Effects of Urban Sprawl On Obesity

Zhenxiang Zhao

Robert Kaestner

July 2010

Zhenxiang Zhao
University of Illinois at Chicago
1747 West Roosevelt Road
Chicago, IL 60608
zzhao5@uic.edu

Robert Kaestner
Institute of Government and Public Affairs
University of Illinois
815 West Van Buren Street, Suite 525
Chicago, IL 60607
kaestner@uic.edu

The authors thank Nathan Anderson, Eric Finkelstein, Evelyn Lehrer, Lee Mobley, Daniel McMillen, Lisa Powell and two anonymous reviewers for helpful comments.

Abstract

In this paper, we examine the effect of changes in population density—urban sprawl—between 1970 and 2000 on BMI and obesity of residents in metropolitan areas in the US. We address the possible endogeneity of population density by using a two-step instrumental variables approach. We exploit the plausibly exogenous variation in population density caused by the expansion of the U.S. Interstate Highway System, which largely followed the original 1947 plan for the Interstate Highway System. We find a negative association between population density and obesity, and estimates are robust across a wide range of specifications. Estimates indicate that if the average metropolitan area had not experienced the decline in the proportion of population living in dense areas over the last 30 years, the rate of obesity would have been reduced by approximately 13%.

Introduction

Over the past forty years, the prevalence of obesity in the U.S. has increased dramatically. Between the early 1960s and 2004, the percentage of U.S. adults who were obese more than doubled from 13.4% to 32.2% (Flegal at al., 2002; Ogden et al., 2006). This dramatic increase is disconcerting because obesity has been linked to several diseases including type 2 diabetes, hypertension, high cholesterol, and asthma (Kopelman, 2000; Koplan et al., 1999; Peeters et al., 2003; Wellman et al., 2002). Further, it has been estimated that obesity contributed to between 100,000 and 300,000 deaths in just one year—2000 (Mokdad et al., 2004 and 2005; Flegal et al., 2005). Finally, obesity-related morbidity has been estimated to account for a significant share of total annual medical expenditures in the US.[1]

The spatial distribution of the population in the U.S. has also changed markedly over the same period. Between 1950 and 2000, the share of the population living in metropolitan areas grew from 56% to 80% (Transportation Research Board Special Report 282, 2005). While a greater proportion of the population is now living in urban areas broadly defined, all of the growth in metropolitan areas has occurred in suburban areas, as central cities actually declined in population (Baum-Snow, 2007). In 1950, the population of metropolitan areas was roughly evenly divided between the suburban fringe and the central city; currently, approximately two-thirds of the population of metropolitan areas resides in the suburbs, and this proportion has been rising (Pisarski, 2001). Table 1 shows that between 1970 and 2000 the population density for 53 major metropolitan areas fell over 19% with more dramatic declines observed for the densest parts of metropolitan areas.

[1]The Surgeon General's report has estimated the total cost of obesity to the U.S. was $117 billion in 2000 (USDHHS, 2001).

The growth in obesity and decline in population density observed over roughly the same periods have led researchers to investigate the association between these two trends more fully, and to assess whether the association is causal. Urban sprawl, characterized by low-density development patterns and changes in the built environment, has been found to be positively associated with rates of obesity, although the evidence is not uniform (Ewing, 2003; Frank et al., 2004; Lopez, 2004; Vandegrift and Yoked, 2004; Platinga and Bernell, 2007; Eid et al., 2008; Black and Macinko 2008; Moon 2009).[2]

Several explanations of this association have been proposed. Suburban residential location increases the distance between home and destination (e.g., job), increases the reliance on automobiles, and minimizes walking. Lack of sidewalks and bicycle trails, and the cul-de-sac street layouts that are typical in suburban areas may decrease physical activity (Cervero et al, 1995; Handy, 1996; Hess et al., 1999; Crane et al., 1998; Boarnet et al., 2000; Saelens et al., 2003, Frank, 2000; Berrigan et al., 2002). In addition, greater presence of large retail stores (i.e., big box stores) in the suburbs leads to lower costs of food in suburban areas due to economies of scale and relatively cheaper land prices. Lower food prices have been found to be correlated with dietary intake patterns (French, 2005; French et al. 1997, 2001; Jeffery et al., 1994; Powell et al. 2007a; Beydoun et al., 2008; Larson et al., 2009). Offsetting these aspects of urban sprawl that have been hypothesized to increase obesity of suburban dwellers, is the greater availability of healthy foods in large suburban stores that may decrease obesity (Sallis et al., 1986; Horowitz et al., 2004; Jetter et al., 2006; Chung et al., 1999).

[2] In this paper, urban sprawl is defined population density and the two terms are used interchangeably. Urban sprawl is a loosely defined term, which is a point frequently noted by urban planners. Commonly referred to aspects of urban sprawl include low population density, dispersed employment, separate zones of land use (e.g., retail here and housing there), and dependence on automobile for travel. Most researchers interested in studying the extent and consequences of urban sprawl have used population density as part of a measure of urban sprawl. Ideally, urban sprawl would be defined by more detailed measures of land use and characteristics of the built environment. However, these data are not available for the years and places we include in the analysis.

Urban sprawl may also affect the weight of central city residents. Land in densely populated cities is not sufficient to accommodate parking and other structures of large grocery stores, which are more likely to carry fresh fruit and vegetables. Thus, access to healthy foods is decreased in central cities. Instead, central cities are more likely to be characterized by small grocery stores and a higher density of fast-food outlets sometimes resulting in what has been referred to as "food deserts" in which poor urban residents cannot buy affordable, healthy food (Cummins and Macintyre 2002; Ignami et al., 2006; Feldstein, 2007; Ford et al., 2008). Finally, urban sprawl is associated with a concentration of poverty and higher crime rates in some parts of the central city, which makes outdoor activities more dangerous and may limit opportunities for physical activity (Lumeng et al., 2006; Stafford et al., 2007; Mujahid et al., 2008.) All of these changes are hypothesized to result in an increase in obesity.

Results from previous studies of the association between urban sprawl and obesity have led the Centers for Disease Control and Prevention (2003), the World Health Organization (2004) and others (e.g., Smart Growth America) to advocate using community (re)design as a tool to curb obesity. For example, the CDC recommends that communities should: "improve geographic availability of supermarkets in underserved areas"; "improve access to outdoor recreational areas"; "enhance infrastructure supporting biking"; "enhance infrastructure supporting walking"; "support locating schools within easy walking distance of residential areas"; "zone for mixed use development"; and "enhance personal safety in areas where persons are or could be physically active".[3] Although policy makers are calling for action to combat urban sprawl, a key policy question, which previous research has not adequately addressed, is to what extent urban sprawl causes people to be obese.

