Volume Title: Economic Aspects of Obesity

Volume Author/Editor: Michael Grossman and Naci H. Mocan, editors

Volume Publisher: University of Chicago Press

Volume ISBN: 0-226-31009-4
ISBN13: 978-0-226-31009-1

Volume URL: http://www.nber.org/books/gros09-1

Conference Date: November 10-11, 2008

Publication Date: April 2011

Chapter Title: Studying the Child Obesity Epidemic with Natural Experiments

Chapter Authors: Robert Sandy, Gilbert Liu, John Ottensmann, Rusty Tchernis, Jeff Wilson, O. T. Ford

Chapter URL: http://www.nber.org/chapters/c11820

Chapter pages in book: (181 - 221)

7

Studying the Child Obesity Epidemic with Natural Experiments

Robert Sandy, Gilbert Liu, John Ottensmann, Rusty Tchernis, Jeff Wilson, and O. T. Ford

7.1 Introduction

One of the broad strategies for reducing child obesity is to alter the built environment near children's homes to either promote calorie expenditures or reduce calorie consumption. Some of the proposals in this built environment strategy are to: (a) build sidewalks, (b) construct recreational amenities such as pools, soccer fields, basketball courts, and trails (c) require mixes of residences and retail outlets through zoning laws, (d) locate schools within walking distance of homes, and (e) limit fast-food restaurants (King et al. 1995; Sallis, Bauman, and Pratt 1998; Margetts 2004; Committee on Environmental Health 2009). The two crucial questions in implementing this built-environment strategy are which proposals will reduce child obesity, and the degree to which crime counteracts any benefits of changing the built environment.

This chapter uses clinical data from 1996 to 2006 to estimate the effect of

Robert Sandy is assistant executive vice president of Indiana University and professor of economics at Indiana University–Purdue University Indianapolis (IUPUI). Gilbert Liu is assistant professor of pediatrics at the Indiana University School of Medicine. John Ottensmann is professor of public and environmental affairs and director of Urban Policy Research at Indiana University-Purdue University Indianapolis (IUPUI). Rusty Tchernis is associate professor of economics at the Andrew Young School of Policy Studies, Georgia State University, and a faculty research fellow of the National Bureau of Economic Research. Jeff Wilson is associate professor and chair of the department of geography at Indiana University-Purdue University Indianapolis (IUPUI). O. T. Ford is a graduate student in the department of geography at the University of California, Los Angeles.

We thank Shawn Hoch, Zhang Ya, Megan McDermott, Bikul Tulachan, and Jonathan Raymont for research assistance. This study was funded under NIH NIDDK grant R21 DK075577-01. We thank the participants at the Purdue University Department of Agricultural Economics seminar, and at the NBER preconference and conference on Economic Aspects of Obesity for many helpful comments, and Kristen Butcher for a careful and insightful review.

changes in nearby physical environment and crime levels on child weight. Electronic medical records for patients who received care at a large academic health care system in Indianapolis between 1996 through 2006 provide anthropometric, demographic, and geographic data for over 60,000 children between the ages of three and sixteen years of age. The identifying assumption in our research design is that changes in the physical environment are exogenous to children who were at the same address before and after the change. This natural experiment approach addresses a major limitation of cross-sectional studies of the associations between environment and child obesity. Cross-sectional associations are confounded when families select their locations. For example, families that highly value exercise are more likely to live near a recreation trail. In cross-sectional regressions, a variable for proximity to a recreation trail could pick up the unobserved effects of the selection.

Controlling for selection is important in identifying the determinants of body weight. For example, in the adult body weight context there are many cross-sectional studies that find that urban sprawl is associated with obesity. However, these results were not supported by studies of migrants between high- and low-sprawl cities (Plantinga and Bernell 2007) and for cities that changed their sprawl levels over time (Ewing, Brownson, and Berrigan 2006). In the child obesity context sprawl was associated with a higher probability of overweight in cross-sectional data, but the initial level of sprawl did not have a significant effect on weight gains over three years (Ewing, Brownson, and Berrigan 2006).

The built environment changes we study are fast-food restaurants, convenience stores, supermarkets, recreation trails, and thirteen publicly accessible recreational amenities, such as public outdoor basketball courts and outdoor pools. We also studied the direct effect of crime levels on child obesity and interactions between changes in crime levels and the built environment on obesity. We test our exogeneity assumption by comparing the trend in body mass index (BMI) for children who will gain an amenity in the future to the trend for children who never gain an amenity. We use the term "amenities" to mean any built environment factor, desirable or undesirable. We found that, except for supermarkets, all of the amenities had largely the same trends for children who would in the future gain an amenity and those who would not.

The remaining sections of the chapter are a literature review, a description of the data, the estimation strategy, results, and conclusions.

7.2 Literature Review

Although the entire literature on the effect of the built environment on child obesity is quite recent, it is growing rapidly. A comparison of three literature review articles published in 2005, 2007, and 2010 illustrates this rapid growth. The 2005 review article by Booth, Pinkston, and Carlos Poston covers the years 1998 and 2004, and includes two studies on child obesity/overweight and the built environment. The 2007 review article by Papas and colleagues discusses six articles published over a four-year period, beginning in 2002. A 2010 review article by Galvez, Pearl, and Yen describes fifteen studies published over a twenty-month period beginning in January of 2008. All of the child obesity and built environment studies in these review articles use cross-sectional data. Based on these three review articles, the rate of production of articles increased from 0.33 per year in 1998 to 2004, to 1.5 per year in 2002 to 2006, and to nine per year in 2008 through the first eight months of 2009.

In light of the aforementioned three review articles, this section will only highlight a few important studies and make three observations about the literature that connect to this study. The first observation is a need for weight or weight status as an outcome. Second, there is a lack of understanding as to what are the distances between homes and amenities that should be considered in defining the relevant built environment surrounding a child's residence. Third, analyses should be of specific amenities rather than aggregations of amenities. The review article by Galvez, Pearl, and Yen describes thirty-three additional articles that estimate associations between the built environment and behaviors thought to affect weight, for example, walking to school. A behavioral response to a change in the built environment is a necessary, but not a sufficient, condition for changes in body weight. For example, a potential weight loss due to behavior observed at a recreation amenity can be offset if the child either reduces other exercise or eats more. Measures of body weight are required to confirm that observed behaviors do reduce body weight. A related problem with many studies of behaviors tied to specific environmental amenities is that, because they are expensive, they are usually done at a small scale. An example of small-scale observations of behavioral responses is using trained observers to record activity levels in three city parks and categorize whether the children's play is sedentary, moderate, or vigorous (Tester and Baker 2009). Even if body weight measures were available in such small-scale studies it would be difficult to detect a statistically significant effect.

While some nationally representative data sets only include self-reported or parent-reported height and weight data (e.g., the Youth Behavioral Risk Factor Surveillance System), there are a few that include directly measured height and weight (e.g., National Health and Nutrition Examination Survey). However, these national data sets have severe limitations for studies examining associations between weight status and environment. One limitation, due to restrictions on providing home address information in these data sets, is the requirement to summarize amenities at relatively large geographies around the center of an area known to include the child's home. Many studies have found that children are unlikely to walk or bike to school if the distance is beyond a mile (McDonald 2007; Timperio et al. 2004; Schlossberg et al. 2006; Merom et al. 2006). Presumably, recreational amenities need to be within a mile of the child's home to generate frequent usage by children who walk or ride bicycles. If being driven or taking public transportation is the primary mode for children to access an amenity, then the transportation time is much more important than the Euclidean distance. A study by Gordon-Larsen et al. (2006) illustrates the coarse spatial resolutions that arise from limitations in national probability samples for health surveillance. They were restricted to locating study subjects based on Census block groups, and then defined amenities as relevant if they were within a five-mile radius of the Census block group. Census block groups vary widely in size based on density of settlement, but optimally contain approximately 1,500 persons.

Another problem in studies that use national data sets is the necessity of combining different types of recreational amenities when developing variables to the point where there is little chance of being able to discern which amenity types are beneficial. Singh, N. Siahpush, and M. D. Kogan (2010) combine parent/guardian-reported data on availability of (a) sidewalks and walking paths, (b) parks and playgrounds, (c) recreation centers, community centers, and boys' and girls' clubs, and (d) libraries and bookmobiles into an index of built environment resources. They found that higher values of this index were associated with a lower probability of overweight. The inclusion of libraries and bookmobiles is puzzling in that they seem to be resources that encourage reading, a sedentary activity. Even if libraries and bookmobiles somehow reduced body weight, the index approach blurs that effect with those of the other amenities.

The effect of fast-food restaurants on obesity has probably received more public attention than any other food or recreation amenity. The presumed link between fast-food and child obesity has prompted a ban on new restaurants in parts of Los Angeles (Los Angeles Times 2008). This link, however, may reflect in part decisions to locate fast-food restaurants in areas in which residents have a higher than average taste for high-caloric foods, and hence a larger than average percentage of the population is overweight. Anderson and Matsa (2007) have one of the few obesity papers that address the endogeneity of location issue for any type of amenity. However, it is for adults living in rural areas. They used location near an interstate highway exit as an instrument for fast-food location. Their conclusion is that fast-food has no causal effect on adult BMI.

A second paper on fast-food restaurants and obesity that addresses endogeneity is Currie et al. (2009). They have 3.06 million student-year observations for ninth graders in California over the years 1999 and 2001 through 2007, with precise locations of their schools and the fast-food restaurants. They do not have data on individual children. Obesity rates are reported for all ninth graders in a school. The measurements on the children are taken during the spring semester and represent approximately thirty weeks of exposure to any near-to-the-school fast-food restaurant for a child entering high school. They find a 5.2 percent increase in the incidence of obesity, relative to the mean of 32.9 percent, for schools that have a fast-food restaurant within 0.1 miles. They found no effect for fast-food within analytic buffers of 0.25, 0.5, and 1.0 miles from the school. They attribute the statistically significant and economically important effect within 0.1 miles as being due to the ninth graders having to be able to quickly walk to the fast-food restaurants for it to influence their weight status.

Since Currie and colleagues have no data on individuals, they cannot know if the children who gained a nearby fast-food restaurant by enrolling in a particular high school had also recently gained a fast-food restaurant near their homes. They also could not exclude the possibility that a higher proportion of the children whose schools have fast-food restaurants within 0.1 miles entered the ninth grade being already obese. Going across years, they find no trend in obesity rates at schools that will gain a fast-food restaurant in the future. However, they have very little temporal variation at the level of the school in the number of fast-food restaurants within 0.1 miles.

