when linear distance ring-specific trends are added to the model. To summarize, the literature suggests that identifying Walmart's impacts using the interactions of time and distance from Walmart's headquarters is an intuitively appealing strategy that is likely appropriate in some contexts but not in others.

Our paper uses a variation of the distance-time interactions identification strategy to estimate the effect of Walmart Supercenters – as opposed to Walmart stores in general – on body weight. Our strategy is based on the observation that the proliferation of Supercenters beginning in 1988 follows a similar expansion path as the aforementioned proliferation of Walmart discount stores that began in 1962. As Neumark et al. (2008) note, Walmart's expansion occurs in essentially concentric circles around Bentonville, Arkansas for a large part of the company's history. Figures 2-10 show that Walmart Supercenters follow a similar pattern of expansion.

We begin with a naïve fixed effects specification:

$$WT_{ict} = \beta_0 + \beta_1 S_{ct} + \sum_{j=1}^{J} \beta_{2j} X_{jict} + \sum_{y=1}^{Y} \tau_y YR_y + \sum_{s=1}^{S} \alpha_s CY_s + \varepsilon_{ict} \qquad (10)$$

where WT_{ict} is a measure of weight – either BMI or an indicator for obesity status – for

individual i living in county c in year t. S_{cy} is the number of Supercenters per 100,000 residents

in the county. X_{jict}, where $j = 1, ..., J$, represents the individual-level control variables discussed

in Section 3. The remaining summation terms reflect year and county fixed effects.[18]

Approximately two-thirds of entering Supercenters during the sample period were conversions

from discount stores, while the other third were entirely new stores. The parameter of interest –

β_1 – should therefore be interpreted as the weighted average effect of new and converted

Supercenters. This is equivalent to the effect of a new Supercenter if discount stores have no

effect on weight, a plausible assumption given the limited quantity of food sold by discount

stores.[19] We compute heteroskedasticity-robust standard errors clustered by county.

The identifying assumption for β_1 is that Supercenter presence is uncorrelated with the

error term. Formally,

$$Cov(\varepsilon_{ict}, S_{ct}) = 0. \tag{11}$$

There is little reason to suspect unobserved individual characteristics to be correlated with

county-level Supercenter presence conditional on unobserved county characteristics. Therefore,

the identifying assumption is effectively that changes over time in Supercenter presence are

uncorrelated with changes over time in unobservable county characteristics that impact body

weight. If Supercenter locations are the result of a two-part process in which Walmart selects

first a distance ring and then a county within the distance ring, there are two reasons for concern

---

[18] Because of the large sample size, we estimate linear probability models for these binary dependent variables.
Results are robust to the use of probit models. For all years, we scale the number of Supercenters by county
population in 1997 obtained from the U.S. Census Bureau. We fix county population so that the variation over time
comes entirely from variation in the number of Supercenters. Allowing population to vary over time would allow
the possibility that the results are driven by changes in population instead of changes in stores. Nonetheless, our
results are very similar using county population from the survey year.
[19] Another option would be to isolate the effect of new Supercenters by controlling for the number of discount
stores. We do not do this as it would create the need to instrument for a second endogenous variable in the
subsequent instrumental variable analysis.

with this assumption. Most obviously, the choice of store locations within a distance ring is endogenous. If Walmart chooses counties trending upward in terms of population, income, and economic activity, $\hat{\beta}_1$ could be biased downward because of the aforementioned negative association between income and weight. Alternatively, rapid economic development may increase availability of restaurant food, in which case $\hat{\beta}_1$ could be biased upward.

A second, though less obvious, concern is that Walmart's targeted distance ring(s) may be trending differently in obesity than other areas for reasons aside from Walmart. Basker's (2006) argument about differential trends could apply to body weight if differences in demographic characteristics between coastal and inland areas translate to different growth rates of obesity. A specific concern is that southern states have higher obesity rates on average than other states (Centers for Disease Control and Prevention, 2009).[20] Of course, some of this difference could be the causal effect of Walmart, while much of the remaining regional variation should be captured by the county fixed effects and controls for potential confounders such as income and education. Also, while Bentonville is technically in the south, its location in the northwest corner of Arkansas is less than 500 miles from the geographic center of the U.S. Columbus, GA and Denver, CO are nearly equidistant from Bentonville, as are Shreveport, LA and Saint Louis, MO as well as Knoxville, TN and Madison, WI (Carden et al., 2009a). Nonetheless, the possibility for bias remains.

We next eliminate the first of these two concerns by estimating an instrumental variables model using interactions of distance with each of the year fixed effects as instruments for Supercenter presence. The first stage estimates

---

[20] Indeed, there is even an annual conference—the Southern Obesity Summit—aimed at reducing Southern obesity. Their website is www.southernobesitysummit.org.

$$S_{ct} = \gamma_0 + \sum_{j=1}^{J} \gamma_j X_{jict} + \sum_{y=1}^{Y} \sigma_y YR_y + \sum_{s=1}^{S} \theta_s CY_s + \sum_{y=1}^{Y} \varphi_y (DIST_c * YR_y) + \mu_{ict} \qquad (12)$$

while the second stage estimates equation (10) using predicted instead of actual Supercenter presence.[21] $DIST_c$ is county c's distance in miles from Walmart's headquarters in Bentonville, AR. Different specifications for distance have been used in the literature. Neumark et al. (2008) constructed a set of dummy variables reflecting 100 mile distance increments (i.e. less than 100 miles from Benton County, 100 to 200 miles, etc.). Dube et al. (2007) used dummy variables for 10 distance rings in their baseline model, though they obtained similar results with a variety of alternative specifications. Carden et al. (2009a) used linear and quadratic distance specifications. Our results remain very similar regardless of the functional form for distance, so we use the simplest specification – linear distance – in our baseline model and experiment with alternatives in the robustness check section.

The instrumental variables model is valid under two assumptions. First, the distance-year interactions must impact Supercenter presence conditional on the controls and fixed effects. This assumption is empirically testable and seems likely to hold given the pattern of Supercenter expansion displayed in Figures 2-10. The second – and more controversial – assumption is that the distance-year interactions can be excluded from the second-stage model (10). In other words, they do not affect weight other than though their impact on Supercenter presence. The instrumental variables model therefore improves on the fixed effects model in that it does not rely on the strong assumption that Supercenter locations within distance rings are exogenous. It does, however, still rely on the assumption that the different layers of distance have a common

---

[21] We estimate the model using the Stata module xtivreg2 by Schaffer (2005).

trend in unobservable characteristics that affect body weight. We examine the validity of this assumption in Section 6 with a series of robustness checks.

5. Results

Table 3 reports the coefficient estimates for the distance-year instruments from the first-stage regression, along with the F statistic from a test of joint significance of these variables. The distance*1996 interaction is omitted to prevent perfect collinearity; the coefficient estimates can therefore be interpreted as the difference between the effect of distance in the given year and the effect of distance in 1996. Each distance-year interaction is significant at the 1% level. The coefficient estimates are all negative and become increasingly negative each year, meaning that distance from Bentonville became a stronger predictor of Supercenter presence over time during our sample period. The F statistic for joint significance of the instruments is 21.75, safely over commonly-accepted levels at which a weak instrument problem can be ruled out.

[INSERT TABLE 3 HERE.]