[3] http://www.cdc.gov/mmwr/preview/mmwrhtml/rr5807a1.htm, website last accessed May 10, 2010.

While there have been several studies that have examined the associations between urban sprawl and obesity, these studies have been limited in their ability to provide estimates of a causal relationship. The primary limitation of previous studies is that the relationship between urban sprawl and obesity is likely to be confounded by non-random selection of residents into neighborhoods; factors that affect where a person live may also affect their weight. There are many possibilities. For example, those who live in the suburbs because of a preference for large houses that are facilitated by relatively cheaper land prices may also prefer to eat healthier, exercise and maintain a healthier weight. Moreover, preferences of these suburbanites will attract providers of healthy foods and cause local governments to provide amenities consistent with such preferences. In this example, urban sprawl will be associated with lower weight. However, if we could account for unmeasured preferences, urban sprawl—dispersed employment, separate zones of land use (e.g., retail here and housing there), and dependence on automobile for travel—may have increased obesity. The upshot is that any observed association between urban sprawl and obesity may reflect selection rather than a causal relationship in which sprawl causes obesity. Importantly, identifying the causal effect, and not just the association between urban sprawl and obesity, is essential for designing effective public policy. If the association between urban sprawl and obesity is due to selection, then policies such as those promoted by the CDC and others aimed at curbing urban sprawl would have little effect on obesity.

Four previous studies have attempted to address the selection issue (Eid et al., 2008; Plantinga and Bernell, 2007a; Plantinga and Bernell, 2007b; Ewing et al., 2006). Three of these studies used fixed-effect methods with longitudinal data on individuals from the National Longitudinal Survey of Youth (1979 and 1997 cohorts). Findings from these studies were

mixed; Plantinga and Bernell (2007a) found that sprawl lowered BMI, and Eid et al. (2008) and Ewing et al. (2006) found no statistically significant associations between sprawl and BMI.[4]

While these studies arguably improve over cross-sectional analyses that ignore selection, there are some notable limitations. First, these studies use samples with narrow age ranges that limit the applicability of their findings. Eid et al (2008) used a sample of persons between the ages of 23 and 36; Ewing et al. (2006) used a sample of young adults ages 18 to 23; and Platinga and Bernell (2007a) used a sample of persons ages 33 to 41. Second, while fixed-effects methods eliminate time-invariant factors, there is the possibility that omitted time-varying factors will bias estimates, particularly given the young cohort used in these studies. It is noteworthy that observed time-varying factors such as marital status and work have relatively large and statistically significant effects on BMI in the Eid et al.(2008) study. This raises the possibility that omitted time-varying factors may also be important confounders. Third, as noted above, urban sprawl may affect the weight of both central city residents and suburban residents, so studying the effects of moving from places that are characterized by more or less sprawl may obscure the effects of sprawl. Changes over time and the growth of urban sprawl may have affected the "obesity environment" in both types of places, and thus, studies of moving from one type of place to the other misses the effect of sprawl. However, observing changes over time in both places due to sprawl, as we do, will reveal an effect of sprawl on obesity.

Plantinga et al. (2007b) use instrumental variables to address selection. Instruments for their measure of urban sprawl are marital status and family size. The validity of instrument is questionable because it is likely that marital status and family size are correlated with BMI. For

[4] Eid et al. (2008) use proportion of undeveloped land in the square kilometer around residence as one of measures of sprawl, and the other two papers use sprawl index at county level where population density is one of the components.

example, single people may be less likely to be obese because they are active in the marriage market (Helmchen 2004).

In this paper, we take a fundamentally different approach to the problem than the previous literature. We examine the effect of changes in population density, which is almost universally used, at least in part, as a measure of urban sprawl, between 1970 and 2000 on BMI and obesity of residents in metropolitan areas in the US. The long time period is an advantage because it incorporates large changes in population density and the built environment in both central cities and suburbs. It is these changes that are most directly tied to the debate over the use of community (re)design to combat obesity. We address the possible endogeneity of population density by using a two-step, instrumental variables approach. We exploit the plausibly exogenous variation in population density (urban sprawl) caused by the expansion of the U.S. Interstate Highway System, which largely followed the original 1947 plan for the Interstate Highway System. We use the original 1947 plan for the Interstate Highway System to instrument for population density.

We find a negative association between the proportion of an MSA that lives in dense areas and obesity, and estimates are robust across a wide range of specifications. Estimates indicate that if the average metropolitan area had not experienced the decline in the proportion of population living in dense areas over the last 30 years, the rate of obesity would have been reduced by approximately 13%. Sub-group analyses show that those living in the central city seem to experience slightly greater changes in their weight status compared to their counterparts living in suburbs, although these results are more tentative.

Research Design

The goal of the empirical analysis is to obtain estimates of the association between the proportion of a metropolitan area living in dense areas (i.e. census tracts), which we use as a proxy for urban sprawl and the built environment, and the weight of residents in that area. We use the following model specification:

$$BMI(OBESE)_{ijt} = \alpha_j + \gamma_t + \beta PopDen_{jt} + \phi X_{ijt} + \lambda Z_{jt} + \nu_{ijt} \quad (1)$$

$i = 1,...,N$ (persons)

$j = 1,...,J$ (MSAs)

$t = 1976,...,2001$ (years)

In equation (1), the body mass index (BMI), or obesity (BMI>30), of person i in MSA j and year t depends on MSA fixed effects (α_j); year fixed effects (γ_t); individual characteristics (X) such as age, race, sex, marital status and education; MSA population density ($PopDen$); and time-varying MSA level characteristics (Z_{jt}) such as median family income, education and employment rates.[5] Population density is measured as the proportion of the population living in dense areas where different thresholds are used to define dense areas to capture the degree to which populations are centered in high density living areas. While not a perfect measure of urban sprawl, this measure is consistent with previous literature and more sophisticated measures (Ewing et al., undated).