The Currie and colleagues paper shows the importance of being able to define highly precise specifications of the environment around a child's home when studying how the built environment is associated with obesity. The first law of geography according to Waldo Tobler is, "Everything is related to everything else, but near things are more related than distant things" (Tobler 1970). The likelihood of children interacting with amenities is probably inversely proportional to the cost of travel to them. Or stated slightly differently, physical activity or eating opportunities that are very nearby likely have stronger influence than those opportunities that are as far as five Euclidean miles distant.

Summing up this literature review section, several of the gaps in the existing literature are addressed in our study. We address the endogeneity of location issue. We draw tight circles of 0.1, 0.25, 0.5, and 1.0 miles around the child's home. We estimate the individual effects of different types of recreational amenities and food vendors. We utilize directly measured heights and weights, which are much more reliable than parent/guardian-reported heights and weights. We do not, however, have observations on child behavior. To our knowledge, no study to date has a comparably large sample, direct observations of behavior at recreation and food amenities, and the geographic resolution needed to analyze the effect of amenities near children's residences. Privacy concerns in national data sets plus the expense of either directly observing child behavior or administering a survey about exactly what amenities children utilize would make doing such a study difficult.

7.3 Data

The main sources of our data are: (a) clinical records from pediatric ambulatory visits to the Indiana University Medical Group between 1996 and 2006; (b) annual inspections by the Marion County Health Department of all food establishments; (c) aerial photographs, used to identify and verify recreational amenities; (d) reports of violent crimes from the Indianapolis Police Department and the Marion County Sheriff's Department; (e) birth certificates; and (f) the U.S. Census. These six data sources are described in more detail below.

7.3.1 Clinical Records From Well-Child Visits

The Indiana University Regenstrief Medical Records System (RMRS), in existence since 1974, is an electronic version of the paper medical chart. It has now captured and stored 200 million temporal observations for over 1.5 million patients. Because RMRS data are both archived and retrievable, investigators may use these data to perform retrospective and prospective research. The RMRS is distributed across three medical centers, thirty ambulatory clinics, and all of the emergency departments throughout the greater Indianapolis region. The RMRS supports physician order entry, decision support, and clinical noting, and is one of the most sophisticated and most evaluated electronic medical record systems in the world.

Using the RMRS, we identified medical records in which there are simultaneous assessments of height and weight in outpatient clinics for children aged three to eighteen years inclusive. We then decided to exclude children above the age of sixteen for reasons described in the section on descriptive statistics. For these clinic visits, we extracted the visit date, date of birth, gender, race, insurance status, and visit type (e.g., periodic health maintenance versus acute care). We found that too few patients had private insurance for this variable to have any predictive power. Because height and weight measurements are routinely performed as part of pediatric health maintenance, these measures should be present for virtually all children receiving preventive care at each of the study sites. The data generated by pediatric visits in the RMRS include higher representation of low-income and minority households compared to the demographics of the study area because the associated clinics serve a population that is mostly publicly insured or has no insurance. The overrepresentation of minorities and low-income households in the RMRS, we contend, is a decided advantage. Poorer households are more sensitive to their immediate neighborhood because they face financial constraints against motorized transit (e.g., reduced car ownership, less money for gasoline or bus fares). Indianapolis has a vestigial public transportation system. It has been described as the worst city system in the Midwest (Quigley 2003). If the built environment has any effect on child weights it should be most readily observed in poorer households in a city with minimal public transportation. Moreover, obesity is more prevalent in poorer households and among poorer children. Knowing what interventions reduce and exacerbate child overweight in this population would be valuable.

National guidelines for well-child visits advocate annual visits for ages three to six years and at least biannual visits thereafter. We extracted ICD-9 codes or other diagnoses list data for identifying children who may have systematic bias in growth or weight status (i.e., pregnancy, endocrine disorders, cancer, congenital heart disease, chromosomal disorders, and metabolic disorders), and excluded observations for such children. We also excluded patient encounters prior to 1996 because the RMRS did not archive address data before this date.

7.3.2 Food Establishment Data

We received annual inspection data on 8,641 food service establishments in Indianapolis that received permits from the Marion County Board of Health between 1993 and 2007, inclusive. Of these, 5,550 are restaurants and 1,507 are in the grocery category. Fast-food establishments have been a particular focus of research on adult obesity and child overweight. Defining and identifying fast-food restaurants is problematic. Fast-food establishments in our study have been defined in two ways. Chou, Rashad, and Grossman (2008) identified a set of forty-one national fast-food chains when they studied the effect of local advertising on child overweight. We will refer to that as the "national chains" list. The national chains are of special interest because they advertise more than local restaurants and local chains, and their restaurants are generally larger, in higher-traffic locations, and more likely to have a drive-up window. The second method of identifying fast-food relied on the Census Bureau's counts of restaurants by Standard Industrial Classification (SIC) codes. Chou, Grossman, and Saffer (2004) used the Census Bureau data for state-level counts of establishments in the SIC 5812/40. These are establishments with a limited menu of items such as pizza, barbecue, hotdogs, and hamburgers. We refer to these restaurants as limited-service restaurants. Full-service restaurants (SIC 5812/10), in contrast, have at least fifteen seats, table service, and serve prepared food from a full menu.

We have 735 establishments in Indianapolis on the national chains list and 1,138 establishments on the broader limited-service list. Data-cleaning challenges included repeated counting due to slight changes in names of restaurants at the same address. Of the 735 fast-food restaurants on the national chains list, 393 were opened between 1994 and 2004, which allows a natural experiment investigating change in food environment as a possible cause of change in child body mass index.

There were 1,507 retail food establishments in the data. Again, we had to do some data cleaning. From the perspective of a Marion County food inspector, a sushi retailer that rents space in a supermarket is a separate inspection entity, but from the consumer's perspective it is part of the supermarket. After a first cut at data cleaning, there were 114 supermarkets. The Indianapolis market, not atypically, has been roiled by the entry of supermarket chains, as well as discount stores with embedded supermarkets such as Meijer, Walmart, and Target. The city's largest chain, Kroger, has had a substantial expansion. Some of the entrants failed, such as Cub Foods, and have left behind stores that are still empty. Among supermarkets there is even more variation, proportionally, than among the national chains' fast-food establishments. Fifty of the 114 supermarkets would satisfy the temporal requirement for a natural experiment because they opened after the first year, closed before 2004, or both.

7.3.3 Recreational Amenities

The study began with a geographic database of recreational amenities and associated features (such as parking lots), in vector form, developed from 2001 data provided by the Indianapolis Parks and Recreation Department. Each individual amenity, such as a basketball court or soccer field, was included as a feature in the database. We incorporated three other similar databases for later periods, also provided by the Indianapolis Parks and Recreation Department.

Additional recreational amenities were identified for the years 1995 through 2006 through the interpretation of aerial photographs. We chose thirteen specific recreational amenities for identification. These were thought to be the most likely to be used by children in the study population and to be amenable to identification from aerial photographs, as well as sufficiently numerous to measure an effect. The chosen categories and their quantifications within 0.1, 0.25, 0.5, and 1 mile buffers centered on the child's home are:

1. Baseball and softball fields, count of fields in buffer
2. Outdoor basketball courts, count of hoops in buffer
3. Family centers (indoor recreation center), area of facilities in buffer
4. Fitness courses, area in buffer
5. Football fields, count of fields in buffer
6. Kickball fields, count of fields in buffer
7. Playgrounds without permanent equipment, area in buffer
8. Playgrounds with permanent equipment, area in buffer
9. Swimming pools, area of water in buffer
10. Soccer fields, total area available for playing in buffer
11. Tennis courts, count of courts in buffer
12. Track and field facilities, area of facilities in buffer
13. Volleyball courts, count of courts in buffer

Nine photo interpreters participated in the process; they were assigned specific areas of the county, generally strips half a mile wide running north-to-south. To control for quality of interpretation, amenities lying on the borders of assigned interpretation areas were to be analyzed by both relevant interpreters. The resulting border features were then compared to each other and to the photographs. Where the features differed, the more accurate interpretation was selected for the final data set, and corrected if necessary. Additionally, errors that appeared in this process were treated as potential systematic errors; the other features interpreted by the responsible interpreter were examined for evidence of the same error repeated. If present, such errors were corrected, and if errors were found while the process was ongoing, the interpreter was retrained to avoid the error. An appendix contains the full details of the photo interpretation process.

7.3.4 Crime Data

When the city limits of Indianapolis were expanded to the border of Marion County in 1970 (while excluding certain small municipalities), the original police jurisdictions were not affected. Therefore, during the study period, the primary law enforcement responsibility for Marion County was divided between the Indianapolis Police Department (IPD), which had responsibility for the area within the original Indianapolis boundary, the Marion County Sheriff's Department (MCSD), which had responsibility for most of the outlying areas of the county, and the police departments of the four small excluded municipalities of Speedway, Lawrence, Southport, and Beech Grove. In 2007, the Indianapolis Police Department and the Marion County Sheriff's Department were merged into the Indianapolis Metropolitan Police Department.

From the Indianapolis Police Department, for the IPD service area in which they had primary responsibility, we have a data set of the geocoded locations of all crimes reported for the Federal Bureau of Investigation's Uniform Crime Reports (UCR), from 1992 through 2005. From the Marion County Sheriff's Department, for the area in which they had primary responsibility, we have a data set on the point locations of a wide range of crimes and other incidents, including the UCR crimes, from 2000 through 2005. We are using information on the crimes from both data sets that are included in the UCR violent crime categories: criminal homicides, rapes, robberies, and aggravated assaults. The data set includes the date and time of the crime, and more detailed information on the specific type of crime within each of those four categories. Because of the manner in which these data have been assembled, we have reason to believe that these are accurate locations and that the classification of the type of crime is accurate.

To summarize, we have the following coverage for violent crimes:

1. Through 1999, for the IPD service area only.
2. From 2000 through 2005, for both the IPD service area and the MCSD jurisdiction.

No crime data are available for any time period for the jurisdictions of the four small excluded municipalities that are within Marion County.

7.3.5 Birth Certificate Data From the Marion County Health and Hospital Corporation

We matched children's clinical data with Marion County Health and Hospital Corporation data on birth certificates by date of birth, gender, mother's surname, and child's given name. We were able to match 34.3 percent of the children in the clinical data. For a match to be possible, the child must have been born in Marion County. The birth certificate data include birth weight, sex, race, mother's age and intention to breastfeed, parents' marital status, and one or both parents' education, race, and eligibility for Women, Infants, and Children (WIC) aid (all, of course, at time of birth). In the few cases where reported race changed between the birth certificate and the clinical record, we used the race identified in the clinical record.