Table 4 reports the coefficient estimate of interest from the two second-stage regressions, along with the corresponding OLS estimates. We also report results from a Hausman test of the consistency of the OLS estimator, as well as the p-value from the test of the overidentifying restrictions. The OLS estimates suggest Supercenters have little to no effect on BMI or obesity. The IV estimates, however, are both statistically significant at the 1% level and much more strongly positive than the OLS estimates. An additional Supercenter per 100,000 residents increases BMI by an average of 0.251 units – 1.6 pounds at the sample mean height – and P(Obese) by 2.4 percentage points. Both Hausman tests strongly reject the consistency of the OLS estimator, while the overidentification tests fail to reject the null hypothesis that the set of

17

instruments is valid. (Admittedly, the results from the overidentification test are of limited

usefulness since all instruments are constructed using the same distance variable.)

[INSERT TABLE 4 HERE.]

We next evaluate the economic significance of these results by using the estimated effect

on obesity to answer two questions: 1) What portion of the rise in obesity since in 1988 – when

the first Walmart Supercenter opened – can be attributed to Supercenters, and 2) How much of

consumers' savings from shopping at Supercenters are offset by additional medical expenses

from the increased obesity rate? For the first question, define ΔO as the change in the BRFSS

obesity rate between 1988 and 2005, 14.3 percentage points. $\bar{S}$ is the sample mean for

Supercenters per 100,000 residents in 2005, and $\frac{dO}{dS}$ is the estimated effect of Supercenters on

obesity from the baseline IV regression. The percentage of the rise in obesity that can be

attributed to Supercenters is

$$P_S = \frac{\bar{S} * \frac{dO}{dS}}{\Delta O} * 100\% = \frac{0.651 * 0.024}{0.143} * 100\% = 10.9\%. \tag{13}$$

Our results therefore imply that the existence of Walmart Supercenters accounts for 10.9% of the

rise in obesity since the late 1980s.

The percentage of savings from Walmart Supercenters offset by the increased medical

expenditures is

$$P_{OFFSET} = \frac{\bar{S} * \frac{dO}{dS} * MED}{H * SAV} * 100\%. \tag{14}$$

$\bar{S} * \frac{dO}{dS}$ represents the percentage increase in obesity from Supercenters. MED represents the total

annual medical expenditures in the U.S. from obesity, estimated at $117 billion (U.S.

Department of Health and Human Services, 2001). The total annual savings from Walmart is

equal to the number of households H times the average savings per household SAV. For H, we use the number of households in the U.S. according to the 2000 census, 105,480,101.[22] In Appendix A, we calculate $SAV = \$177$ based on estimates of Walmart's price effects from the literature and information from the 2002 *Census of Retail Trade* and Walmart's 2002 annual report. Since this estimate is based on 2002 information, we also use 2002 levels of Supercenter prevalence for $\bar{S}$. Ultimately,

$$P_{OFFSET} = \frac{0.390 * 0.024 * \$117,000,000,000}{105,480,101 * \$177} = 5.9\%. \tag{15}$$

Our estimates therefore suggest that medical expenditures from the increase in obesity offset approximately 5.9% of consumers' savings from shopping at Walmart Supercenters. While our calculations are admittedly crude, they are sufficient to illustrate the point that, while Walmart Supercenters appear to have been an important contributor to the rise in obesity, the obesity effect only offsets a small portion of the savings from their lower prices.

6. Robustness Checks

In this section, we evaluate the sensitivity of our baseline IV estimates to a number of alternative specifications. Our robustness checks fall into four categories: 1) alternate functional forms for the Supercenter variable, 2) alternate functional forms for the distance-year instruments, 3) checking whether the estimates are confounded by Walmart discount store entry or Sam's Club entry, and 4) evaluating the potential for bias from differential trends in weight on the basis of distance from Bentonville.

While our scaling of the number of Supercenters in the county by population follows much of the Walmart literature, papers in the literature have also used the raw number of stores

---

(Basker and Noel, 2007) and a binary variable reflecting whether there is at least one store in the county (Basker, 2005b). In Panel A of Table 5, we report the results using these two functional forms for Supercenters, as well as another specification that scales the number of Supercenters by land area in hundreds of square miles. The effects of Supercenters on BMI and P(Obese) remain positive in all regressions, and are statistically significant in five of the six. The magnitudes are not comparable across specifications because a one unit increase in the Supercenter variable represents a different shock to the market depending on which functional form is used. For instance, an additional Supercenter per 100,000 residents is an increase of 1.26 standard deviations, while an additional (unscaled) Supercenter is an increase of 0.74 standard deviations. Nonetheless, the results from Panel A show that the conclusion that Supercenters increase BMI and obesity does not depend on the functional form used for Supercenters. We therefore return to the per capita Supercenter measure for all subsequent regressions.

[INSERT TABLE 5 HERE.]

We next examine the sensitivity of the results to the different specifications for the distance-year interactions used in the Walmart literature. As discussed in Section 4, these include a quadratic functional form for distance (Carden et al., 2009a), distance rings of 100 miles (Neumark et at, 2008), and 10 distance rings (Dube et al., 2007). We construct distance rings of 100 miles by creating a dummy variable equal to one if the county is less than 100 miles from Bentonville, another if the county is 100 or more miles away but less than 200, etc., up to the last ring of 1600 miles or more.[23] We then create the set of instruments by interacting each of the distance ring dummies with each of the year fixed effects. In the models with 10 distance rings, each ring represents 170.31 miles, reflecting the maximum distance from Bentonville in

_____

[23] The maximum distance from Bentonville in the sample is 1703.1 miles; we combine the few observations for whom distance is greater than 1700 miles with the 1600-1700 mile category.

the sample divided by 10. As shown in Panel B of Table 5, the estimates are virtually identical

to the baseline estimates from Table 4. The results are therefore not sensitive to different

functional forms for distance from Bentonville, and we return to the linear specification in

subsequent regressions unless otherwise indicated.

Next, we consider the possibility that our estimates are confounded by Walmart discount

store or Sam's Club entry. Since these two chains are also operated by Walmart Stores, Inc.,

their expansion patterns were likely influenced by distance from Bentonville. If these stores

impact body weight, our IV estimator of the effect of Supercenters could potentially be biased.

However, it seems unlikely that bias from discount stores or Sam's Clubs is driving our results

for two reasons. First, we restricted the sample to the years 1996 and later because, as

documented by Neumark et al. (2008), Walmart discount stores had spread to all corners of the

U.S. by 1995, after which point new discount stores entered to fill gaps instead of expand the

chain to new territory. Distance from Bentonville therefore predicts discount store entry much

more weakly in the post-1995 period than in the pre-1995 period. Similarly, Sam's Clubs had

also spread to all corners of the U.S. by 1995.

Second, even if distance from Bentonville did strongly predict discount store and Sam's

Club entry during the sample period, it seems unlikely that these stores have meaningful effects

on body weight. Discount stores sell primarily consumer goods such as clothing and electronics,

so they probably do not significantly impact market prices for food. Sam's Clubs primarily

served small businesses during the sample period, as evidenced by their slogan "We are in

business for small business."[24] It is therefore unlikely that they affect consumers' eating habits

enough to lead to meaningful changes in weight. Nonetheless, we conduct robustness checks to

---

[24] See http://www.fundinguniverse.com/company-histories/Sams-Club-Company-History.html for a brief history of
Sam's Club that discusses its focus on small businesses beginning in 1994.

test whether discount stores and Sam's Clubs confound our estimates for Supercenters. We control for the per capita number of discount stores and Sam's Clubs opened in the county, and we also estimate models dropping counties that experienced discount store or Sam's Club entry during the sample period. As shown in Panel C of Table 5, controlling for these stores has essentially no effect on the estimates, while dropping counties that experienced their entry actually increases the estimates somewhat. We therefore find no evidence that neglecting to account for discount stores or Sam's Clubs led to the estimation of a spurious positive effect of Supercenters on weight in the baseline regression.