Ideally, we would like to interpret estimates of the association between population density and weight obtained from equation (1) as causal, and toward this goal, we have included controls for potentially confounding influences (e.g., MSA fixed effects and time-varying MSA characteristics). However, there still may be omitted MSA-level factors that are correlated with population density and weight of residents. For example, changes over time in the education, family structure, age, and income distributions of residents in an MSA may result in different

[5] All fixed effects are specified with an omitted group as reference category, for example 1976 is omitted year.

land use regulations, development patterns and rates of obesity. In this case, urban sprawl is not a cause of obesity, but rather the association between sprawl and obesity is due to confounding from omitted resident characteristics. To address this problem, we use a two-step instrumental variables approach. Specifically, we use the original 1947 Interstate Highway System plan to predict population density and use this predicted measure of population density in equation (1). The equation used to predict population density is:

$$PopDen_{jt} = \rho_1 TIME_t + \rho_2 TIME_t^2 + \pi[(YEAR_t - 1947) * HWPLAN_j] + \sigma Z_{jt} + \mu_j + \xi_{jt} \qquad (2)$$

$j = 1, \ldots, J$ (MSAs)

$t = 1970, 1980, 1990, 2000$ (years)

In equation (2), population density of MSA j in year t depends on MSA fixed effects (μ_j), time-varying MSA level characteristics (Z_{jt}), a quadratic time trend (TIME and TIME^2), and the number of planned highway rays from the 1947 Interstate Highway System plan in MSA j ($HWPLAN_j$) multiplied by the number of years since 1947. A planned highway passing through the central city counts as 2 rays. For those planned highways ending in or near the central city, one ray is counted. Note that we use aggregate data for the four decennial census years to estimate equation (2), and that individual characteristics (X in equation 1) are not included in equation (2).[6] We are not using a two-stage least squares procedure, but a two-step instrumental variables procedure. The underlying logic of this approach is that the highway plan originating in 1947 was to be implemented over time and therefore population density of an MSA would be affected in a diffused way over time. This diffusion process is captured by the specification of equation (2) by the term [(Year$_t$-1947)*HWPLAN$_j$], which allows the effect of the highway plan to vary over time in a linear fashion. Other specifications (e.g., quadratic) of the diffusion

[6] The purpose of equation (2) is to derive a predicted measure of population density for each year between 1976 and 2001. Therefore, we use a quadratic time trend instead of dummy variables for each of the four Census years.

process were tried and statistical tests could not reject the linear specification, nor did estimates of interest differ in these alternative specifications.

Models similar to equation (2) have been estimated by others. Specifically, Baum-Snow (2007) showed that the greater the number of planned highways in a metropolitan area, the greater the loss of central city population over the last half of the 20th century. This paper finds similar effects with respect to population density, which are presented below. Thus, in assessing the validity of the instrumental variables approach, we have a sufficiently strong first stage association. The interstate highway system is associated with a significant change in the spatial distribution of the population with an MSA.

The next question related to the validity of the instrumental variables approach is whether the use of the 1947 plan meets the exclusion criterion. We believe it is reasonable to think it does. The original Interstate Highway System plan of 1947 was motivated by concerns related to national defense, was designed to connect faraway places and not intended to facilitate local commuting. The Federal-Aid Highway act of 1944 called for designation of a National System of Interstate Highways, to include up to 40,000 miles "…so located as to connect by routes, as direct as practicable, the principal metropolitan areas, cities, and industrial centers, to serve the national defense, and to connect at suitable border points with routes of continental importance in the Dominion of Canada and the Republic of Mexico." On August 2, 1947, Commissioner of Public Roads Thomas H. MacDonald and Federal Works Administrator Philip B. Fleming announced the selection of the first 37,700 miles. The routes had been proposed by state highway agencies and reviewed by the Department of Defense to meet the needs of national defense. Therefore, it is plausible that, conditional on controlling for MSA fixed effects (initial conditions that may have influenced 1947 plan), the original planned number of highways is exogenous—

uncorrelated with unmeasured determinants of changes in obesity within an MSA. Some evidence that this is in fact the case is that the inclusion of (relevant) time-varying MSA characteristics has little effect on instrumental variables estimates even though these variables are often significant predictors of obesity and BMI. Nevertheless, we recognize that it is difficult to obtain valid instrumental variables estimates and that our findings need to be interpreted in this light.

Data

A variety of data sources were used for the analysis. Individuals' demographic and socioeconomic information, weight and height information, and metropolitan area of residence were obtained from the National Health Interview Survey (NHIS) for each year from 1976 to 2001. The normalized census tract level population data were obtained from the Neighborhood Changing Data Base for years 1970, 1980, 1990 and 2000. The planned number of highway rays for metropolitan areas was calculated from the General Location of National System of Interstate Highways (1956). Median family income, education and employment rates at metropolitan areas level came from the Current Population Survey (CPS) March file from 1976 to 2001.

NHIS

The NHIS for the years 1976 to 2001 was used to obtain demographic information, weight, height and metropolitan area of residence.[7] We limited the sample to persons age 18 and over. The NHIS provided the following information: age; race (four race/ethnicity categories: non-Hispanic white, non-Hispanic black, Hispanic and others); sex; education (five education categories describing the highest grade individual completed: elementary school, some high school, high school graduate, some college and college graduate); income (0-4,999, 5-9,999, 10-

[7] We used 1976 as the first year of the sample because 1976 is the first year NHIS has individual's weight and height information in the core data set. We used 2001 as the ending year because the construction of the interstate highway was almost finished by 2001 and it is the last year that NHIS public use data have MSA identifiers.

14,999, 15-24,999, 25,000 or more, and missing); marital status (four marital status categories: single, married, separated/divorced, and widowed); and metropolitan area of residence. Definitions, means, and standard deviations of all variables employed in the NHIS are available from authors' home page.

The metropolitan area is the lowest geographic identifier available in the public-use NHIS data.[8] There are two points worth mentioning regarding the metropolitan area identifier. First, the NHIS only identifies large metropolitan areas and the number of metropolitan areas identified in the NHIS increased approximately every ten years.[9] For the years 1976-1984, 31 metropolitan areas (SMSA/SCAs) were identified. Between 1985 and 1994, 33 metropolitan areas and 16 sub-areas (PMSAs) were identified. For the years 1995-1996, 52 metropolitan areas and 32 sub-areas were identified. Finally, for the years 1997-2001, 45 metropolitan areas and 19 selected sub-areas were identified. We excluded metropolitan areas that appeared only in 1995-96 and those with no central city.[10] Honolulu, HI and Anchorage, AK are also excluded from the sample as Hawaii and Alaska became U.S. states in 1959 and no interstate highway was planned in the original 1947 highway plan. This leaves 53 unique metropolitan areas in the final sample: 31 of them were identified in each year between 1976 and 2001; an additional 7 were

[8] Respondents who either live in non-MSA areas or not self-representing MSA areas are excluded from the estimation sample due to lack of their geographical location information.