7.3.6 Neighborhood Characteristics

Neighborhood characteristics were estimated for 0.5 mile and 1.0 mile buffers surrounding each residence. These include five variables derived from Census 2000 data: population density, proportion African American, proportions graduated from high school and from college, and median family income. The first two are estimated from the block data, the remainder from the block group data. Two additional neighborhood characteristics are measures of the density and of the interconnectedness of the road network. Land use diversity in the area is represented by the proportion of land in commercial and residential use. Detailed information on the data sources and procedures used to create these variables are provided in an appendix.

Data Cleaning

In examining the height and weight data from the clinical records we found highly improbable patterns, such as a child shrinking five inches in height from one well-child visit to the next. We calculated z-scores for height and weight measures based on year 2000 U.S. Centers for Disease Control and Prevention (CDC) growth charts. We used CDC statistical programs to identify biologically implausible values for heights and weights (Centers for Disease Control and Prevention 2000). Figure 7.1 shows the histograms of heights and weights, excluding biologically implausible values with z-scores greater than $+ 3.0$.

Visually, there is a small amount of truncation for the heights in the right tail of its distribution. As can be seen in the second graph, the truncation in the right tail of the body weight distribution is substantial. The CDC Growth Chart reference population spans the period 1963 to 1994, and thus does not fully cover the epidemic in child obesity of the past two decades. Another visual indicator of the extent of the epidemic is how much the distribution has shifted to the right relative to the mean of the reference population. We

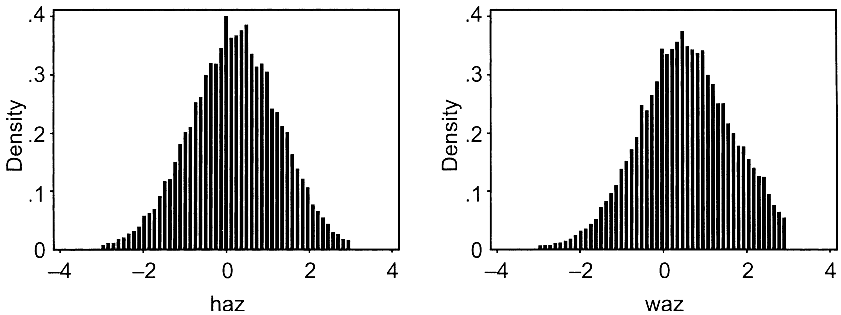

Fig. 7.1 Histograms of standardized height (haz) and weight (waz) scores after dropping observations with *z*-scores at or above 3

dropped observations with weight-for-height, weight-for-age *z*-scores, and BMI *z*-scores either below –3 or above + 5.0 as under the assumption that they were likely due to data entry error or measurement errors.

7.4 Descriptive Statistics

The histogram of the ages of all children at the time of their well-child visits is shown in figure 7.2. Since well-child visits for preschool-age children are more frequent, the higher bars for ages three, four, and five were expected, as well as the steady decline thereafter. Prior to age sixteen the genders at the well-child visits are split nearly fifty-fifty. For age seventeen, the ratio is more than eighty-twenty females-to-males. One of the authors of this chapter works as a pediatrician in these clinics. His explanation for the high ratio of females-to-males among the sixteen- and seventeen-year-olds is the requirement that the girls must have physicals before they can obtain birth control pills.

The sixteen- and seventeen-year-olds are clearly a different population. To simplify our analysis, we restricted our sample to children under the age of sixteen. The literature on child obesity and the built environment has papers on a variety of age ranges. There is no established basis for splitting up our broad range of ages into subsamples. There is not an obvious age at which children in the range of three through fifteen gain significantly more control over their food and exercise choices. Although obtaining a driver's license does give a child much more independence, our exclusion of children aged sixteen or greater eliminates this factor. Lacking any a priori or literature comparison basis for splitting the sample by age, our split was dictated by the data. Almost exactly 50 percent of our observations are below age eight. For each amenity, we tested whether the coefficients for children younger than eight were the same as for children eight or older. Some of our amenities are clearly suited for younger children, such as playground equipment,

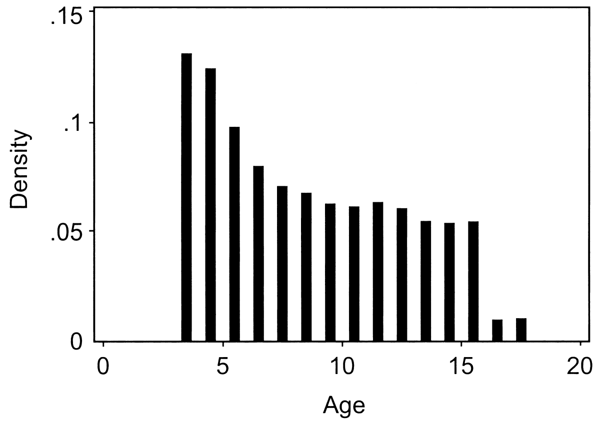

Fig. 7.2 Histogram of age at time of clinic visit

and others for older children, such as volleyball courts, tennis courts, and football fields.

Most studies of child weight determinants use the body mass index percentile compared to the pre-epidemic population sampled in the National Health and Nutrition Examination Survey (NHANES) I from 1971 to 1974. Absolute thresholds for overweight or obesity are not applied to children because the amount of body fat changes with age and differs between girls and boys. The mean body mass index percentile in our data is 65.5. A disadvantage to using the percentiles of the reference population is truncation. For the 205 well-child visits in our sample that were at the one-hundredth percentile, we would not be able to observe any responses to amenities that increase their weight. For the 5,049 well-child visits at or above the ninety-ninth percentile, there would be limited ability to observe weight increase responses. To avoid the truncation problem inherent in the percentiles, we used the BMI *z*-score as the dependent variable.

Table 7.1 has the descriptive statistics for the data used in the cross-sectional analysis. The definitions and procedures used to create the neighborhood characteristics variables are in an appendix. Again, all of the data were restricted to children whose age was under sixteen, and who had an absolute value of the *z*-score of height relative to the reference population less than three, a *z*-score of weight relative to the reference population between –3 and 5, and a *z*-score of BMI relative to the reference population also between –3 and 5.

Table 7.2 has the environmental variables that are based on the annual Marion County food establishment inspections, Indianapolis Department of Parks and Recreation records, the Indianapolis and Marion County crime reports, and our photo interpretation of recreational amenities. These are reported within buffers of 0.1 mile, 0.1 to 0.25 miles, 0.25 to

Table 7.1 Descriptive statistics for the cross-sectional analysis

Clinical observations by year	
Year	*N*
1996	1,811
1997	10,744
1998	13,437
1999	13,289
2000	12,242
2001	12,034
2002	11,945
2003	11,502
2004	8,135
2005	7,795

	N	Mean	St. d.	min.	max.
		Clinical data			
BMI *z*	102,955	0.68	1.17	–2.99	4.99
Well-child visit	102,955	0.82	0.38	0.00	1.00
Female	100,937	0.48	0.50	0.00	1.00
White	102,955	0.29	0.45	0.00	1.00
Black	102,955	0.53	0.50	0.00	1.00
Hispanic	102,955	0.13	0.33	0.00	1.00
		Neighborhood characteristics			
Population density	102,955	9.32	4.23	0.00	27.35
Percent black	102,954	0.43	0.35	0.00	0.98
Percent high school	102,954	0.70	0.13	0.38	1.00
Percent college	102,954	0.12	0.11	0.00	0.85
Median family income	102,954	37,540	12,302	11,202	157,951
Road network density	102,955	3.60	1.24	0.00	9.27
Number of road intersections	102,955	25.10	15.78	0.00	109.00
Commercial land use	102,955	0.07	0.09	0.00	0.89
Residential land use	102,955	0.64	0.21	0.00	1.00
		Marion County birth certificate data:			
Child's birth weight (g)	54,066	3141.55	630.21	170.00	5443.00
Father's age	26,827	27.23	7.37	14.00	91.00
Father's years of education	25,562	11.66	2.10	1.00	26.00
Mother's age	54,171	23.10	5.64	11.00	50.00
Mother's education	53,095	11.21	1.90	1.00	24.00
Intention to breastfeed	50,487	0.22	0.41	0.00	1.00
Marital status (1 = married)	54,986	0.28	0.45	0.00	1.00
WIC Eligibility (1 = yes)	50,890	0.74	0.44	0.00	1.00

0.5 miles, and 0.5 to 1 mile. The table reports the average values by buffer over the study period. The definitions of the recreational amenities are in an appendix.

One striking number from the descriptive statistics is the amount of crime. The violent crimes included are criminal homicides, rapes, robberies,

Table 7.2 **Amenity variables**

	Mile radius	Mean	St. d.	min.	max.	*N*
Fast-food restaurants	.1	0.03	0.20	0.00	4.00	98,541
	.1–.25	0.22	0.66	0.00	7.00	98,541
	.25–.5	0.87	1.40	0.00	11.00	98,541
	.5–1	3.44	2.99	0.00	20.00	98,541
Supermarkets	.1	0.01	0.08	0.00	1.00	98,541
	.1–.25	0.06	0.25	0.00	3.00	98,541
	.25–.5	0.22	0.48	0.00	4.00	98,541
	.5–1	0.78	0.82	0.00	4.00	98,541
Convenience stores	.1	0.03	0.18	0.00	2.00	98,541
	.1–.25	0.17	0.43	0.00	3.00	98,541
	.25–.5	0.54	0.83	0.00	7.00	98,541
	.5–1	1.98	1.85	0.00	11.00	98,541
Trails (m)	.1	12.72	106.14	0.00	2100.67	102,955
	.1–.25	108.83	478.02	0.00	7110.42	102,955
	.25–.5	503.92	1323.73	0.00	12646.57	102,955
	.5–1	2419.53	4091.35	0.00	40526.93	102,955
Violent crimes (annual)	.1	4.18	5.37	0.00	49.00	98,541
	.1–.25	16.09	17.56	0.00	135.00	98,541
	.25–.5	47.16	49.13	0.00	354.00	98,541
	.5–1	155.80	144.66	0.00	739.00	98,541
Baseball diamonds	.1	0.07	0.35	0.00	7.00	102,955
	.1–.25	0.47	1.05	0.00	12.00	102,955
	.25–.5	1.67	2.09	0.00	16.00	102,955
	.5–1	6.31	4.09	0.00	27.00	102,955
Basketball hoops	.1	0.25	0.78	0.00	9.00	102,955
	.1–.25	0.98	1.63	0.00	18.00	102,955
	.25–.5	3.05	2.97	0.00	21.00	102,955
	.5–1	10.40	5.94	0.00	56.00	102,955
Family centers (m^2)	.1	1.43	34.52	0.00	1430.10	102,955
	.1–.25	13.72	130.18	0.00	3483.40	102,955
	.25–.5	88.92	389.36	0.00	3483.40	102,955
	.5–1	357.66	755.85	0.00	5517.90	102,955
Fitness areas (m^2)	.1	2.37	63.89	0.00	4099.10	102,955
	.1–.25	23.09	266.94	0.00	5423.90	102,955
	.25–.5	69.10	462.40	0.00	11786.00	102,955
	.5–1	390.63	1343.81	0.00	11786.00	102,955
Football fields	.1	0.02	0.16	0.00	2.00	102,955
	.1–.25	0.10	0.34	0.00	5.00	102,955
	.25–.5	0.27	0.61	0.00	8.00	102,955
	.5–1	1.03	1.19	0.00	11.00	102,955
Kickball diamonds	.1	0.01	0.11	0.00	3.00	102,955
	.1–.25	0.06	0.26	0.00	4.00	102,955
	.25–.5	0.21	0.50	0.00	5.00	102,955
	.5–1	0.65	1.01	0.00	7.00	102,955
Playgrounds, no equipment (m^2)	.1	72.99	316.35	0.00	4559.50	102,955
	.1–.25	309.05	684.31	0.00	10268.30	102,955
	.25–.5	982.75	1300.46	–0.10	12972.40	102,955
	.5–1	2981.93	2373.51	0.00	14253.70	102,955