Finally, we estimate a number of additional models to examine the validity of the assumption that underlying trends in body weight do not vary systematically on the basis of distance from Bentonville. As discussed in Section 4, Bentonville is located close to the center of the U.S. and is equidistant from very diverse locations, so it is not immediately obvious that this assumption would be violated. As noted by Basker (2006), however, large population centers – which may experience different economic shocks than other areas – are located disproportionately near the coasts and further from Bentonville. We therefore test for differential trends on the basis of population and economic activity by estimating models including interactions of the year fixed effects with county population density, unemployment rate, grocery stores per capita, and general merchandise stores per capita.[25] We also run regressions dropping the furthest 10% of counties from Bentonville (those beyond 1119.74 miles), the closest 10% of counties to Bentonville (those within 264.06 miles), and counties with a population of over 1,000,000. This helps to ensure that our results are not driven by large coastal population centers and/or rural areas close to Bentonville. Next, we estimate two models that account for differential trends more generally: one including linear 100 mile distance ring-specific time

---

[25] Ruhm (2000, 2005) documents a negative relationship between unemployment rate and obesity.

trends, and another including interactions of the year fixed effects with dummy variables reflecting the nine U.S. Census Bureau census divisions.[26] Ring-specific time trends and census division-by-year effects eliminate much of the variation in the instruments if the linear specification for distance is used; we therefore use the most flexible of the distance specifications – 100 mile distance rings – for the instruments in these regressions. Finally, we consider a different identification strategy altogether: we use interactions of the year effects with distance from the nearest Walmart food distribution center as the instruments. This approach is less susceptible to concern about differential trends between coastal and inland areas because Walmart food distribution centers are scattered throughout the country. However, food distribution center locations and Supercenter locations may be jointly determined. We also estimate a model using both the distance from Bentonville-year interactions and the distance from the nearest food distribution center-year interactions as instruments.

Panel D of Table 5 displays the results from these additional robustness checks. Supercenters are statistically significant in 20 of the 22 regressions, and the estimates are similar across the different specifications. An additional Supercenter increases BMI by 0.21-0.35 units and P(Obese) by 1.4-3.4 percentage points. The 95% confidence intervals include the point estimate from the baseline IV regression in all 22 regressions, so there is no evidence that the baseline estimator is inconsistent. We only report the overidentification test p-value for the regression with the distances from both Bentonville and the nearest food distribution center as instruments, as this is the only model with a plausible overidentifying restriction. For both BMI and obesity status, the overidentification test fails to reject the null hypothesis that the set of instruments is valid.

---

[26] Census division classifications are available at http://www.census.gov/geo/www/us_regdiv.pdf.

7. Heterogeneity

We next return to the baseline IV model and evaluate the potential for heterogeneity on the bases of gender, income, and county population. Stratifying by gender is common in the obesity literature. We stratify by income and population as well since Walmart's shoppers are disproportionately low-income individuals living in rural areas or small towns. For income and population, we divide the sample into those below the 25th percentile, between the 25th and 75th percentiles, and above the 75th percentile. When splitting the sample by income, we also stratify by marital status since household income depends largely on the number of adults in the family.

Table 6 reports the results. The effect of Supercenters on BMI is stronger for women than men, but the effect on P(Obese) is stronger for men. This suggests that the impact on men is skewed more heavily toward the right tail of the BMI distribution. Supercenters have the strongest effect on the weights of low-income married individuals, while there is no clear connection between the income of single individuals and the strength of the effect. The most striking result shown in the table, however, is the strong impact on the weight of people living in the least populated areas. An additional Supercenter per 100,000 residents increases average BMI by 0.7 units (4.5 pounds at the sample mean height) and P(Obese) by 6.2 percentage points among individuals below the 25th percentile in county population. These large effects may reflect the relative lack of shopping alternatives in rural areas.

[INSERT TABLE 6 HERE.]

8. Conclusion

This paper exploits the unique geographic pattern of Walmart Supercenter expansion to identify the effect of county-level Supercenter presence on individual BMI and obesity status. We find evidence that Supercenters increase both BMI and obesity, with effects that are largest

for women, low-income married individuals, and those living in the least populous counties. The estimates imply that the proliferation of Walmart Supercenters explains 11% of the rise in obesity since the late 1980s, but that the increase in medical expenditures offsets only 6% of consumers' savings from shopping at Walmart.

We contribute to the Walmart literature by providing information about one of many factors that communities should consider when incentivizing Walmart entry. We also contribute methodologically by showing that distance from Bentonville predicts Supercenter entry in the post-1995 period in much the same way that it predicted discount store entry in the pre-1995 period. This may prove useful in identifying other effects of Supercenters. More generally, our exhaustive robustness testing contributes to the ongoing debate about the distance from Bentonville instrument by showing that it is likely appropriate in at least some contexts.

We contribute to the obesity literature by identifying another mechanism – improved efficiency in food distribution – through which technological progress likely contributed to the rise in obesity. Our findings epitomize an ongoing debate about the appropriateness of policy interventions designed to reduce obesity. The results are consistent with arguments made by Lakdawalla et al. (2005) and others that the rise in obesity can be largely attributed to rational responses to changing incentives from welfare-enhancing technological change, in which case government interference would be socially wasteful. However, the potential exists for market failures to lead to sub-optimal weight outcomes. Medical expenditures create a negative externality as individuals rarely pay their own medical bills (Bhattacharya and Sood, 2005). Consumers therefore receive the benefits from shopping at Walmart but do not internalize the costs, potentially leading to distortions. Also, self-control problems likely play a role in obesity

(Cutler et al., 2003), and cheap food from Supercenters may exacerbate these self-control problems.

Our findings point to natural avenues for future research. Future work should investigate the mechanisms through which Walmart Supercenters impact body weight. The BRFSS data do not include detailed information on eating habits during our sample period, preventing such an analysis here. Of particular interest is whether the increase in weight results from an across-the-board increase in food consumption as opposed to a substitution from healthy to unhealthy foods. The former would mean that Supercenters improve nutrient intake, potentially offsetting some of the health consequences of the rise in obesity. Research should also consider the health effects of Supercenters' reduced prices for cigarettes, alcohol, and prescription drugs. Finally, future research should examine the effect of other big box retailers on body weight. For instance, Target caters to a somewhat more affluent clientele than Walmart, and income and substitution effects for food could differ across the income distribution. This paper therefore provides a first step in understanding the health consequences of big box retailing, but much is left to learn about the subject.

References

Anderson, Michael L. and David A. Matsa. 2007. Are Restaurants Really Supersizing America? Working Paper, University of California-Berkeley and Northwestern University.

Basiotis, P.P. and Lino, M. (2002) Food Insufficiency and Prevalence of Overweight among Adult Women. *Nutrition Insights 26*, USDA Center for Nutrition Policy and Promotion.

Basker, Emek. 2005a. Job Creation or Destruction? Labor-Market Effects of Wal-Mart Expansion. *Review of Economics and Statistics* 87(1):174-183.

Basker, Emek. 2005b. Selling a Cheaper Mousetrap: Wal-Mart's Effect on Retail prices. *Journal of Urban Economics* 58(2):203-229.

Basker, Emek. 2006. When Good Instruments Go Bad: A Reply to Newumark, Zhang, and Ciccarella. University of Missouri Department of Economics Working Paper 07-06.