[9] Metropolitan areas are defined by the U.S. Office of Management and Budget and definitions change approximately every ten years basedon Census data. The general concept of a metropolitan area is that of a core area containing a substantial population nucleus, together with adjacent communities having a high degree of economic and social integration with that core. Changes in definitions of metropolitan areas since the 1950 Census have chiefly resulted from the recognition of new metropolitan areas when the requirement on population was reached; the addition of counties to existing metropolitan areas; transfer of counties from one area to another; and dropping of counties from an area due to changes in population or the economic or social tie to the central counties of metropolitan areas. The large majority of changes have taken place on the basis of Decennial Census data.

[10] Those metropolitan areas which are only identified for two years (1995, 1996) in NHIS are: Dayton, OH MSA; Ventura, CA PMSA; Vineland, NJ PMSA; Wilmington, DE PMSA; Richmond, VA MSA; Santa Cruz, CA PMSA; Santa Rosa, CA PMSA; Albany, NY MSA; Bridgeport, CT PMSA; Stamford, CT PMSA; Danbury, CT PMSA; Jersey City, NJ PMSA; Dutchess County, NY PMSA; New Haven, CT PMSA; New Burgh, NY PMSA; Trenton, NJ PMSA; Waterbury, CT PMSA; Birmingham, AL MSA; and Vallejo-Fairfield-Napa, CA PMSA. Those metropolitan areas that have no central cities are also excluded from the sample: Bergen-Passaic, NJ PMSA; Middlesex, NJ PMSA; Nassau-Suffolk, NY PMSA.

identified in the years between 1985 and 2001; and another 15 were identified from 1995 to 2001.

Because of historical changes in geographic definitions, caution must be taken in comparing data for these statistical areas from different years. For example, most metropolitan areas encompass less territory during earlier years than in later years, and those newly included areas are generally less densely populated than those already included areas. If, for some unobserved reasons, people living in less dense areas are more or less likely to be obese, changing definition alone would lead to a non-zero association between population density and obesity.

To address this issue, we include a MSA fixed effect for each unique definition of an area. Thus, equation (1) is modified as follows:

$$BMI(OBESE)_{ijt} = \alpha_{j1} + \alpha_{j2} + \alpha_{j3} + \gamma_t + \beta PopDen_{jt} + \phi X_{ijt} + \lambda Z_{jt} + v_{ijt} \qquad (1`)$$

$j1 = 1,...,31$ (MSAs - 1971 definition)

$j2 = 1,...,38$ (MSAs - 1983 definition)

$j3 = 1,...,53$ (MSAs - 1993 definition)

$t = 1976,...,2001$ (years)

The important aspect of equation (1`) is that there are now three sets of MSA fixed effects; fixed effects are specific to three periods: 1976-1984 when MSA codes are based on the 1971 metropolitan area definition in NHIS, 1985-1994 when MSA codes are based on the 1983 metropolitan area definition in NHIS, and 1995-2001 when MSA codes are based on the 1993 metropolitan area definition in NHIS. We refer to the specification of equation (1`) as unrestricted MSA fixed effects.

Census Data From Neighborhood Changing Data Base

We used U.S. Censuses data from the Neighborhood Changing Data Base (NCDB) to calculate population density in 1970, 1980, 1990, and 2000 for each of the 122 unique MSA

definitions; for example, we calculated the population density in each of the four decennial census years for each of the three geographical definitions of the Atlanta MSA. Thus, there were 12 measures of population density related to Atlanta. To construct these figures, we used constant-geography (boundary) census tract data on population that were obtained from NCDB. We aggregated these census tract population data to the county level because MSAs consist of county groups and county definitions rarely change.[11] Then, we created a county-MSA cross walk file and aggregated the county level data to the MSA level using the historical metropolitan area definitions. In addition, MSA level median family income, employment rates, and education were also created at the constant-MSA level using NCDB data.

The population density of the MSA was measured as the proportion of the population living in census tracts with a specified threshold of population density, for example, the proportion of persons living in census tracts with a population density of 5,000 or more persons per square mile. Previous studies used various population density thresholds to measure the degree to which population is centered in high density living areas (Lopez, 2004; Ewing, 2003). We use several thresholds to define dense areas[12]: 5,000 or more people per square mile (about 50% of census tracts in the constant-geography MSAs have population density less than this), 9,000 or more people per square mile (about 75% of census tracts in the constant-geography MSAs have population density less than this), and 12,500 or more people per square mile, which is the lower limit of density needed to support mass transit (Ewing, 2003).

The changing geographical definitions of MSAs in the NHIS and the corresponding implications it has for measuring population density requires modification of equation (2) to

[11] No county was consolidated and relatively few counties were broken off. A few counties with changing boundaries were restored to form their original counties when matching counties to MSAs.

[12] The threshold of 3,500 or more people per square mile was initially adopted in the analysis to indicate when people begin to use non-automobile mode of transportation (Lopez, 2004). Analysis of the data showed that all of the decreases in the population density occurred in areas with density greater than 5,000 people per square mile.

incorporate the multiplicity of definitions. The modified specification is given by:

$$PopDen_{jt} = \rho_1 TIME_t + \rho_2 TIME_t^2 + \pi[(YEAR_t - 1947) * HWPLAN_j] + \sigma Z_{jt} + \mu_j + \xi_{jt} \qquad (2`)$$

$j = 1,...,122$ (MSAs)

$t = 1970, 1980, 1990, 2000$ (years)

Note that, in equation (2`), there is a fixed effect for each of the 122 unique definitions of MSA found in the NHIS. Thus, there are 488 observations: 122 MSAs by 4 Census years. Equation (2`) is used to predict the population density of each of the 1030 MSA-year combinations in equation (1`).[13]

Highway Plan Data

The Federal-Aid Highway Act of 1944 authorized designation of a 40,000-mile Interstate Highway System, but did not create a program to build it until 1947. On August 2, 1947, the Public Roads Administration designated approximately 37,700 miles of the nation's principal highways. The process was completed in September 1955 when the General Location of National System of Interstate Highways was published by the Bureau of Public Roads.

Based on information from the 1947 National System of Interstate Highways plan, we use the number of planned highway rays emanating from the largest central city of an MSA for the HWPLAN variable of equation (2). Although the authorized Interstate Highway System by the Federal-Aid Highway Act of 1944 is still under construction, virtually the entire 1947 planned Interstate Highway System was built by 2000 (FHWA, 2001).