Table 7.2 (continued)

	Mile radius	Mean	St. d.	min.	max.	*N*
Playgrounds with equipment (m^2)	.1	133.70	372.93	0.00	6818.80	102,955
	.1–.25	486.26	797.68	0.00	10165.20	102,955
	.25–.5	1408.51	1617.50	–0.10	15141.80	102,955
	.5–1	5170.40	3557.66	0.00	21986.30	102,955
	.1	22.24	85.19	0.00	1717.80	102,955
Pools (m^2)	.1–.25	69.27	193.09	0.00	2825.10	102,955
	.25–.5	201.78	344.25	0.00	5526.10	102,955
	.5–1	710.51	811.61	0.00	7247.30	102,955
	.1	39.42	464.84	0.00	23207.30	102,955
Soccer (m^2)	.1–.25	481.89	2275.35	0.00	77155.50	102,955
	.25–.5	1937.60	6346.88	0.00	137783.10	102,955
	.5–1	8364.20	15137.73	0.00	193082.40	102,955
	.1	0.10	0.51	0.00	12.00	102,955
Tennis	.1–.25	0.45	1.18	0.00	32.00	102,955
	.25–.5	1.48	2.67	0.00	35.00	102,955
	.5–1	5.41	5.63	0.00	47.00	102,955
	.1	47.92	467.45	0.00	15158.00	102,955
Track and field (m^2)	.1–.25	394.62	1725.75	0.00	19316.10	102,955
	.25–.5	1052.36	2940.19	0.00	25371.70	102,955
	.5–1	4024.17	6037.80	0.00	39704.60	102,955
	.1	0.03	0.17	0.00	2.00	102,955
Volleyball	.1–.25	0.10	0.35	0.00	4.00	102,955
	.25–.5	0.30	0.61	0.00	5.00	102,955
	.5–1	0.92	1.12	0.00	9.00	102,955

and aggravated assaults. The maximum value for the tenth-mile buffer was forty-nine.

7.5 Estimation Strategy

Our initial data set consists of fixed information on the child (race, sex, family composition at birth), changing information on the child (height, weight, and age at each clinic visit), fixed information on the parents (race, mother's and possibly father's education at the child's birth), changing information on the family (residence), the built environment near the residence in each year, crime counts by year within buffers around the child's home, and some information on neighborhood characteristics for buffers around each residence, including information on the road network, and land use, and the population density from the census.

As was mentioned earlier, to control for the variations in BMI as the child ages, we use age-sex adjusted base-period BMI z-scores as the dependent variable. We estimate two main types of models, Ordinary Least Squares (OLS) and Fixed Effects (FE) for a child at a stable address across serial clinic visits.

For the FE estimation we assume that households that stay at the same location after an amenity is placed near their residence retain the same preferences they had before the amenity was added. Under this assumption, the household fixed effect would remove constant-over-time preferences for location amenities and any other unobserved variables that did not change for each household. For example, the parents' discount rate over future consumption by either themselves or their children and their altruism toward their children would wash out in the fixed effects specification.

The key potential criticism of the FE estimation is that there are unobserved variables common to households that are located near the new amenity. If the households in a neighborhood lobbied the parks department to obtain the playground or pool built near them, then there would be some common-to-the-neighborhood but unobserved-to-the-econometrician interest in exercise that would bias the estimates. A pool placed near a neighborhood where the parents had lobbied (presumably because they were anxious to have their children use the new pool) would have a smaller effect on child overweight. This is the endogeneity problem in another guise.

More problematic is the location of privately owned amenities such as fast-food restaurants or supermarkets. These types of firms often employ market researchers to identify areas where households will be the most receptive to a fast-food outlet or the most likely to buy fresh produce. We can use robust estimators that yield consistent estimates of the standard errors when there are common-but-unobserved differences at a neighborhood level, but without the original information that was in the hands of the market researchers, we cannot fully control for differences among households in receptiveness to fast-food or fresh produce. At least the direction of any potential bias is clear. We will have upper-bound estimates on child overweight effect of these privately owned amenities. Thus, if any of them turn out to have a negligible estimated effect, we can be confident that public policy aimed at increasing or reducing these amenities would have no impact. Further, by looking at the trends in BMI *z*-score (BMI *z*) before amenities such as fast-food arrived, we can test whether the children who will gain an amenity in the future differ from those who will not.

7.6 Results

To see how much of the variation in BMI *z* can be accounted for by the fixed mother and child characteristics, BMI *z* was regressed on all of the variables in table 7.3, using robust standard errors clustered on the child's identification data (ID). Three of the year-indicator variables were significant at the 10 percent level, but these are omitted. The results are also reported separately for children under age eight and over age eight. Age is measured at the time of the clinic visit and is a continuous variable.

Table 7.3 OLS regression of fixed mother and child characteristics and neighborhood characteristics

Variable	All ages	Age < 8	Age ≥ 8
Age	0.098***	0.218***	0.220***
Age squared	–0.003***	–0.014***	–0.008***
Well-child visit	–0.052*	–0.035	–0.074*
Female	0.027	–0.050*	0.140***
White	–0.155**	–0.125*	–0.156
Black	–0.066	–0.083	–0.034
Hispanic	0.356***	0.336***	0.251
Mother's weight gain	0.001	0.001	0.001
Mother's age	0.009***	0.008***	0.012***
Birth weight	0.379***	0.434***	0.293***
WIC	0.033	0.016	0.063
Mother's marital status	0.012	0.006	0.02
Intention to breastfeed	0.008	0.02	–0.023
Mother's education	–0.015*	–0.025***	0.003
Population density	0.004	0.005	0.003
Proportion black	–0.146***	–0.226***	–0.021
Proportion high school grad.	–0.208	–0.165	–0.249
Proportion college grad.	0.268	0.340*	0.148
Median family income	–0.004**	–0.004*	–0.004*
Road network density	–0.014	–0.015	–0.018
Number of road nodes	0.457	1.395	–0.248
Prop. commercial land	0.007	0.081	–0.094
Prop. residential land	0.055	0.018	0.099
Constant	–0.890***	–1.271***	–1.581***
Observations	42,890	25,436	17,420
R^2	0.07	0.08	0.05

Note: Robust standard errors in parentheses.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

About 7.4 percent of the overall variation in BMI *z* can be accounted for by fixed child characteristics, mother's characteristics, and neighborhood characteristics. The explanatory power is 8 percent for the younger children and 5 percent for the older children. The increased explanatory power of the model for younger children may be attributable to the birth certificate data more accurately representing the current socioeconomic environment of the study subject. The well-child visit indicator is significant overall and for the older children; the sign is negative. The negative association between child weight and well-child care is counterintuitive. The well-child variable, in theory, represents the health status of the child, with poorer-health-status children having systematically lower body mass index. Recall that visits with diagnostic codes known to affect body weight were dropped from the data set (pregnancy, endocrine disorders, cancer, congenital heart disease, chromosomal disorders, and metabolic disorders).

The well-child variable may reflect behavior practices of the child's caregivers. Caregivers who less frequently access routine health maintenance for their children, or primarily bring their children in for sick-child visits, may be less supportive of child health behaviors associated with optimal child weight (e.g., promoting routine physical activity or a nutritious diet). A health maintenance organization (HMO) study found that overweight children were less likely to have well-child visits (Estabrooks and Shetterly 2007).

Relative to the reference population, BMI z is increasing rapidly with age. As children age from the sample minimum of three to the maximum of sixteen years, their predicted BMI z increases by 0.6. The mean BMI z-score for children in the age range three to four is 0.43, while for children in the age range fifteen to sixteen it is 0.85. The age and age squared specification is parsimonious, but a histogram, in Figure 7.3, suggests a rapid increase up to age thirteen, and level BMI z thereafter.

The BMI z gain appears to be largely a permanent cumulative process. Children (Wilfley et al. 2007) and adults (Jeffery et al. 2000) are often able to lose weight in the short-term, but find it much more difficult to sustain any loss from their peak weight over a long period. Large recorded z-score gains are rarely reversed at later visits.

To assess how often large gains in weight were reversed, we looked at the subset of children with large BMI z-gains, defined as + 0.5 in BMI z from the first visit to the second visit. At the mean of the reference population a z-score gain of 0.5 would be 27 percentile points. The count of big gainers that have at least three visits was 3,381. Among these big gainers, the count of those who were above their initial z score by the third visit was 2,743. The big gainers who were at or below their first-visit z-score by the third visit was 638. Only 19 percent of the big gainers recovered.

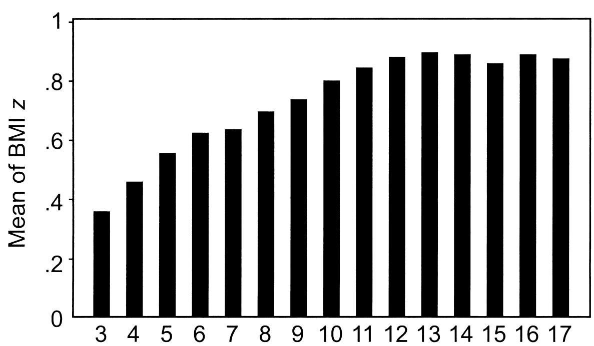

Fig. 7.3 Mean BMI *z* by age

There is some noise in the weight and the height data that are likely due to data entry or measurement errors at the clinics. If observations of big gainers between the first and second visits were primarily due to such errors, we would expect them to largely disappear by the next visit. Because only a small proportion of the big gains were reversed, we can be confident that they are not primarily due to recording errors. Further, the irreversibility of most of the big gains supports our conclusion that the weight gains are largely cumulative.