Basker, Emek. 2007. The Causes and Consequences of Wal-Mart's Growth. *Journal of Economic Perspectives* 21(3):177-198.

Basker, Emek and Michael Noel. 2007. The Evolving Food Chain: Competitive Effects of Wal-Mart's Entry into the Supermarket Industry. Working Paper, University of Missouri and University of California-San Diego.

Bhattacharya, Jay and Neeraj Sood. 2005. Health Insurance and the Obesity Externality. National Bureau of Economic Research, Working Paper 11529.

Carden, Art and Charles Courtemanche. 2009. Wal-Mart, Leisure, and Culture. *Contemporary Economic Policy* 27(4):450-461.

Carden, Art, Charles Courtemanche, and Jeremy Meiners. 2009a. Does Wal-Mart Reduce Social Capital? *Public Choice* 138:109-136.

Carden, Art, Charles Courtemanche, and Jeremy Meiners. 2009b. Painting the Town Red? Wal-Mart and Values. *Business and Politics* 11(2) article 5.

Cawley, J. 1999. Rational Addiction, the Consumption of Calories, and Body Weight. PhD Dissertation, University of Chicago.

Centers for Disease Control and Prevention. 2009. Estimated County-Level Prevalence of Diabetes and Obesity—2007. *Morbidity and Mortality Weekly Report* 58(45):1259-1263.

Chou S., Grossman M., and Saffer H., 2004. An economic analysis of adult obesity: Results from the Behavioral Risk Factor Surveillance System. *Journal of Health Economics* 23: 565-587.

Chou S., Grossman M., and Saffer H., 2006. Reply to Jonathan Gruber and Michael Frakes. *Journal of Health Economics* 25: 389-393.

Courtemanche, Charles, 2009. Rising Cigarette Prices and Rising Obesity: Coincidence or Unintended Consequence? *Journal of Health Economics*, forthcoming.

Cutler, Glaeser, and Shapiro, 2003. Why Have Americans Become More Obese? *Journal of Economic Perspectives*.

Drewnowski, A. and Specter, S.E. (2004). Poverty and obesity: the role of energy density and energy costs. *American Journal of Clinical Nutrition* 79(1): 6–16.

Dube, A., T. Lester, and B. Eidlin (2007). Firm Entry and Wages: Impact of Wal-Mart Growth on Earnings Throughout the Retail Sector. Institute of Industrial Relations Working Paper No. iirwps-126-05.

Dunn, Richard, 2008. Obesity and the Availability of Fast-food: an instrumental variables Approach. Working paper, Univeristy of Wisconsin-Madison.

Eid, Jean, Henrey G. Overman, Diego Puga, and Matthew A. Turner. 2008. Fat City: Questioning the Relationship Between Urban Sprawl and Obesity. *Journal of Urban Economics* 63(2):385-404.

Ewing R, Schmid T, Killingsworth R, Zlot A, Raudenbush S. 2003. Relationship Between Urban Sprawl and Physical Activity, Obesity, and Morbidity. *American Journal of Health Promotion* 18(1):47–57.

Finkelstein E., I. Fiebelkorn, G. Wang. National medical spending attributable to overweight and obesity: How much, and who's paying? *Health Affairs*, Web Exclusives: W219-W226.

Fishman, Charles. 2006. *The Walmart Effect*. New York: Penguin Press.

Flegal, K., M. Carroll, R. Kuczmarski, and C. Johnson. 1998. Overweight and Obesity in the United States: Prevalence and Trends, 1960-1994. *International Journal of Obesity*, 22(1): 39-47.

Flegal K., B. Graubard, D. Williamson, and M. Gail. 2005. Excess deaths associated with underweight, overweight, and obesity. *Journal of the American Medical Association* 293: 1861-1867.

Frank L., Andresen M., and Schmid T. 2004. Obesity Relationships with Community Design, Physical Activity, and Time Spent in Cars. *American Journal of Preventive Medicine* 27(2):87.96.

Gelbach J., Click J., and Stratmann T., 2007. Cheap Donuts and Expensive Broccoli: The Effect of Relative Prices on Obesity. Florida State University College of Law, Public Law Research

Paper #261.

Giles-Corti, Billie and Robert J. Donovan. 2003. Relative Influences of Individual, Social Environmental, and Physical Environmental Correlates of Walking. *American Journal of Public Health* 93: 1583-1589.

Global Insight. 2007. The Price Impact of Wal-Mart: An Update Through 2006. Available http://www.globalinsight.com/MultiClientStudy/MultiClientStudyDetail2438.htm.

Goetz, Stephan and Anil Rupasingha, 2006. Wal-Mart and Social Capital. *American Journal of Agricultural Economics*, American Agricultural Economics Association, vol. 88(5), pages 1304-1310

Goetz, Stephan and Hema Swaminathan, 2006. Wal-Mart and County-Wide Poverty. *Social Science Quarterly* 87(2): 211-226.

Gruber, Jonathan and Frakes, Michael, 2006. Does falling smoking lead to rising obesity? *Journal of Health Economics* 25: 183-187.

Hausman, Jerry and Ephraim Leibtag. 2004. CPI Bias from Supercenters: Does the BLS know that Wal-Mart Exists?" NBER Working Paper 10712.

Hausman, Jerry and Ephraim Leibtag. 2005. Consumer Benefits From Increased Competition in Shopping Outlets: Measuring the Effect of Wal-Mart. NBER Working Paper 11809.

Hicks, Michael J. 2007. Job Turnover and Wages in the Retail Sector: The Influence of Walmart. *Journal of Private Enterprise* 22(2):137-159.

Holmes, Thomas J. 2008. The Diffusion of Wal-Mart and Economies of Density. NBER Working Paper w13783.

Karjanen, David. 2006. The Walmart Effect and the New Face of Capitalism: Labor Market and Community Effects of the Megaretailer, in Nelson Lichtenstein, ed. *Wal-Mart: The Face of Twenty-First Century Capitalism*. New York: The New Press, pp. 143-162.

Lakdawalla, Darius and Tomas Philipson. 2002. The Growth of Obesity and Technological Change: A Theoretical and Empirical Investigation. NBER Working Paper 8965.

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

Lord, J. Dennis. 2006. Walmart Supercenter Market Share of Grocery Retailing in U.S. Metropolitan Areas, in Stanley D. Brunn (ed.), *Walmart World: The World's Biggest Corporation in the Global Economy* . Oxford: Routledge, pp. 55-62.

Magrabi, F., Y. Chung, S. Cha, S. Yang. 1991. *The Economics of Household Consumption*. New York: Praeger Paperback.

National Center of Health Statistics. 2008. "Prevalence of overweight, obesity and extreme obesity among adults: United States, trends 1976-80 through 2005-2006." Available http://www.cdc.gov/nchs/data/hestat/overweight/overweight_adult.pdf.

Neumark, D., J. Zhang, and S. Ciccarella. 2008. The Effects of Wal-Mart on Local Labor Markets. *Journal of Urban Economics* 63(2): 405-430.

Nonnemaker, J., E. Finkelstein, M. Engelen, T. Hoerger, M. Farrelly. 2008. Have Efforts to Reduce Smoking Really Contributed to the Obesity Epidemic? *Economic Inquiry*, forthcoming.

Ogden, C., M. Carroll, L. Curtin, M. McDowell, C. Tabak, K. Felgal. Prevalence of Overweight and Obesity in the Unived States, 1999-2004. *Journal of the American Medical Association*, 295, 2006, 1549-55.

Philipson, T. and R. Posner. 2003. The Long Run Growth of Obesity as a Function of Technological Change. *Perspectives in Biology and Medicine* 46(3):S87-S108.