Current Population Survey

The CPS March file was used to calculate (weighted) MSA level median family income, employment rates, and education level in an MSA that matched to NHIS individual level data

[13] In the NHIS, there are 31 MSAs in 1976-1984 period for a total of 279 MSA-year combinations. There are 38 MSAs in the 1985-94 period for a total of 380 MSA-year combinations. There are 53 MSAs in the 1995-2001 period for a total of 371 MSA-year combinations. Adding all totals yields 1030 MSA-year combinations.

from 1976-2001. Definitions of MSA identifiers in CPS also changed over time based on OMB's definitions. The CPS data were merged with the appropriate MSA definition.

Results

The first estimates we discuss are from equation (2`), which is used to predict population density using the 1947 highway plan. Table 2 presents these estimates. Estimates of the effect of the 1947 highway plan are negative and statistically significant. More planned highways in 1947 are significantly associated with a decrease in population density over time. In terms of magnitudes, an additional highway ray is associated with a 5% decrease every 20 years in the proportion of the population in an MSA that lives in census tracts with 5000 or more people per square mile. Larger effects are found for measures of greater population density; an additional highway ray is associated with a 9% decrease in the proportion of the population in an MSA that lives in census tracts with 9,000 or more people, and an additional highway ray is associated with a 10% decrease in the proportion of the population in an MSA that lives in census tracts with 12,500 or more people.

To assess whether estimates are sensitive to the inclusion of time-varying MSA characteristics, we estimated models with and without controls for median family income, employment rates, and education level (proportion of population with college degree of the MSA). Columns (3), (5) and (7) of Table 2 show these estimates. The inclusion of these controls has little effect on estimates of the effect of planned highway rays even though median family income has significant and positive effect on population density. These results are evidence in support of the validity of the instrumental variables approach because the instrument is uncorrelated with observed time-varying MSA characteristics.

We now turn the discussion to the estimates of the effect of population density on obesity

and BMI. Table 3 presents the estimates for obesity and Table 4 estimates for BMI. Estimates are obtained from the two-step, instrumental variables procedure. Standard errors have been constructed using methods that account for the predicted nature of population density and the potential non-independence (clustering) of observations within MSA-year (Murphy and Topel, 1985; Hardin 2002; Hardin et al 2003). Estimates in Table 3 are statistically significant at the 10% level and indicate that for each additional percentage point decrease in the proportion of population living in dense areas, obesity is approximately 0.1 to 0.2 percentage points higher. Estimates are larger for higher density thresholds. For example, estimates associated with a density of 12,500 are twice as large as estimates associated with a density of 5,000. Notably, estimates are not sensitive to the inclusion of time-varying MSA controls for median family income, employment rates and education. This provides some evidence supporting the identification assumption underlying the instrumental variables approach, which is that changes in population density caused by the 1947 highway plan are uncorrelated with changes in other attributes of the MSA that are correlated with changes in weight status.

Table 4 presents estimates of the association between population density and BMI. Estimates indicate that a one percentage point decrease in the proportion of population living in dense areas increases BMI by about 0.01 units, but parameter estimates are not statistically significantly at conventional levels. Interestingly, comparing the effects on BMI and obesity suggests that instead of shifting the whole weight distribution to the right, the decline in population density has a larger effect on the upper tail of the weight distribution. This is notable because previous studies that have addressed the selection problem have not used obesity as an outcome and only used BMI. Here we find no effect on BMI, but an association between population density and obesity.

It is worth mentioning that coefficients on individual characteristics generally have the expected signs and are consistent with findings in the previous literature (estimates are not presented). Age has an inverted U-shaped effect on the probability of being obese and BMI. Black and Hispanic persons are more likely to be obese and have higher values of BMI than whites, while persons of other races are less likely to be obese and have lower values of BMI than whites. Men are less likely than women to be obese, but have a higher average BMI than women. Compared to single (never married) individuals, married and divorced individuals are less likely to be obese and widowed persons are more likely to be obese. Years of formal schooling completed has a negative effect on the probability of being obese and BMI.

Sensitivity Analyses

Two-stage Least Squares Estimates

To investigate whether estimates are sensitive to the method of estimation, we obtained standard two-stage least squares estimates. To conduct this analysis, we used data from the NHIS in only the census years of 1980, 1990, and 2000. For these years, census data can be directly matched to the individual-level data from the NHIS. Summary statistics for this "matched" NHIS sample are available from authors' home page. The matched sample's summary statistics are very similar to those of the full sample. We obtain estimates of the association between the proportion of the population living in dense areas and obesity (BMI) using the conventional two-stage least squares approach in which planned highway rays are the excluded instrument.

The two-stage least-squares model is illustrated below:

$$BMI(OBESE)_{ijt} = \alpha_j + \gamma_t + \beta PopDen_{jt} + \phi X_{ijt} + \lambda Z_{jt} + \nu_{ijt} \quad (3)$$

$$PopDen_{jt} = \rho_t + \pi[(YEAR_t - 1947) * HWPLAN_j] + \kappa X_{ijt} + \sigma Z_{jt} + \mu_j + \xi_{jt} \quad (4)$$

$j = 1,...,53$ $(MSAs)$

$t = 1980, 1990, 2000$ $(years)$

Note that individual characteristics are now included in the first-stage regression (equation 4). In addition, we cannot control for changes in MSA definitions in this analysis because unrestricted MSA fixed effects would be perfectly correlated with the planned highway measures, which vary only by MSA and (census) year. Therefore, each MSA is treated as if it was the same in each census year. Finally, the second stage (equation 3) model includes year fixed effects instead of a quadratic time trend.

Estimates obtained by the two-step IV procedure are virtually identical to those obtained using the standard 2SLS procedure and tables of estimation results are available from authors' home page.

Central City versus Suburban Residents

As noted earlier, there are reasons to expect that urban sprawl may have affected the weight of both central city and suburban residents. We investigated this hypothesis by allowing the effect of population density to differ by place of residence within the MSA. The NHIS provides information on whether individuals live in central cities in the years 1976 to 1996. However, the definition of central city also changes with the change of the definition of metropolitan areas. Changes in the definitions of central cities result from: 1) old central cities expanded geographically; 2) the recognition of new central cities when the requirement on population was reached; or 3) dropping of central cities from an area due to changes in population or the economic or social ties to metropolitan areas. Given these measurement problems, we view this analysis as suggestive, particularly because choice of location within the

MSA is likely endogenous.