The age effect, due to tendency of children to accumulate weight relative to the reference population and rarely lose any of it, is quite strong. Consequently, we will include age and age squared in subsequent regressions, which always have a maximum age of sixteen. What these regressions can tell us is the extent to which the addition or removal of an amenity can alter the pronounced BMI *z*-age pattern.

The birth certificate variables are available for only one third of the data set. This is too small a sample size to detect many amenity effects. In table 7.4, we drop the birth certificate variables and add the amenity variables.

The amenities that are significant show up at various distances. There is no reason to expect the real effects of different amenities to operate over the same distance. Also, in the smaller circles there may be real effects but too few observations to yield statistically significant results. In discussing the results in table 7.4, we concentrate on the signs and significance levels rather than the values of the coefficients. We do this because we are primarily interested in how the OLS results contrast with the FE results. We think the OLS results are telling us more about who chooses to live near an amenity, such as a school with open recreational facilities or a fast-food restaurant that is near a major road.

Age and age squared are highly significant in every OLS regression. In cross-section, the well-child visit indicator is always negative. It is significant in six of the nine regressions, including all of the regressions on the over age eight children. The female indicator variable should not be significant because the BMI *z* variable adjusts for gender in the reference population. The consistent negative and significant coefficients for the younger children and the positive and significant coefficients for the older children indicate that relative to boys in the same age range, the younger girls are not gaining BMI *z* as fast while the older girls are gaining BMI *z* faster than older boys. The differential may be due to a trend toward an earlier age of puberty for girls.

In terms of racial differences, the striking result is the high BMI *z* values for Hispanics relative to the omitted category—Asian and other. There are two variables related to African Americans. One refers to the race of the child. Older black children are significantly heavier than the omitted category and than whites. The second variable refers to the neighborhood. Other things equal, living in a neighborhood with a higher proportion of

Table 7.4 **OLS with birth certificate variables excluded**

	Tenth of mile			Quarter mile			Half mile			One mile		
Variables	All ages	Age < 8	Age > 8	All ages	Age < 8	Age > 8	All ages	Age < 8	Age > 8	All ages	Age < 8	Age > 8
Age	0.093***	0.185***	0.217***	0.093***	0.185***	0.216***	0.093***	0.184***	0.217***	0.093***	0.188***	0.212***
Age squared	–0.003***	–0.012***	–0.008***	–0.003***	–0.012***	–0.008***	–0.003***	–0.012***	–0.008***	–0.003***	–0.012***	–0.008***
Well-child visit	–0.026	–0.001	–0.053**	–0.027	–0.001	–0.054***	–0.028*	–0.003	–0.054***	–0.028*	–0.004	–0.054***
Female	0.002	–0.072***	0.081***	0.002	–0.071***	0.080***	0.003	–0.070***	0.081***	0.002	–0.071***	0.079***
White	0.02	–0.013	0.078*	0.018	–0.017	0.078*	0.02	–0.012	0.076*	0.022	–0.009	0.077*
Black	0.007	–0.089**	0.114***	0.005	–0.094**	0.117***	0.005	–0.091**	0.115***	0.01	–0.087**	0.119***
Hispanic	0.366***	0.371***	0.298***	0.367***	0.369***	0.300***	0.368***	0.372***	0.300***	0.365***	0.371***	0.295***
Population density	0.004	0.004	0.003	0.004*	0.003	0.006*	0.004	0.004	0.005	0.007***	0.008***	0.007**
Proportion black	–0.110***	–0.221***	0.01	–0.122***	–0.228***	–0.003	–0.108***	–0.216***	0.017	–0.141***	–0.217***	–0.053
Proportion with college	–0.057	–0.038	–0.08	–0.065	–0.048	–0.063	–0.026	0.029	–0.063	–0.215*	–0.148	–0.275
Family income	–0.002*	–0.001	–0.002*	–0.001	–0.001	–0.002*	–0.001	–0.001	–0.002	–0.001	–0.001	–0.001
Proportion residential	–0.061	–0.087	–0.052	–0.062	–0.069	–0.076	–0.073*	–0.091*	–0.08	–0.057	–0.073	–0.058
Fast-food	0.082**	0.041	0.134***	0.019*	0.01	0.030**	0.003	0.002	0.003	–0.003	–0.004	–0.002
Supermarkets	–0.185**	–0.195**	–0.168	–0.013	–0.036	0.021	0.023	0.027	0.019	0.011	0.015	0.006
Convenience stores	0.07	0.022	0.121**	–0.02	–0.01	–0.03	0.012	0.012	0.015	–0.007	–0.003	–0.011
Trails	–0.746	–1.321*	–0.066	–0.266**	–0.331**	–0.128	–0.053	–0.032	–0.043	0.030*	0.035	0.028
Crime	–0.009	0.21	–0.171	0.033	0.084	–0.015	0.009	0.015	0.003	0.003	0.001	0.006
Baseball/softball	0.006	–0.02	0.037	–0.002	–0.001	–0.003	0.002	–0.001	0.006	–0.003	–0.008***	0.003
Basketball	–0.004	0.01	–0.018	–0.002	0.005	–0.010*	0.003	0.009***	–0.003	0.001	0.001	0.002
Family centers	–1.544	0.038	–3.811	0.614	0.402	0.736	0.645***	0.871***	0.342	0.204**	0.370***	0.002

Fitness areas	–0.716	–1.539*	–0.318	0.25	0.071	0.442	0.124	–0.068	0.355*	0.029	–0.002	0.067
Football	–0.078	–0.046	–0.119	0.016	0.031	–0.002	0.009	0.022	–0.006	0.001	0.007	–0.006
Kickball	0.112	0.172*	0.038	0.042	0.073**	0.005	0.007	0.016	–0.007	0.006	0.001	0.01
Playgrounds (no equipment)	–0.329	–0.355	–0.238	–0.056	–0.063	–0.056	–0.043	–0.026	–0.082	–0.036	0.008	–0.085*
Playgrounds (with equipment)	–0.042	–0.301	0.311	0.018	0.024	0.04	0.001	–0.026	0.028	–0.022	–0.001	–0.038
Pool	–0.808	–1.292	–0.08	–0.517	–0.119	–0.929*	–0.215	–0.267	–0.126	0.154	0.144	0.151
Soccer	0.055	0.237	–0.134	0.002	–0.014	0.015	–0.009	–0.004	–0.01	0.002	0.004	0.002
Tennis	–0.014	–0.029*	0.004	–0.002	–0.005	–0.001	–0.003	–0.003	–0.004	–0.003**	–0.002	–0.004**
Track and field	0.294	0.09	0.729***	0.004	–0.007	0.022	–0.013	–0.011	–0.016	0.011	0.022	–0.002
Volleyball	0.068*	0.094**	0.031	0.009	0.002	0.01	0.016	0.008	0.018	0.009	–0.001	0.019**
Constant	0.242***	0.053	–0.557**	0.235**	0.043	–0.550**	0.196**	0.008	–0.605***	0.164	0.041	–0.620***
N (observations)	96,522	50,503	45,951	96,522	50,503	45,951	96,522	50,503	45,951	96,522	50,503	45,951
R^2	0.03	0.04	0.01	0.03	0.04	0.01	0.03	0.04	0.01	0.03	0.04	0.01

Notes: Robust standard errors in parentheses.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

blacks is associated with a lower BMI z. Since proportions run from 0 to 1, the interpretation of the coefficient is straightforward. For children under the age of eight, hypothetical neighborhoods with no African Americans have higher BMI z, by about 0.22, than neighborhoods that are entirely African American. The proportion of residents in the neighborhood with a college education is almost never significant (one of nine at the 0.10 level). The median family income in the neighborhood and proportion of dwellings that are residential are similar (both have two of nine at the 0.10 level). In the neighborhoods our children live in, college education is rare and incomes are generally low.

The fast-food variable is significant in four of the nine OLS regressions. The significant coefficients are always positive. This positive effect on BMI is the conventional result for fast-food in cross-sectional regressions. The significant coefficients are for the closer buffers, within 0.1 and 0.25 miles. The supermarket variable (OLS) also has the conventional result that supermarkets are associated with lower BMI z when they are close. The signs on the first five coefficients are negative. Of those, just two are significant. Supermarkets also tend to be located on major roads that are on commuting routes. Households that live near a supermarket are likely to differ from households that live far from the nearest supermarket. The convenience store variable has little explanatory power in the OLS regression. Only one of the nine coefficients is significant. Crime is never significant in the OLS regressions.

Very few of the recreational amenities have a significant negative sign. These include trails (< 8 at 0.1 miles, all ages at 0.25 miles, and < 8 at 0.25 miles), baseball/softball (< 8 for 1.0 mile), pools (> 8 at 0.25 miles), and tennis (< 8 at 0.1 miles, all ages at 1.0 miles). Even some of the results that do have a negative and significant sign are counterintuitive; for example, how many children under eight play tennis?

The problems with the OLS results are that they have little explanatory power, most of the demographic variables have limited policy implications, and most importantly, it is impossible to know if the associations are causal. For example, track and field facilities and football fields are almost all located at middle schools and high schools. Even if they had been statistically significant, would the BMI z differences associated with these variables be due to children using these amenities, or simply to unobserved differences in the families that chose to live near these schools? The fast-food restaurants, supermarkets, and even convenience stores tend to be located on major roads that are commuting routes from the city center to the suburbs. Later, we provide some clear maps showing these amenities lined up on the commuting routes. Are the BMI z associations of these amenities due to proximity to these food sources or to unobservable differences in households living near major roads?

Table 7.5 Counts of children having any change by amenity and by buffer

	0.1 mile	0.25 mile	0.5 mile	1 mile
Fast-food	29	342	1,446	4,980
Supermarkets	8	79	337	1,290
Convenience stores	33	270	1,066	3,686
Trails	73	258	715	2,085
Crime	14,643	17,782	18,923	19,946
Baseball/softball	50	371	1,168	4,252
Basketball	179	1,041	3,519	9,655
Family centers	0	16	429	614
Fitness	18	39	271	477
Football fields	4	64	232	1,070
Kickball	35	187	622	2,112
Playgrounds no equipment	143	572	3,596	5,764
Playground with equipment	483	1,942	7,561	13,601
Pools	93	329	2,702	3,447
Soccer fields	19	192	1,733	3,847
Tennis	84	250	734	2,133
Track and field	7	28	835	939
Volleyball	14	78	299	1,019

7.6.1 FE Regressions

Before reporting the FE regressions we report in table 7.5 following how many children had changes in each of the amenities.