Plantinga, Andrew J. and Stephanie Bernell. 2007. The Association Between Urban Sprawl and Obesity: Is it a Two-Way Street? *Journal of Regional Science* 47(5):857-879.

Ranney, C.K. and McNamara, P.E. 2002. Do healthier diets cost more? Working Paper 02-22, Department of Applied Economics and Management, Cornell University.

Rashad I., Grossman M., Chou S., 2006. "The Super Size of America: An Economic Estimation of Body Mass Index and Obesity in Adults." *Eastern Economic Journal* 32(1): 133-148.

Rosin, Odelia. 2008. The Economic Causes of Obesity: A Survey. *Journal of Economic Surveys* 22(4):617-647.

Ruhm, Christopher. 2000. Are Recessions Good For Your Health? *Quarterly Journal of Economics*, Vol. 115, No. 2, 617-650.

Ruhm, Christopher. 2005. Healthy Living in Hard Times, *Journal of Health Economics*, 24(2), March 2005, 341-63.

Satia, Jessie, Joseph Galanko, and Anna Maria Siega-Riz, 2004. Eating at Fast-food Restaurants is Associated with Dietary Intake, Demographic, Psychosocial and Behavioural Factors among African Americans in North Carolina. *Public Health Nutrition* 7:1089-1096.

Schaffer, Mark, 2005. XTIVREG2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. Available http://ideas.repec.org/c/boc/bocode/s456501.html.

Sobel, Russell and Andrea Dean, 2008. Has Wal-Mart Buried Mom and Pop?: The Impact of Wal-Mart on Self Employment and Small Establishments in the United States *Economic Inquiry* 46(4): 676-695.

Strum, R., 2002. The Effects of Obesity, Smoking, and Drinking on Medical Problems and Costs. Health Affairs 21: 245-53.

U.S. Department of Health and Human Services, 2001. The surgeon general's call to action to prevent and decrease overweight and obesity.

Figure 1 – Changes over Time in Number of Walmart Supercenters and the Obesity Rate
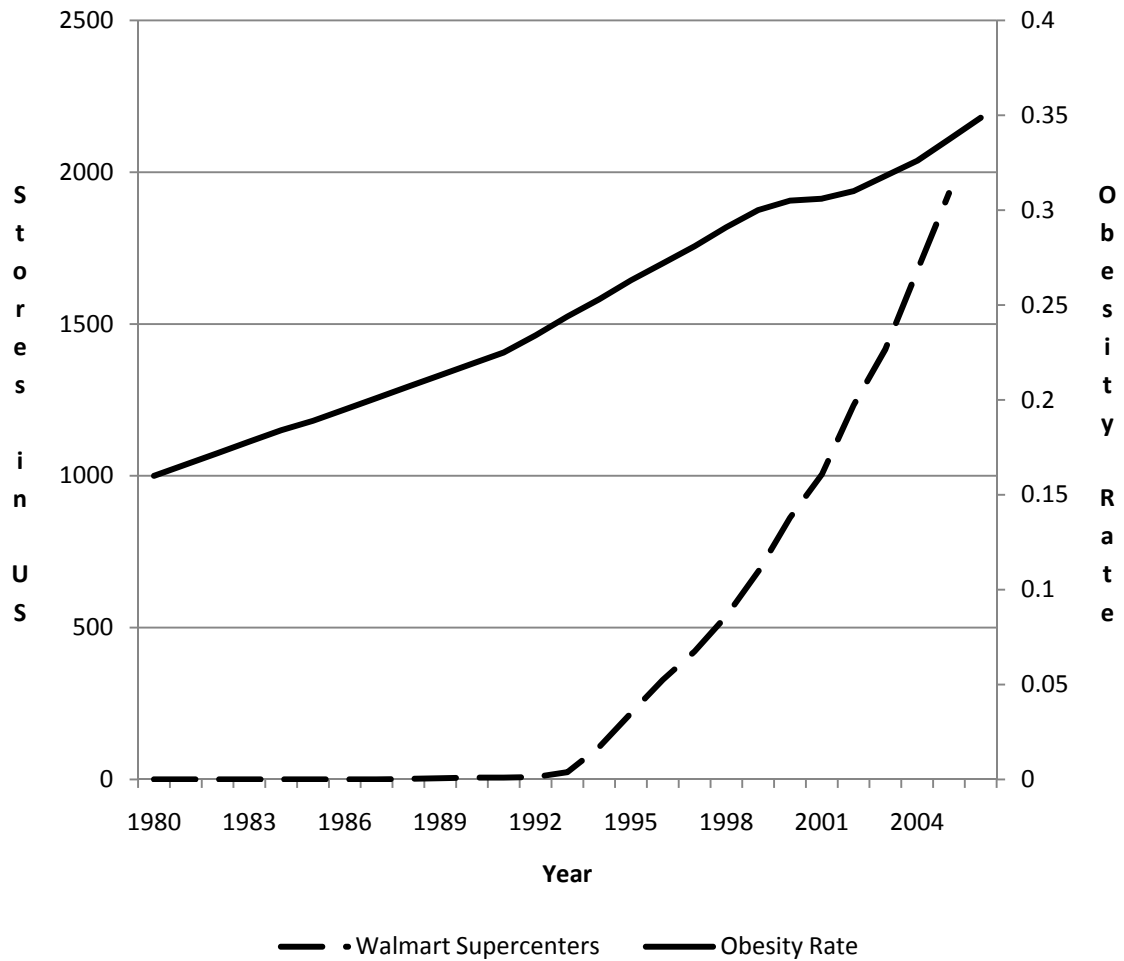
Notes: Obesity rates are from Flegal et al. (1998), Ogden et al. (2006), and National Center for Health Statistics (2008). Total numbers of stores in the United States are computed using our data.
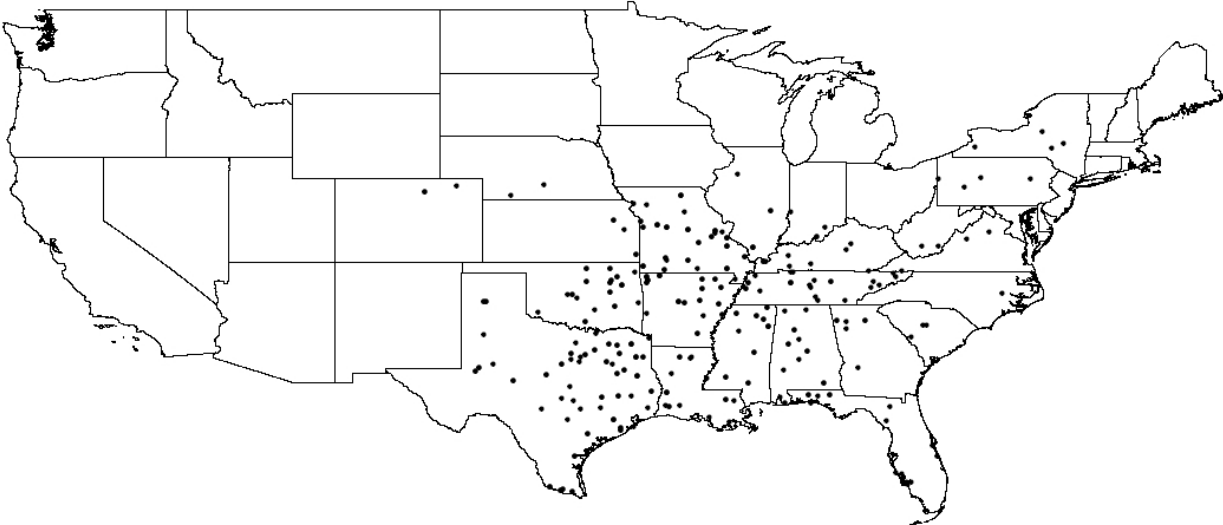
Figure 2—Walmart Supercenters, 1990
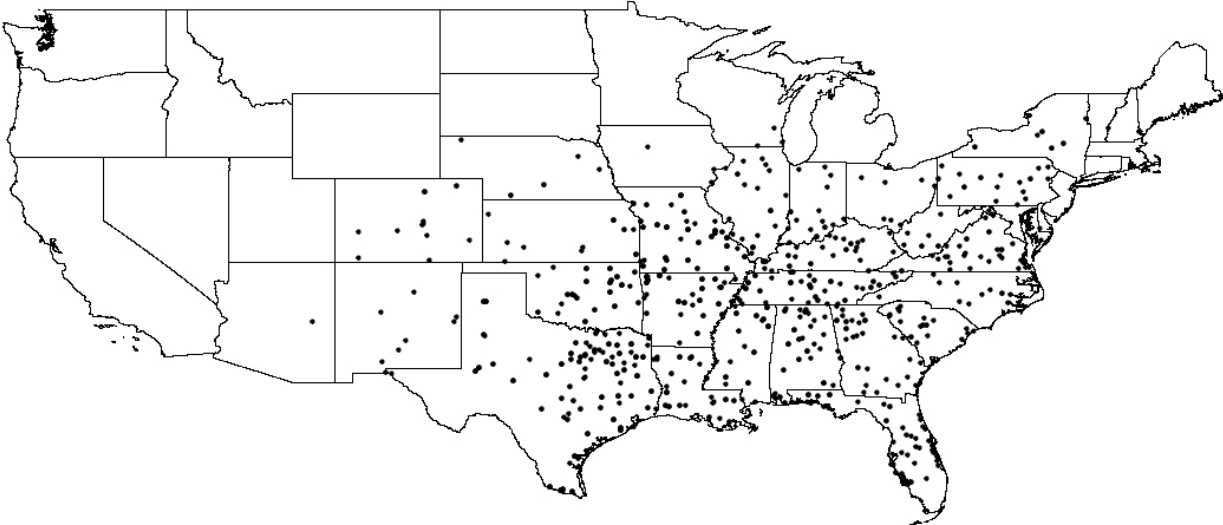
Figure 3—Walmart Supercenters, 1993