Estimates of the associations between population density and weight for central city and suburban residents are presented in Table 5. Focusing first on the results for obesity, coefficients on the interaction term between population density and central city residence indicate that city dwellers experience slightly greater changes in their weight status as population density changes, implying approximately 0.01 to 0.02 percentage point difference for each additional percentage point change in the proportion of population living in dense areas. In addition, the central city-suburban differences are smaller both in magnitude and significance for higher density thresholds. The estimate associated with a density of 12,500 are half the size of estimate associated with a density of 5,000 and is not significant at conventional levels, yet the weight status change for people living in central cities is still statistical significant. Considering the fact that the effect of population density on obesity is larger for higher density thresholds, city dwellers are affected more when the extent of the shifts of population to less dense areas are broader. Turning to the results for BMI, estimates in Table 5 show a pattern similar to that of obesity. City dwellers experience slightly greater changes in their BMI than suburbanites, though effects on BMI for both city dwellers and suburbanites are not significantly different from zero.

Conclusion

Previous research has documented a positive association between obesity and urban sprawl (population density). Whether this association represented a causal relationship, however, has not been established. In this paper, we address the causality issue by using a plausibly exogenous source of urban sprawl—the decline in population density caused by the Interstate Highway System plan. Estimates indicate that a one percentage point decrease in the share of

population living in dense areas increased the prevalence of obesity by 0.1 to 0.2 percentage points depending on which threshold of dense area is used. Further analyses that examined how changes in the entire population distribution of the MSA affected obesity show that loss of population in the central city, as well as growth in the suburbs, increases obesity.

To place these results in context, we evaluate the importance of urban sprawl in explaining the rising trend in obesity by examining the counterfactual obesity trend if population density had remained at its 1970 levels. On average, according to the NHIS sample used in this paper, obesity rates increased by 145% from about 8.3 percent to 20.3 percent between 1976 and 2001. We take the predicted values of population density from the first stage to determine the "exogenous urban sprawl" by differencing population density between 2001 and 1976 resulting from 4 rays of planned highways (4 rays are the mean of the planned highway rays). The resulting number is about 10 percentage points. We then multiply this number by 0.0015, the coefficient of the population density with the threshold of 9,000 people per square mile, to derive the average percentage point difference in obesity that can be attributed to the exogenous change in population density. We divide the estimated difference in obesity rates due to exogenous urban sprawl (1.5 percentage points) with the observed difference in obesity rates between 2001 and 1976 (12 percentage points). The estimates thus indicate that about 13% of the increases in the obesity rate can be attributed to urban sprawl.

Overall, the results of this study suggest that urban sprawl did cause an increase in obesity, but its effect was relatively modest. Thus, policy makers may want to look elsewhere for solutions to the obesity problem, particularly if urban and community redesign are costly. While this study has identified a relationship between population density and obesity, it did not identify the underlying mechanisms that link urban sprawl to weight. Future research is warranted to

better understand the mechanisms through which urban sprawl has caused the changes in obesity rates documented in this paper.

References

Baum-Snow, N.: Did highways cause suburbanization? Quarterly Journal of Economics 122;775-805: 2007.

Berrigan D, Troiano R.P.: The association between urban form and physical activity in U.S. adults. American Journal of Preventive Medicine 23;74–79: 2002.

Beydoun, M., Powell, L. and Wang, Y.: The association of fast food, fruit and vegetable prices with dietary intakes among US adults: Is there modification by family income? Social Science & Medicine 66;2218-2229: 2008.

Black, J. L., Macinko, J. Neighborhoods and obesity. Nutrition Reviews, 66**,** 2-20:2008.

Boarnet, M., Greenwald, M.: Land use, urban design, and nonwork travel: Reproducing other urban areas' empirical test results in Portland, Oregon. Transportation Research Record 1722;27-37: 2000.

Centers for Disease Control and Prevention: CDC's role in promoting healthy lifestyles. Statement by Julie L. Gerberding before the Senate Committee on Appropriations, Subcommittee on Labor, HHS, Education and Related Agencies: 2003.

Cervero, R., Roger G.: Commuting in transit versus automobile neighborhoods. Journal of the American Planning Association 61;210-25:1995.

Chung, C., Samuel, L. Myers Jr.: 1999. Do the poor pay more for food? an analysis of grocery store availability and food price disparities. Journal of Consumer Affairs 33;276–296: 1999.

Crane, R., Crepeau, R.: Does neighborhood design influence travel? A behavioral analysis of travel diary and GIS data. Transportation Center Working paper No. 374, University of California, Berkeley:1998.

Cummins, S., Macintyre, S. "Food deserts"—evidence and assumption in health policy making. British Medical Journal, 325**,** 436-8: 2002.

Eid, J., Overman, H.G., Puga, D., Turner, M.A.: Fat city: questioning the relationship between urban sprawl and obesity. Journal of Urban Economics 63;385-404: 2008.

Ewing, R., Pendall, R., Chen, D.: Measuring sprawl and its impact. Smart Growth America, http://www.smartgrowthamerica.org/sprawlindex/MeasuringSprawl.PDF, website last accessed 5/10/10.

Ewing, R., Schmid, T., Killingworth, R., Zlot, A., Raudenbush, S.: Relationship between urban sprawl and physical activity, obesity and morbidity. American Journal of Health Promotion 18;47-57: 2003.

Ewing, R., Browson, R., Berrigan, D.: Relationship between Urban Sprawl and Weight of United States Youth. American Journal of Preventive Medicine. 31(6):464–474: 2006

Ford, P. B., Dzewaltowski, D. A. Disparities in obesity prevalence due to variation in the retail food environment: three testable hypotheses. Nutrition Reviews, 66**,** 216-228: 2008.

Feldstein, L.: Linking land use planning and the food environment: smart growth Online, 2007.

F. H. W. A. U.S. Department of Transportation: Route log and finder list. U.S. Government Printing Office. Washington, DC, 2001.

Finkelstein, E.A., Ruhm, C.J., Kosa, K.M.: Economic causes and consequences of obesity. Annual Review of Public Health 26;239-57: 2005.

Flegal, K.M., Carroll, M.D., Ogden, C.L., Johnson, C.L.: Prevalence and trends in obesity among U.S. adults, 1999–2000. Journal of the American Medical Association 288;1723–1727: 2002.

Flegal, K.M., Graubard, B. I., Williamson, D., Gail, M.: Excess deaths associated with Underweight, Overweight, and Obesity. Journal of the American Medical Association 293;1861–1867: 2005.

Frank, L.: Land use and transportation interaction: Implications on public health and quality of life. Journal of Planning Education and Research 20;6-22: 2000.

Frank, L., Andresen, M., Schmid, T.: Obesity relationships with community design, physical activity, and time spent in cars. American Journal of Preventive Medicine 27;87-96: 2004.

French, S. A., Story, M., Jeffery, R. W., Snyder, P., Eisenberg, M., Sidebottom, A.: Pricing strategy to promote fruit and vegetable purchase in high school cafeterias. Journal of the American Dietetic Association 97;1008-1010: 1997.