The large number of children having changes in the amounts of crime is due to the underlying variable counting individual crimes. At the smallest buffer, within 0.1 miles, many of the amenities have so few children facing changes that it is unlikely we would observe any effect. These include supermarkets, family centers, fitness areas, football fields, soccer fields, track and field, and volleyball courts. Except for the family centers, by the 0.25-mile buffer there are enough children with observed changes that if changes in the amenity indeed had an effect on BMI, at that distance we would have a good chance of observing the effect. We added the 0.1-mile buffer for all amenities because the Currie and colleagues paper found a fast-food effect within 0.1 miles of a child's school. The vast majority of the changes in amenities that are in counts were a gain or loss of one unit; for example, one fast-food restaurant. Of these, the modal change was from 0 to 1 unit with the next most frequent being from 1 unit to 0.

Data on individual children have more variation than data on the average of the all ninth graders in high schools. Our sample of children is 3.2 percent of the Currie and colleagues sample. Their sample is based on observations of 3.06 million student-years, and our data on 98,541 clinic visits among children with two or more visits while residing at the same address. We have twenty-nine changes in the number of fast-food restaurants within the 0.1-mile buffer compared to twenty-two in the Currie and colleagues sample.

In the FE regressions reported in table 7.6, we dropped all variables that are constant at the level of the child. Again, this sample is restricted to observations in which a child remains at the same address between clinic visits. The same restrictions on age and biologically implausible values of BMI *z*, height, and weight used in the cross-section regression are applied in the FE regressions. The covariates to the environmental/amenity variables are age and age squared, year of clinic visit indicator variables, an indicator for a well-child visit, and crime.

Again, the coefficients on the year dummy variables are not reported. In the FE regressions, the children under eight years of age are gaining BMI *z* faster relative to the reference population than the children over age eight, roughly by 0.13 BMI *z* units a year. This younger versus older child differential did not appear in the OLS regressions. The well-child variable is now positive and significant at the 10 percent level for all children. This FE result sharply contrasts with the OLS result for the well-child visit variable, which was always negative, and significant in six of the nine OLS regressions. Thus, for a given child a non–well-child visit is associated with a lower weight than a well-child visit.

There are very few overlaps from the OLS to the FE results of the same amenity being significant at the same distance. Adding a fast-food restaurant within a quarter mile of the same child appears to significantly *reduce* the child's BMI *z*. Recall that in the cross-sectional results at the tenth-mile buffer, the association between BMI *z* and fast-food was positive.

From a public policy perspective, the FE results for the recreational amenities are somewhat discouraging. The variables with negative and significant coefficients are fitness areas for all children and younger children at 0.25 miles; kickball for all children and younger children at 0.1 mile, all children at 0.25 miles, and older children at the 0.5 and 1.0 mile buffers; playgrounds without equipment for younger children at 0.5 miles; tennis for older children at 0.25 miles; and volleyball for older children at the 0.1 mile buffer and older children at the mile buffer. The division across amenities that might be associated with reducing BMI *z* in younger versus older children appears plausible; for example, younger children use playgrounds and kickball fields, while the older ones use volleyball courts.

As a check on whether their estimated fast-food effects on percentages of boys in a high school who were overweight (defined as the eighty-fifth percentile) could plausibly be due to the calories from an extra fast-food meal per day, Currie and colleagues calculated the weight *gain* required for a median height fourteen-year-old boy to move from the eightieth to the eighty-fifth percentile of the BMI distribution. This weight gain was 3.6 pounds. To get a sense of what our estimated coefficients imply for weight gains we will use boys, to match Currie and colleagues, but change the age to

Table 7.6 Fixed effects regressions

	Tenth of mile			Quarter mile			Half mile			One mile		
Variables	All ages	Age < 8	Age > 8	All ages	Age < 8	Age > 8	All ages	Age < 8	Age > 8	All ages	Age < 8	Age > 8
Age	0.117***	0.279***	0.151***	0.117***	0.280***	0.153***	0.117***	0.280***	0.153***	0.118***	0.279***	0.155***
Age squared	–0.003***	–0.017***	–0.005***	–0.003***	–0.017***	–0.005***	–0.003***	–0.017***	–0.005***	–0.003***	–0.017***	–0.005***
Well-child visit	0.014*	0.019	0.007	0.014*	0.02	0.008	0.013*	0.018	0.008	0.013*	0.017	0.007
Fast-food	–0.134	–0.074	–0.109	–0.077***	–0.084	–0.038	–0.021*	–0.037	–0.012	0.015**	0.024**	0.003
Supermarkets	0.052	–0.169	–0.255	–0.046	–0.054	–0.096	0.028	0.044	0.042	0.01	0.028	0.043***
Convenience stores	0.009	–0.096	–0.004	0.029	0.024	0.011	0.004	0.036	–0.025*	0.013*	0.036**	–0.007
Trails	–0.557	–1.214	1.802**	0.014	–0.333	0.368**	0.04	–0.056	0.088	0.017	0.023	0.033*
Crime	–0.098	0.069	–0.186	–0.096**	–0.162**	–0.057	–0.050***	–0.088**	–0.023	–0.013	–0.031**	0.002
Baseball/softball	0.081*	0.187**	–0.008	0.013	–0.006	0.026	–0.001	0.015	–0.011	–0.005	–0.013	0.008
Basketball	0.001	0.01	–0.035	–0.01	–0.007	–0.015	–0.003	–0.007	0.004	0.001	–0.002	0.004*
Family centers	–	–	–	–0.812	–6.09	0.659	1.099	1.124	1.122	–0.818*	–0.184	–0.18
Fitness areas	–12.278	–62.440**	25.365	–2.262***	–4.813***	0.651	0.095	–0.247	0.385	0.07	0.182	0.077
Football	0.433	0.507	–0.082	0.09	0.074	–0.007	0.104***	0.116*	0.015	–0.006	–0.001	–0.01
Kickball	–0.322***	–0.416**	–0.049	–0.084**	–0.103	–0.046	0.008	0.04	–0.047**	–0.004	0.013	–0.048***
Playgrounds (no equip.)	–0.28	–1.434	2.643***	0.08	–0.571	0.464**	–0.007	–0.478**	0.296**	–0.056	–0.112	–0.013
Playgrounds (with equip.)	0.516	1.365	–0.257	0.851***	1.291***	0.393	0.416***	0.706***	0.072	0.029	0.113	–0.037
Pool	–1.49	–3.33	–2.08	–1.149	–2.097	–0.169	–0.147	–0.12	–0.949	0.458	1.228*	–0.205
Soccer	–0.067	0.042	–0.133	0.016	–0.059	0.024	0.015	0.027	0.003	0.015**	0.026**	0.006
Tennis	–0.014	0.004	–0.008	–0.003	0.014	–0.027**	0.005	0.014	–0.003	0.005	0.001	0.006
Track and field	8.515	12.495	–	–0.076	0.193	–0.029	0.156	0.364*	0.143	0.091*	0.201**	0.073
Volleyball	0.09	0.113	–0.904***	–0.018	–0.073	–0.074	0.038	0.051	–0.021	–0.013	–0.026	–0.030*
Constant	–0.218	–0.622*	–0.343***	–0.174	–0.497***	–0.369***	–0.310***	–0.739***	–0.373***	–0.304**	–0.844***	–0.440***
N (observations)	98,541	50,521	47,952	98,541	50,521	47,952	98,541	50,521	47,952	98,541	50,521	47,952
N (child/address)	54,823	30,304	26,615	54,823	30,304	26,615	54,823	30,304	26,615	54,823	30,304	26,615
R^2	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Note: Standard errors in parentheses. Dashed cells indicate that the variable was dropped because there were no children who gained or lost a family center within a tenth of a mile from their home.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

8, which is the median for our data. We will start at the eighty-fifth percentile, their end point, and calculate the implied weight *loss* for some amenities that were estimated to statistically significant effect in reducing weight. Adding a kickball diamond within a tenth of a mile is associated (based on the equation for all ages) with a reduction of 2.8 pounds. The weight reduction for adding a playground within a half mile (based on the under age eight regression) is 4.1 pounds. The weight reduction for adding a volleyball court within a tenth of a mile (based on the age eight or over regression) is 6.9 pounds. Recreational amenities that could reduce the weights of overweight eight-year-old boys within a year of being located near their homes by anything in the range of 2.8 to 6.9 pounds would be economically significant.

Switching to the statistically significant effects for food vendors, at a mile distance the addition of a fast-food restaurant was associated (in the all-ages regression) with a tiny weight gain, 0.14 pounds. The addition of a supermarket within a mile (all-ages regression) is associated with a gain of 0.42 pounds. The addition of a convenience store within a mile (under age eight regression) was associated with a gain of 0.36 pounds. The weight changes associated with adding a food vendor, even when statistically significant, are smaller than the weight losses associated with the few recreational amenities that have negative and significant coefficients.

Some of our results are counterintuitive. Fast-food is associated with weight reduction for all ages at a quarter mile. Trails are only significant for the older children. This trails result is partly intuitive because younger children walking on the trails could wander into the paths of runners, bicyclists, and in-line skaters. We see more young children riding in strollers or in bicycle carriers or on tandem bicycles than those traveling entirely on their own power. However, the counterintuitive part is that all of the coefficients that are significant have a positive sign.

If the reported results were causal effects, then BMI z-reducing policy would be to build fast-food restaurants within a quarter mile of the child's home and surround the child's home with a fitness area, a kickball diamond, and a playground, all at their respective optimal distances. Before much credence can be given to these estimates, the issue of the endogeneity of the placement of these amenities must be addressed.

The FE framework allows for separate consideration of gains and losses in amenities. We tested whether the coefficient on a gain was the same as for a loss for every amenity, and could not reject the null hypothesis of equality in a single case. Also, we looked at assumption of linearity of effects; for example, that a gain from 0 to 1 is the same as a gain from 1 to 2. A very high fraction of all of the changes we observed in counts of amenities is in the range of 0 to 1 or from 1 to 0. We could not reject the null hypothesis of linearity largely because we observed too few higher-order changes.