Figure 4—Walmart Supercenters, 1994

Figure 5—Walmart Supercenters, 1995

Figure 6—Walmart Supercenters, 1998
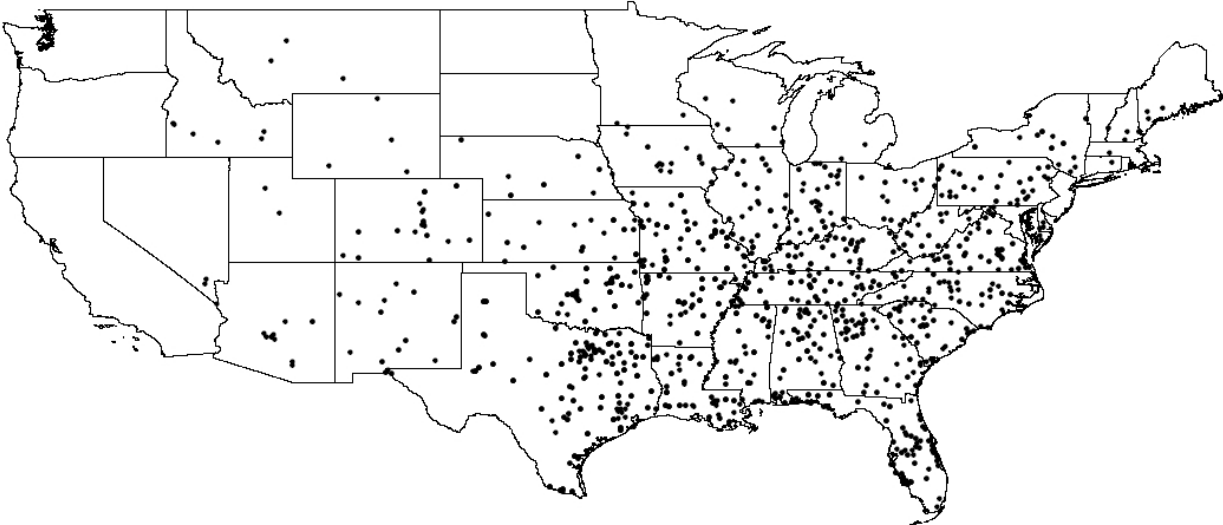
Figure 7—Walmart Supercenters, 2000