French, S. A.: Public health strategies for dietary change: schools and workplaces. Journal of Nutrition 135;910-912: 2005.

Handy, S.: Understanding the link between urban form and nonwork travel behavior. Journal of Planning Education and Research 15;183-98: 1996.

Hardin, J., Carroll, R.: Variance estimation for the instrumental variables approach to measurement error in generalized linear models. The Stata Journal 3(4):342-350: 2003

Hardin, J.: The robust variance estimator for two-stage models. The Stata Journal 2(3): 253-265: 2002.

Hess, P., Moudon, A., Snyder, M., Stanilov, K.: Site design and pedestrian travel. Transportation Research Record 1674;9-19: 1999.

Helmchen, L Marriage Market Incentives to Invest in Health. Ph.D. dissertation. The University of Chicago, Chicago, IL. 2004

Horowitz, C., Colson, K., Hebert, P., Lancaster, K.: Barriers to buying healthy foods for people with diabetes: evidence of environmental disparities. American Journal of Public Health 94;1549 –54: 2004.

Inagami, S., Cohen, D. A., Finch, B. K., Asch, S. M. You are where you shop: grocery store locations, weight, and neighborhoods. American Journal of Preventive Medicine, 31**,** 10-17: 2006.

Jeffery, R.W., French, S.A., Raether, C., Baxter, J. E.: An environmental intervention to increase fruit and salad purchases in a cafeteria. Preventive Medicine 23;788-792: 1994.

Jetter, K., Cassady, D.: The availability and cost of healthier food alternatives. American Journal of Preventive Medicine (30);38–44: 2006.

Kopelman, P.G.: Obesity as a medical problem. Nature 404;635-43: 2000.

Koplan, J.P., Dietz, W.H.: Caloric imbalance and public health policy. Journal of the American Medical Association 282;1579-81: 1999.

Larson, N. I., Story, M. T., Nelson, M. C. (2009) Neighborhood environments: disparities in access to healthy foods in the US. American Journal of Preventive Medicine, 36**,** 74-81:2009.

Lopez, R.: Urban sprawl and risk for being overweight or obese. American Journal of Public Health 94;1574 –1579: 2004.

Lumeng, J. C., Appugliese, D., Cabral, H. J., Bradley, R. H., Zuckerman, B. Neighborhood safety and overweight status in children. Archives of Pediatrics and Adolescent Medicine, 160**,** 25-31: 2006.

Mieszkowski, P., Mills, E.S.: The causes of metropolitan suburbanization. Journal of Economic Perspectives 7;135—147: 1993.

Mokdad, A., Marks, J., Stroup, D., Gerberding, J.: Correction: actual causes of death in the united states, 2000. Journal of the American Medical Association, 293; 293-294: 2005

Maddock, J.: The relationship between obesity and the prevalence of fast food restaurants: state-level analysis. American Journal of Health Promotion, 19**;** 137-143: 2004.

Moon, Graham Residential environments and obesity: estimating causal effects. In, Geographies

of Obesity: Environmental Understandings of the Obesity Epidemic. London, UK, Ashgate. (Geographies of Health Series). 2009. (In Press)

Murphy, K., Topel, R.: Estimation and inference in two-step econometric models. Journal of Business and Economic Statistics 4;370-379: 1985.

Ogden, C.L., Carroll, M.D., Curtin, L.R., McDowell, M.A., Tabak, C.J., Flegal, K.M., 2006. Prevalence of overweight and obesity in the United States, 1999–2004. Journal of the American Medical Association 295;1549–1555: 2006.

Peeters, A., Barendregt, J.J., Willekens, F., Mackenbach, J.P., Al Mamun, A., Bonneux, L.: Obesity in adulthood and its consequences for life expectancy: a life-table analysis. Annals of internal medicine 138;24-32: 2003

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

Plantinga, A.J., Bernell, S. (b): Can urban planning reduce obesity? The role of self-selection in explaining the link between weight and urban sprawl. Review of Agricultural Economics 29;557-563: 2007.

Pisarski, A.: US Roads. In: Millennium Book International Road Federation, 2001.

Powell, L., Auld, C., Chaloupka, F., O'Malley, P., Johnston, L.: Access to fast food and food prices: Relationship with fruit and vegetable consumption and overweight among adolescents. Advances in Health Economics and Health Services Research 17;23-48: 2007

Saelens, B.E., Sallis, J.F., Frank, L.D.: Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. Annals of Behavioral Medicine 25;80–91: 2003.

Sallis, J., Nader, P., Rupp, J., Atkins, C., Wilson, W.: San Diego surveyed for heart-healthy foods and exercise facilities. Public Health Reports 101; 216–9: 1986.

Transportation Research Board: Does the built environment influence physical activity? examining the evidence. Transportation research board special report 282, Institute of Medicine of the National Academies, 2005.

Vandegrift, D., Yoked, T. (2004) Obesity rates, income, and suburban sprawl: an analysis of US states. Health & Place, 10**,** 221-229: 2004.

Wassmer, R., Edwards, D.: Causes of urban sprawl (decentralization) in the United States: natural Evolution, flight from blight, and the fiscalization of land use. Manuscript: 2005.

Wooldridge, J. M.: Econometric analysis of cross section and panel data, MIT Press, 2002.

Wellman, N.S., Friedberg, B.: Causes and consequences of adult obesity: health, social and economic impacts in the United States. Asia Pacific journal of clinical nutrition 11;S705-9: 2002.

World Health Organization: Global Strategy on Diet, Physical Activity and Health. World Health Organization, 2004

Table 1
Change in Population Density between 1970 and 2000 for Large Constant-geography Metropolitan Areas in U.S.

	1970	1980	1990	2000	% Change 1970-2000
Population weighted population density[b]	7,025	5,754	5,640	5,680	-19.2
Proportion of population living in areas with density >3500 people per square mile (%)	56.8	50.3	48.6	47.9	-15.8
Proportion of population living in areas with density >5000 people per square mile (%)	43.3	35.9	34.2	33.8	-19.0
Proportion of population living in areas with density >9000 people per square mile (%)	20.0	14.7	13.6	14.1	-29.9
Proportion of population living in areas with density >12500 people per square mile (%)	12.2	8.5	8.1	8.1	-33.2

[a] The sample includes those 53 largest metropolitan areas identified in the public-use NHIS data based on 1993 metropolitan area definitions.

[b] Population weighted population density is calculated by $\sum_i^I \frac{Pop_i}{Pop_{total}} * \frac{Pop_i}{Land_i}$, where Pop_i refers to the census tract population, Pop_{total} refers to the total metropolitan area population of that census tract, $Land_i$ refers to the square mileage of the census tract.