7.6.2 Endogeneity of Amenity Location

The sharp differences in significance levels and signs between the OLS and FE regressions raise questions about the endogeneity of the location of food vendors. As pointed out in the literature review, vendors may choose to open new fast-food restaurants, for example, in areas in which a larger percentage of the population is overweight. Figures 7.4, 7.5, 7.6, and 7.7 shed light on this issue by showing fast-food, supermarket, limited-service restaurant, and convenience store locations and changes in location, respectively. Some background on the geography of Indianapolis is necessary to interpret these figures. The borders of Indianapolis and Marion County are the same. The metropolitan area includes the surrounding counties, which are not shown in these figures. The population of the metropolitan area in 2006 was 1.66 million. The city proper had a population of 786,000 in same year. The city has a single central business district located at its geographic center. The original street plan for the city had major roads set at angles that can be interpreted as time on a clock. For example, the road running from the center straight north to 12:00, Meridian Street, is the main commuting route from suburbs directly north of the city. The one running from 3:00 to the center of town, East Washington Street, is the major commuting artery for suburbs directly to the east of the city. The roads at approximately 12:30,

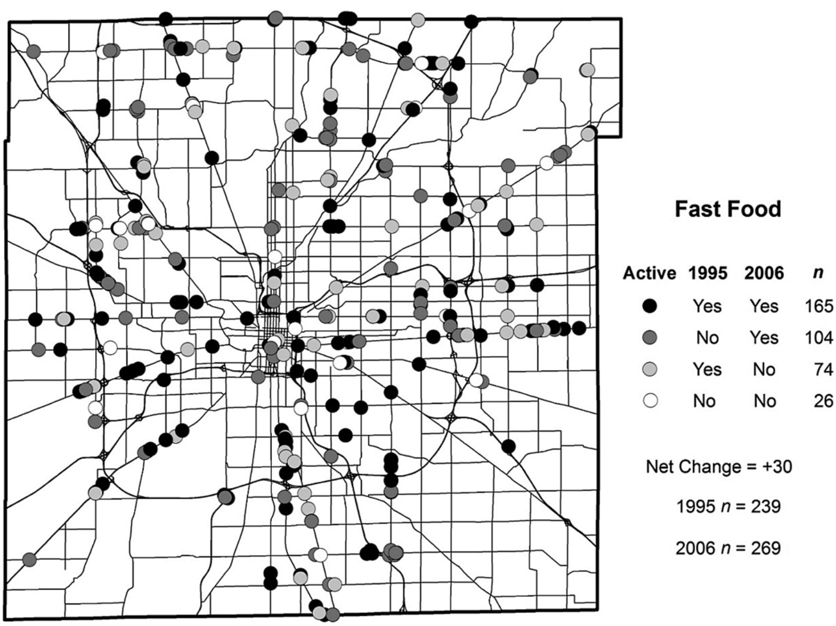

Fig. 7.4 Fast-food locations and changes in Indianapolis

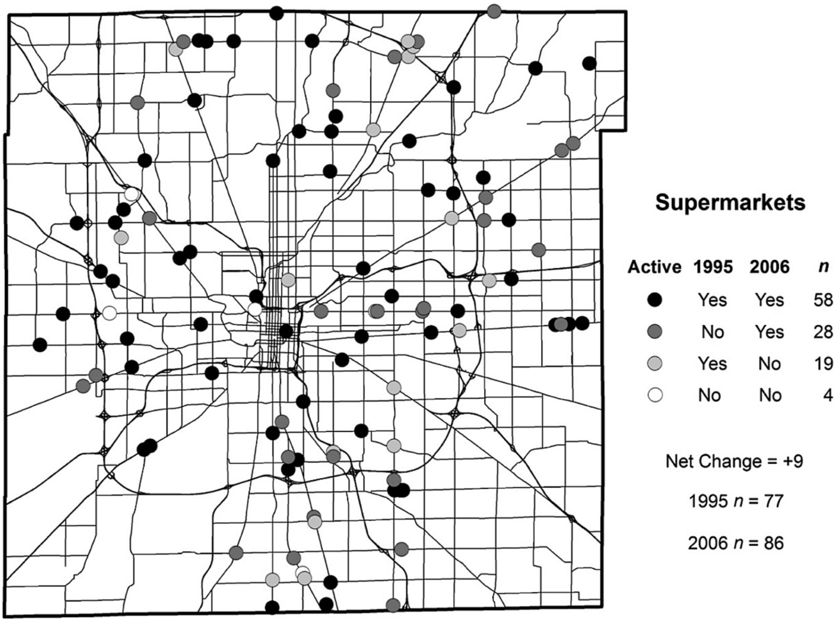

Fig. 7.5 Supermarket locations and changes in Indianapolis

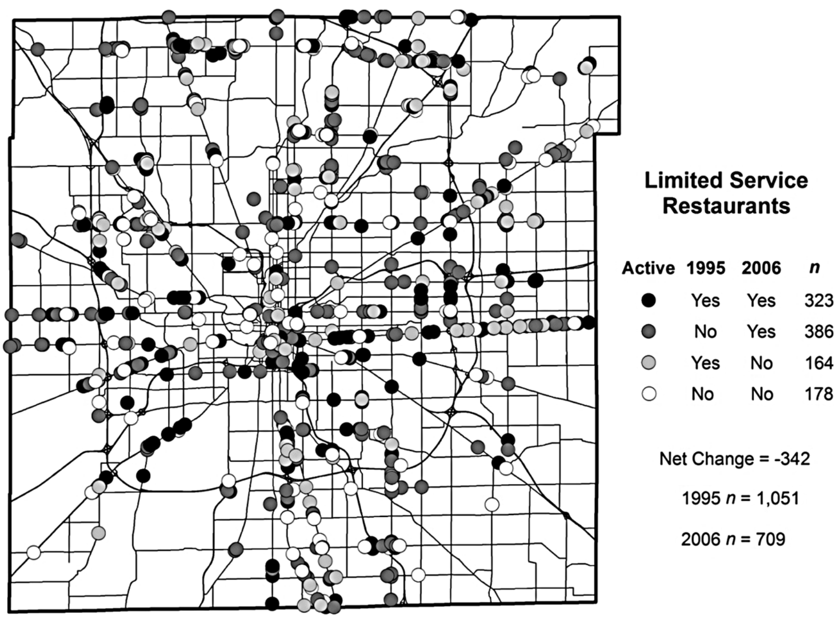

Fig. 7.6 Limited-service restaurant locations and changes in Indianapolis

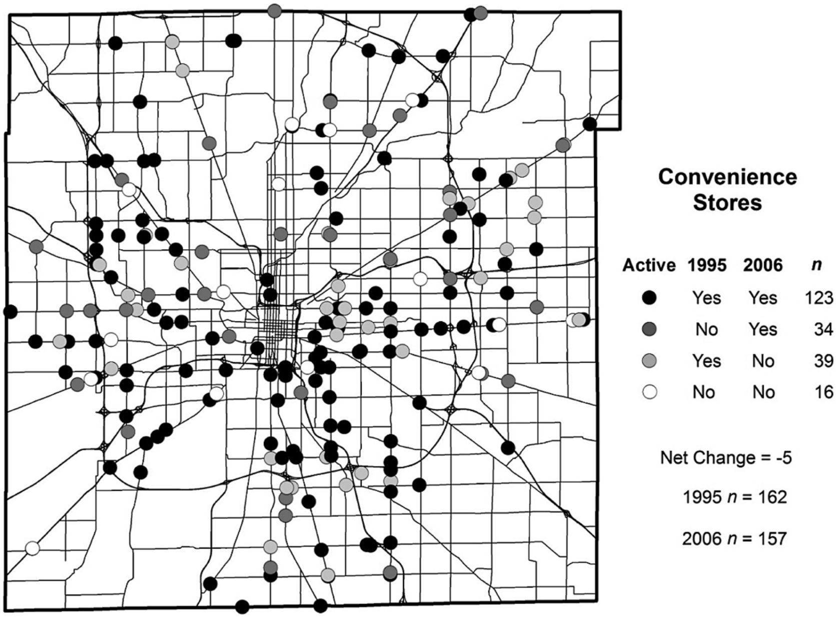

Fig. 7.7 Convenience store locations and changes in Indianapolis

1:30, 4:40, 5:50, 7:30, 8:30, 9:00, 10:30, and 11:00 are also major commuting routes supporting the large traffic to and from the suburbs.

Recall that we defined the fast-food establishments as belonging to national chains. These high-volume restaurants are clearly concentrated on the major surface roads leading in and out of the city center. Traffic flow data would be useful as an instrument to predict fast-food location. Unfortunately, public traffic flow data are outdated and have limited and highly uneven coverage.

The supermarkets, shown in the map in figure 7.5, are also located primarily along major streets. The difference between the supermarkets and the fast-food restaurants is a relative dearth in the inner city (the poorest area). This supermarket "desert" is a residential area immediately north of the central business district. In contrast, the fast-food restaurants are well represented in the inner city.

Figures 7.6 and 7.7 depict locations and changes of limited-service restaurants and convenience stores. The southwest and southeast corners of the county are still largely rural. Other than in those undeveloped areas, there are limited-service restaurants and conveniences stores widely distributed across the county.

One means of addressing the endogeneity of amenity locations is to examine the relationship between the arrival of new amenities in year *T* and children's BMI z-levels and trends in years prior to that year. Under the assumption that the future cannot cause the past, significant relationships indicate that location decisions depend on the body weight of potential customers. To accomplish this, we ran two sets of regressions as follows:

$$\text{(1)} \qquad \text{BMI}z_{i,t} = \alpha + \beta X_{i,T} + e_i, \qquad t = 1, \ldots T\text{-}1,$$

$$\text{(2)} \qquad \text{BMI}z_{i,t} = \alpha + \beta X_{i,T} + \gamma t + \lambda\,(t * X_{i,T}) + e_i,$$

where X_{iT} is a dichotomous indicator that equals 1 if, for example, a fast-food restaurant opened in the area in which the ith child resides in year T. A positive and significant value of β in equation (1) means that the restaurant is more likely to open in an area in which children had high BMIs in the past. A positive and significant value of λ in equation (2) means that the opening of a new restaurant is more likely in an area in which children's BMI had been growing more rapidly in the past.

The results are reported in table 7.7. For each amenity, the top line reports the sign of the estimate of β from equation (1) provided it was significantly different from zero at a 5 percent significance level. Similarly, the second line of each panel reports the sign of significant estimates of λ from equation (2). The results show that only the location of supermarkets is preceded by differences in children's weight, as well as differences in trajectories of children's weight gain.