Figure 8—Walmart Supercenters, 2003
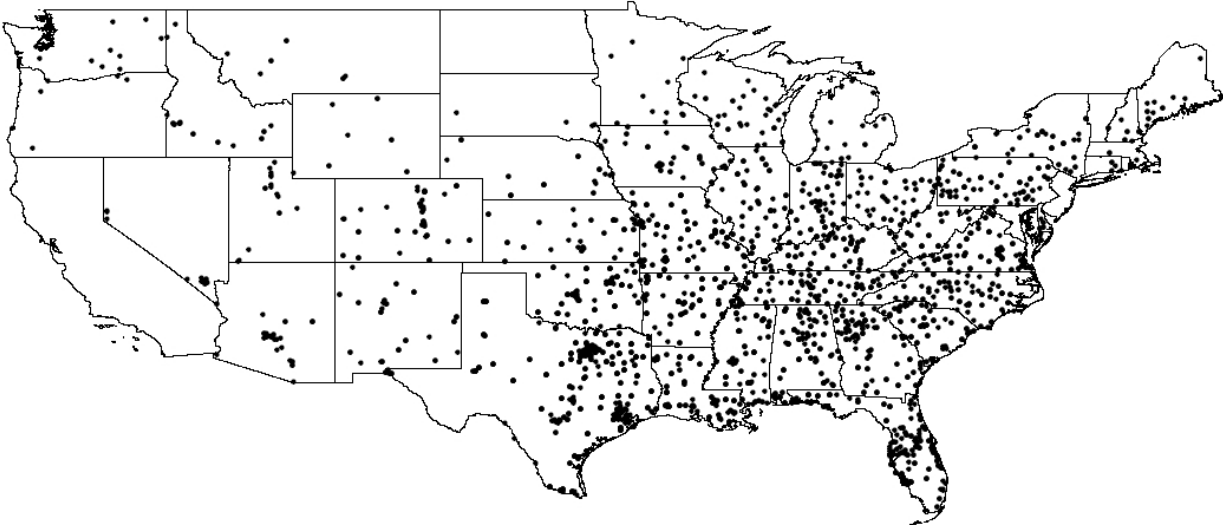
Figure 9—Walmart Supercenters, 2005
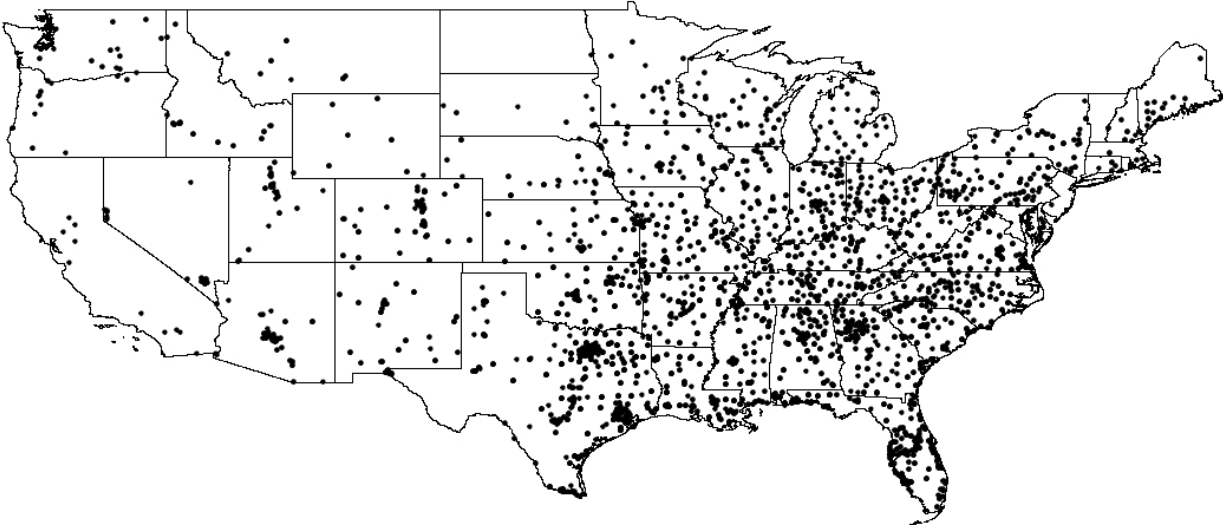
Figure 10—Walmart Supercenters, July 2009
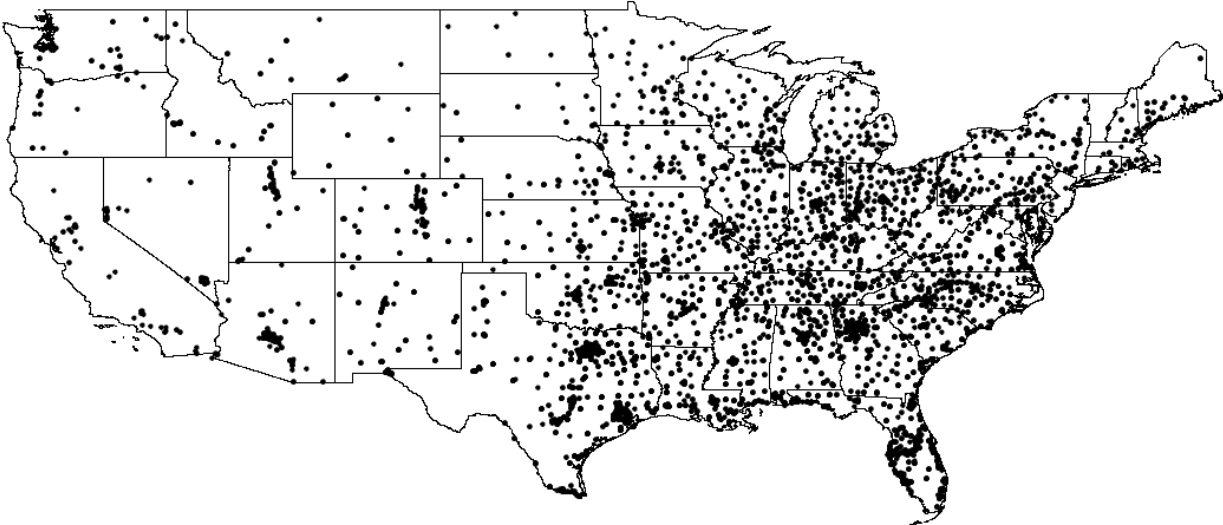
Table 1 – Theoretical Effects of Healthy, Unhealthy, and Restaurant Food Prices on Healthy, Unhealthy, and Restaurant Food Consumption

Price Change	Food Type	Substitution Effect	Income Effect	Total Effect
$P_H \uparrow$	H	↓	↓	↓
$P_H \uparrow$	U	↑	?	?
$P_H \uparrow$	R	↑	↓	?
$P_U \uparrow$	H	↑	↓	?
$P_U \uparrow$	U	↓	?	?
$P_U \uparrow$	R	↑	↓	?
$P_R \uparrow$	H	↑	↓	?
$P_R \uparrow$	U	↑	?	?
$P_R \uparrow$	R	↓	↓	↓

Table 2 – Summary Statistics

	Variable Name	Description	Mean (Std. Dev.)
BRFSS variables	BMI	Respondent's body mass index	26.612 (5.358)
	Obese	1 if BMI≥30 and 0 otherwise	0.212 (0.409)
	Married	1 if married	0.600 (0.490)
	Some high school	1 if attended some high school but did not graduate	0.075 (0.264)
	High school	1 if graduated high school but obtained no further education	0.308 (0.462)
	Some college	1 if attended some college but did not graduate	0.277 (0.448)
	College	1 if graduated from college	0.301 (0.459)
	Female	1 if female	0.496 (0.500)
	Race: black	1 if race is black	0.096 (0.295)
	Race: other	1 if race is neither white nor black	0.161 (0.367)
	Real income	Household income in 2005 dollars	52236 (27918)
	Age	Age in years	44.774 (16.888)
Matched county-level variables	Supercenters	Walmart Supercenters per 100,000 residents	0.333 (0.793)
	Distance	Distance from Bentonville, AR	852.242 (392.532)
	Distance from FDC	Distance from nearest Walmart food distribution center	337.962 (353.921)
	Discount stores	Walmart discount stores ever entered per 100,000	0.918 (1.146)
	Sam's Clubs	Sams Clubs per 100,000	0.190 (0.268)
	Population	Population in units of 100,000	9.376 (18.763)
	Unemployment	Unemployment rate	5.126 (2.011)
	Grocery	Grocery stores per 100,000 residents	55.689 (27.139)
	General	General merchandise stores (besides Supercenters) per 100,000	14.601 (9.081)

Note: Observations are weighted using the BRFSS sampling weights.

Table 3 – Effect of Distance-Year Interactions on Walmart Supercenters per 100,000 Residents (First Stage)

	Supercenters
Distance*1997	-0.00007
	(0.00001)***
Distance*1998	-0.00012
	(0.00002)***
Distance*1999	-0.00019
	(0.00002)***
Distance*2000	-0.00029
	(0.00003)***
Distance*2001	-0.00038
	(0.00003)***
Distance*2002	-0.00046
	(0.00004)***
Distance*2003	-0.00054
	(0.00004)***
Distance*2004	-0.00063
	(0.00005)***
Distance*2005	-0.00077
	(0.00006)***
F statistic	21.75
Observations	1,644,094

Notes: Standard errors are in parentheses. *** indicates statistically significant at the 1% level; ** 5% level; * 10% level. Standard errors are heteroskedasticity-robust and clustered by county. County and year fixed effects are included, as well as the individual-level control variables. Observations are weighted using the BRFSS sampling weights.

Table 4 – Effects of Walmart Supercenters on BMI and P(Obese) (Second Stage)

	BMI		Obese	
	OLS	IV	OLS	IV
Supercenters	0.029	0.252	0.001	0.024
	(0.014)**	(0.086)***	(0.001)	(0.005)***
Hausman test	-	0.223	-	(0.022)
		(0.085)***		(0.005)***
Overidentification test	-	0.167	-	0.578
Observations	1,644,094	1,644,094	1,644,094	1,644,094

See notes for Table 3.