Table 2
Estimates of the Association between 1947 Highway Plan and Population Density

	Density: (>5000 people per square mile)		Density: (>9000 people per square mile)		Density: (>12500 people per square mile)	
Planned Highway Rays	-2.39***	-2.15***	-2.14***	-1.99***	-1.38***	-1.28***
	(0.37)	(0.43)	(0.39)	(0.44)	(0.31)	(0.33)
MSA Employment Rates		-0.10		0.28		0.40
		(0.34)		(0.37)		(0.27)
MSA Median Family Income		0.20***		0.15		0.13*
		(0.08)		(0.09)		(0.07)
MSA % College or More Education		-0.42		-0.14		0.01
		(0.36)		(0.26)		(0.22)
MSA Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	488	488	488	488	488	488
R square	0.95	0.95	0.95	0.95	0.95	0.95
F-statistics for excluded variable	42.18	25.31	29.80	20.54	20.18	15.00
P-value for excluded variable F-test	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Mean of dependent variable in 1970	47.61	47.61	23.25	23.25	14.20	14.20
(standard deviation)	(17.87)	(17.87)	(17.66)	(17.66)	(15.64)	(15.64)

Notes: "Planned Highway Rays" is calculated by multiplying the number of rays in the 1947 plan with year since 1947 and then dividing the resulting number by 20 ($[(Year_t\text{-}1947)*HWPLAN_j]/20$). Standard errors are in parentheses. Standard errors are clustered by MSA/year. All specifications control for time trend and MSA fixed effects. * indicates significant at 10% level, ** indicates significant at 5% level, *** indicated significant at 1% level.

Table 3
Two-Step Instrumental Variables Estimates of the Association between Population Density and Obesity

	Density (>5000 people per square mile)		Density (>9000 people per square mile)		Density (>12,500 people per square mile)	
Population Density	-0.0011**	-0.0011**	-0.0013**	-0.0014**	-0.0021**	-0.0024**
	(0.0005)	(0.0005)	(0.0006)	(0.0006)	(0.0010)	(0.0010)
MSA Employment rates		0.0002		0.0006*		0.0011**
		(0.0003)		0.0003		(0.0005)
MSA Median Family Income		0.0004		0.0004		0.0005
		(0.0003)		(0.0003)		(0.0004)
MSA % College or More Education		-0.0007**		-0.0004*		-0.0002
		(0.0003)		(0.0002)		0.0002
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
MSA Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	703,544	703,544	703,544	703,544	703,544	703,544
Mean of Obesity	0.1096	0.1096	0.1096	0.1096	0.1096	0.1096
(Standard Deviation)	(0.3124)	(0.3124)	(0.3124)	(0.3124)	(0.3124)	(0.3124)

Notes: In all the specifications we control for individual's demographic information: dummy variables for age, race (non-Hispanic Black, Hispanic and others, non-Hispanic White is the reference group), sex, income, education (some high school, high school graduate, some college and college graduate, elementary school is the reference group) and marital status (married, separated/divorced, or widowed, single is the reference group). Table 2 shows the model used to create the predicted population density. Standard errors are corrected using Murphy-Topel estimate of variance-covariance. Standard errors are clustered at MSA/year. The regressions are weighted with NHIS sampling weights. * indicates significant at 10% level, ** indicates significant at 5% level, *** indicated significant at 1% level.

Table 4
Two-Step Instrumental Variables Estimates of the Association between Population Density and BMI

	Density (>5000 people per square mile)		Density (>9000 people per square mile)		Density (>12,500 people per square mile)	
Population Density	-0.0095	-0.0113	-0.0103	-0.0129	-0.0159	-0.0202
	(0.0076)	(0.0081)	(0.0090)	(0.0096)	(0.0141)	(0.0154)
MSA Employment rates		0.0028		0.0054		0.0099
		(0.0041)		(0.0050)		(0.0077)
MSA Median Family Income		0.0114**		0.0113**		0.0120**
		(0.0048)		(0.0048)		(0.0050)
MSA % College or More Education		-0.0145***		-0.0116***		-0.0095***
		(0.0047)		(0.0036)		(0.0033)
Individual Controls	Yes	Yes	Yes	Yes	Yes	Yes
MSA Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Trend	Yes	Yes	Yes	Yes	Yes	Yes
Sample Size	703,544	703,544	703,544	703,544	703,544	703,544
Mean of BMI	24.6591	24.6591	24.6591	24.6591	24.6591	24.6591
(Standard Deviation)	(4.5574)	(4.5574)	(4.5574)	(4.5574)	(4.5574)	(4.5574)

Notes: In all the specifications we control for individual's demographic information: dummy variables for age, race (non-Hispanic Black, Hispanic and others, non-Hispanic White is the reference group), sex, income, education (some high school, high school graduate, some college and college graduate, elementary school is the reference group) and marital status (married, separated/divorced, or widowed, single is the reference group). Table 2 shows the model used to create the predicted population density. Standard errors are corrected using Murphy-Topel estimate of variance-covariance. Standard errors are clustered at MSA/year. The regressions are weighted with NHIS sampling weights. * indicates significant at 10% level, ** indicates significant at 5% level, *** indicated significant at 1% level.

Table 5
Estimates of the Association between Population Density and Weight
Central City and Suburban Residents

	Density (>5,000 people per square mile)		Density (>9,000 people per square mile)		Density (>12,500 people per square mile)	
	Obesity	BMI	Obesity	BMI	Obesity	BMI
Predicted Population	-0.0012**	-0.0072	-0.0015**	-0.0074	-0.0023**	-0.0116
Density	(0.0006)	(0.0087)	(0.0007)	(0.0102)	(0.0010)	(0.0160)
Central City	0.0126***	0.1402***	0.0044**	0.0297	0.0023	0.0032
	(0.0030)	(0.0474)	(0.0020)	(0.0312)	(0.0017)	(0.0268)
Predicted Population	-0.0002***	-0.0034***	-0.0001**	-0.0024***	-0.0001	-0.0021**
Density * Central City	(0.0001)	(0.0009)	(0.0001)	(0.0009)	(0.0001)	(0.0010)
Observations	622,644	622,644	622,644	622,644	622,644	622,644

Notes: In all the specifications we control for individual's demographic information. Standard errors are corrected using Murphy-Topel estimate of variance-covariance. Standard errors are clustered at MSA/year. * indicates significant at 10% level, ** indicates significant at 5% level, *** indicated significant at 1% level.