Table 5 – Robustness Checks

		BMI	Obese
Panel A: Alternate functional form for Supercenter variable	Supercenter variable scaled by land area	0.651 (0.230)***	0.061 (0.017)***
	Unscaled Supercenter variable	0.091 (0.043)**	0.009 (0.004)**
	Binary Supercenter variable	0.244 (0.251)	0.034 (0.017)**
Panel B: Alternate functional form for instruments	Quadratic distance	0.250 (0.085)***	0.023 (0.005)***
	10 distance rings (Dube et al.)	0.239 (0.081)***	0.021 (0.005)***
	Distance rings for 100 miles (Neumark et al.)	0.267 (0.078)***	0.023 (0.005)***
Panel C: Sam's Clubs and discount stores	Control for Sam's Clubs	0.261 (0.087)***	0.024 (0.005)***
	Control for discount stores	0.263 (0.096)***	0.025 (0.006)***
	Drop if Sam's Club entry (n=1,371,068)	0.385 (0.083)***	0.029 (0.006)***
	Drop if discount store entry (n=1,247,730)	0.405 (0.129)***	0.033 (0.009)***
Panel D: Additional robustness checks	Population density*year	0.286 (0.101)***	0.023 (0.007)***
	Unemployment rate*year	0.280 (0.081)***	0.026 (0.005)***
	Grocery stores*year	0.219 (0.089)**	0.023 (0.006)***
	General merchandise stores*year	0.208 (0.103)**	0.018 (0.007)***
	Drop closest 10% of counties (n=1,510,081)	0.224 (0.095)***	0.023 (0.006)***
	Drop furthest 10% of counties (n=1,205,578)	0.347 (0.156)**	0.033 (0.011)***
	Drop if population>1,000,000 (n=1,499,433)	0.278 (0.083)***	0.025 (0.006)***
	Ring-specific trends	0.228 (0.183)	0.034 (0.014)**
	Census-division*year	0.328 (0.141)**	0.029 (0.008)***
	Distance from FDC	0.269 (0.150)*	0.014 (0.009)
	Distances from Bentonville and FDC	0.298 (0.095)*** [0.234]	0.022 (0.006)*** [0.572]

Notes: Sample size is 1,644,094 unless otherwise indicated. P-value from the overidentification test is in brackets. See other notes for Table 2.

Table 6 – Effects of Walmart Supercenters on BMI and P(Obese) for Subsamples

		BMI	P(Obese)
Gender	Women	0.337	0.019
		(0.107)***	(0.007)***
	Men	0.196	0.029
		(0.115)*	(0.008)***
Income: Married	Lowest 25%	0.391	0.020
		(0.190)**	(0.015)
	Middle 50%	0.282	0.017
		(0.137)**	(0.012)
	Highest 25%	0.211	0.016
		(0.144)	(0.012)
Income: Single	Lowest 25%	0.125	0.026
		(0.231)	(0.015)*
	Middle 50%	0.188	0.027
		(0.185)	(0.015)*
	Highest 25%	0.207	0.017
		(0.262)	(0.021)
Population	Smallest 25%	0.696	0.062
		(0.384)*	(0.031)**
	Middle 50%	0.272	0.022
		(0.121)**	(0.009)**
	Largest 25%	0.210	0.023
		(0.252)	(0.016)

See notes for Table 3.

Appendix A – Savings from Wal-Mart

In this appendix, we perform a back-of-the-envelope calculation of consumers' savings from Walmart Supercenters' lower prices. These savings (SAV) are equal to what expenditures per household on consumer goods and groceries (the goods Supercenters sell) would be if there were no Supercenters (C) minus what they are given that there are Supercenters (C_W):

$$SAV = C - C_W \qquad (16)$$

C_W is identical to C for the portion of the population that lives in areas with no Supercenters. People who live in areas with at least one Supercenter can further be divided into those who shop there and those who do not. Those who do not shop at Walmart Supercenters still face lower prices if Walmart affects competitors' prices. People who do shop at Walmart experience this reduction, plus an additional one due to Walmart's price advantage over competitors that remains even after competitors adjust their prices. Therefore,

$$C_W = P_W[C(P_{WW})(1 - PCT_C)(1 - PCT_W) + C(1 - P_{WW})(1 - PCT_C)] + C(1 - P_W) \quad (17)$$

where P_W is the proportion of people who live in a county with at least one Walmart Supercenter, P_{WW} is the proportion of people living in a county with a Supercenter who shop there, PCT_C is the percentage reduction in competitor prices after Supercenter entry (divided by 100), and PCT_W is the percentage savings from shopping at a Supercenter instead of its competitors. Combining (16) and (17) yields

$$SAV = C - P_W[C(P_{WW})(1 - PCT_C)(1 - PCT_W) + C(1 - P_{WW})(1 - PCT_C)] - C(1 - P_W).$$

$$(18)$$

We perform the calculation for 2002 because estimates for all the parameters are available in that year. For C, we first compute total expenditures on the types of goods Supercenters sell using data from the 2002 Census of Retail Trade. We consider Supercenters to

sell goods in the following categories: electronics and appliances; lawn and garden; food and beverage (not counting restaurants or bars); health and personal care; clothing and clothing accessory; sporting goods, hobby and musical instrument; book, periodical, and music; general merchandise; and office supplies and gifts. Expenditures in these categories totaled approximately $1.474 trillion, or 14% of 2002's GDP.[27] C is equal to this number divided by the number of households in the U.S. in 2002 according to the U.S. Census Bureau, 109.297 million. C is therefore $13,488.

For P_W, we use the proportion of people from the 2002 wave of the BRFSS who live in a county with at least one Supercenter, which is 0.37. For P_{WW}, we begin by dividing Walmart Supercenters' domestic sales by $1.474 trillion to determine its approximate market share. We compute a market share of 5.2%.[28] To generate a market share of 5.2% for the entire U.S., which includes counties with no Supercenters where their market share should be 0%, the market share must be 14.1% in the counties that do have at least one Supercenter. We therefore assign $P_{WW} = 0.141$.

For the price reduction parameters, we draw on existing literature. Basker and Noel (2007) find the entry of one or more Walmarts in a city reduces competitors' prices by 1-1.2%. We assign $PCT_C = 0.011$, the midpoint of this range. As discussed in Section 2, studies place Walmart's price advantage over competitors at between 8% and 27% (Hausman and Leibtag, 2004). We therefore assign PCT_W equal to the midpoint, or 0.175. Equation (18) becomes

---

[27] Of course, Wal-Mart sells some goods outside of these categories and sells a limited selection of some of the goods within these categories, so our approach provides only a rough estimate. According to the web site of the St. Louis Federal Reserve Bank, nominal GDP from July 1, 2002 was $10527.4 trillion.

[28] Walmart's 2002 domestic sales at discount stores and Supercenters totaled $139.131 billion according to its annual report, available at http://walmartstores.com/Media/Investors/2002_annualreport.pdf. We approximate the sales at Supercenters by multiplying $139.131 billion by the proportion of Walmart stores that were Supercenters, weighted by the average square footage of discount stores and Supercenters. We obtain the average square footage from http://en.wikipedia.org/wiki/Wal-Mart. Ultimately, we estimate total sales at Supercenters of $77.297 billion.

$$SAV = 13488 - 0.37[13488 * (0.141)(1 - 0.011)(1 - 0.175) + 13488 * (1 - 0.141)(1$$

$$- 0.011)] - 13488(1 - 0.37) = \$176.68. \tag{19}$$

We therefore conclude that Wal-Mart Supercenters saved households $177 per year in 2002.