The positive trends observed at all four buffers for supermarkets undercut the claim that their new locations were selected independently of the prior changes in children's BMI z. Thus, the FE results that supermarkets increase children's weights at the half-mile buffer are suspect. Fast-food restaurants appear to be entering areas with higher child BMI z values and higher rates of child obesity, at least for the quarter- and half-mile buffers. However, these initial differences in levels may not predict the change that will occur after the arrival of a new fast-food restaurant. Our assumption is that gains in BMI z will be the same for a given stimulus over a broad range of initial BMI z. We believe having the same trend in BMI z for children with and without future fast-food gives us an unbiased estimate of the response to the arrival of an amenity. Our negative coefficient quarter-mile fast-food result along with no difference in BMI z trends prior to arrival of the fast-food align with the Anderson and Matsa result cited earlier. While fast-food meals are notoriously calorie-dense, they can have no BMI z effect if children or adults offset the additional calories by eating less food at other meals or by eating fewer meals. Alternatively, there may be so much fast-food in Indianapolis that any child so inclined could readily access a fast-food restaurant whether one was within a tenth of a mile or a quarter mile or not. Either way, as a means of attacking the child obesity epidemic, the Los Angeles freeze on new fast-food restaurants mentioned in the introduction may be misplaced.

Table 7.7 Signs of significant coefficients for future amenities

		Radius			
		.1	.1–.25	.25–.5	.5–1
Fast-food restaurants	BMI *z*		+	+	
	Trend				
Trails	BMI *z*		+		
	Trend			+	
Supermarkets	BMI *z*		+	+	+
	Trend	+	+	+	+
Convenience stores	BMI *z*		+	+	
	Trend			+	
Parks	BMI *z*	–	–	–	–
	Trend	+		–	–
Baseball/softball diamonds	BMI *z*	–			+
	Trend				+
Basketball courts	BMI *z*				–
	Trend		+	+	
Family centers	BMI *z*				–
	Trend				
Fitness centers	BMI *z*				–
	Trend				+
Football fields	BMI *z*				
	Trend			+	
Kickball diamonds	BMI *z*		+		–
	Trend	+		+	
Playgrounds with no equipment	BMI *z*				
	Trend				
Playgrounds with equipment	BMI *z*				
	Trend				
Pools	BMI *z*		+		
	Trend	–			
Soccer fields	BMI *z*				
	Trend	+			
Tennis courts	BMI *z*			+	+
	Trend				
Track and field	BMI *z*		–		–
	Trend				
Volleyball courts	BMI *z*				
	Trend		+		–

Fast-food and supermarkets are the highest-profile amenities. Of the remaining sixty trend terms (fifteen amenities times four buffers), eleven are significant. These are scattered such that none of the other amenities has a significant trend term for more than one buffer. Either the locations of these remaining amenities are not being selected on the basis of differences in BMI *z* trends, or we do not have enough data to detect differences in BMI *z* trends.

7.7 Conclusion

Our first conclusion is that cross-sectional results differ dramatically from the FE results. We believe that the cross-sectional results tell us more about who chooses to live near an amenity than what adding that amenity might do. In cross-section, nearby (tenth-mile) fast-food increases children's BMI z. Our cross-section regression has controls for child's age, race, gender, mother's age at child's birth, mother's education, WIC eligibility, intention to breastfeed, and many neighborhood characteristics. These are as comprehensive a set of covariates as we have seen for child BMI regressions. Other study strengths include directly measured height and weight data for a large sample size that includes high proportions of African American and Hispanic children. Still, in the fixed effects framework, nearby (quarter-mile) fast-food appears to reduce children's weights, with no difference in the trend of BMI z gain prior to the arrival of the fast-food. While we doubt that fast-food really reduces children's BMI z, the results of the fixed effects models cast doubt on the highly publicized policies to reduce fast-food exposure as interventions for preventing obesity.

A second conclusion is that if the arrival of amenities (other than supermarkets) is unrelated to prior trends in BMI z, then there appears to be little in the way of surefire interventions for reducing children's BMI z, through either recreational amenities or food vendors. The best candidates appear to be fitness areas, kickball fields, and volleyball courts. Weight reductions for overweight children (defined as at the eighty-fifth percentile of the pre-epidemic distribution) in the range of three to six pounds, as estimated for eight-year-old boys for these amenities, would be valuable results for an intervention.

Our results look at the short-term. They look for BMI z responses within the year the amenity arrives. It may be that a recreational amenity does have a BMI z-reducing effect on nearby children if it is measured years after its arrival. However, we have few observations with long runs of time after the arrival of an amenity.

Further, our study demonstrates the benefits of an interdisciplinary team of economists, a physician, an urban planner, and geographers. It would have been impossible to assemble these data without the interdisciplinary collaboration.

7.7.1 Future Work

The present chapter is our first effort in using these data. We mentioned earlier that our estimated effects were short-term, specifically within a year, and that we intend to look for persistent effects from changing amenities.

A general assumption of our methods used in this chapter is that proximity is a proxy for exposure to amenities. We do not have direct observational data on whether or not children and their families use (or are even aware of) the amenities we measured. In future prospective work, we hope to collect detailed observational data on spatial and temporal interaction with amenities through survey and global positioning system (GPS) tracking. This may allow us to better infer causal effects of the built environment on children's weight.

National Institutes of Health (NIH)-funded R21 studies are meant to test the feasibility of a new research design. Our study demonstrates that it is feasible to collect detailed longitudinal data on selected components of the built environment surrounding the homes of a large sample of children in a metropolitan area. In total, it took our team about twenty months to assemble and clean the built-environment data used in the analytical portion of this study. As spatial information technologies continue to become more widespread and agencies (such as police, parks, and food safety departments) increasingly collect and organize data on amenities in forms that are easily extended to spatial analysis, it should be easier to extend the methods used in this study both spatially (to include larger populations in multiple cities) and temporally (to include longer-term longitudinal experiments). We plan to seek funding for a six-city extension of the present study. A six-fold increase in the sample size over different regions of the country would provide much more reliable results.

Our study examined associations between BMI z and proximity to amenities within four buffer distances. We used relatively simple methods to measure spatial proximity—straight line (Euclidean) buffers. In future work we will explore more complex measures of proximity, including network buffers and travel time models that consider movement along street networks. This will allow us to test other specifications for built-environment variables, including specific distances or travel times to individual amenities, average distance or time to the closest three amenities of a given type, and more general measures of accessibility to amenities.

Appendix A

Photo Interpretation

Orthorectified aerial photograph mosaics were available for most study years. The primary exception was 1996, for which no photographs of Marion County were available. Small, scattered areas were missing from the 1995 photographs. Photographs were available for 1998 for most of the county. Owing to a problem with the initial flight for 1998, no photography was available for a narrow band of the county running north to south through the center. This area was reflown in early 1999 using the same techniques; the resulting photographs were treated differently in this study from the 1998 photographs, as well as from a complete set of photographs taken later in 1999 using different techniques.

The photographs varied greatly in quality; only the last four years were in color, and later years were generally of higher spatial resolution than earlier years, though the 1998 (and matching early 1999) photographs were significantly coarser than other years. Contrast was also starker and thus of lower quality in earlier years, particularly in 1995.

All photographs were provided to a team of photo interpreters in digital format. In one case, 2005, images were available from Indiana University's spatial data portal, representing the same original photography as the 2005 file photographs, but reproduced at higher resolution. Photographs were examined on computer display in ArcMap 9.2, generally at scales of 1:1000 to 1:2500. Accompanying this was a copy of the amenities database (described earlier in the section on data), in which interpreters were to save any changes.

The county was divided into interpretation areas, usually consisting of a linear strip half a mile wide, running north to south; each area was assigned to a particular interpreter, though much of the county was eventually analyzed by more than one interpreter. The tasks of the interpreter were two: first, to locate recreational amenities of the selected types and add them to the database (through heads-up digitizing), and second, to determine during which years each amenity was present. This information was recorded as attributes of the feature in the database, along with information on the type of amenity, and the source of the feature, whether a particular interpreter or one of the original files. If a previously digitized feature was modified in shape, size, or location by an interpreter, that information was noted as well. If interpreters were unable to determine the presence of an amenity (owing to the absence of photographs), a no-data value of –9 was recorded as that year's attribute. Finally, features that lay on the border between one interpretation area and another were flagged with a special attribute, so that duplication could be avoided. As a quality check on individual interpreters, border features were to be digitized by *both* interpreters, and the results compared.

The study team decided to quantify each amenity type in a way judged most likely to capture the relative recreational opportunities that each provided, with a few practical constraints. In the case of standardized playing areas, such as tennis courts and football fields, a count of amenities was deemed appropriate. In the case of scalable amenities, such as swimming pools and playgrounds, the area of the amenity was deemed the best measure. The opportunity available to each child would be taken as the sum of these measurements, as they fell within a given distance of the child's home address. For example, if 23 m^2 of a swimming pool fell within a half-mile radius around the child's address, this 23 m^2 was added to the value of the child's swimming pool opportunity.

Guidelines for Digitizing

- For baseball/softball, basketball, football, kickball, tennis, track and field, and volleyball, the playing area was to be digitized, as marked where possible.
- For baseball and softball, the boundaries of infield and outfield were to be digitized; where the outfield was unclear, an arc of radius about twice the size of the infield was to be digitized. For football, the field was to be digitized goal line to goal line. For kickball, the infield only was to be digitized. For volleyball, where markings are seldom present, an approximation of the playing area was sufficient.
- In the case of basketball, if no court were marked, a simple polygon around the hoop was to be digitized.
- Backyard amenities, specifically swimming pools and playground equipment, were to be ignored; the inclusion of all such amenities was deemed impractical. Beyond that, no distinction was made between public and private amenities, since the photography would not have informed us whether children could access the amenities or not. Private ownership, as might be determined from a plat overlay, would also not settle the question, as amenities owned by apartment complexes, homeowner groups, and private schools might well be accessible to the public.
- Tennis courts were to be digitized wherever found, to maintain consistency with the practice in creating the original file, and because these were relatively few and unambiguous.
- Equipment playgrounds with a mulched or sandy area surrounding the equipment were to be digitized to that area. In the absence of such an area, a convenient shape, a circle or rectangle, was to be placed around the equipment.
- Swimming pools were to be digitized to the water's area only; any previously added pools in which the deck area was also included were to be modified.
- Family centers were to be digitized to the building's footprint though these facilities were few and no additional ones were identified during the process.
- As soccer fields are not always permanently marked and goals moved frequently as needed, interpreters were instructed to digitize the entire area which, in their judgement, was set aside for playing soccer.
- In the case of fitness, the entire area in which fitness activities take place was to be digitized.
- Any areas where track or field events take place, including tracks, infields, or obviously designated external areas, were to be digitized.
- In cases where a particular area is clearly used for more than one of the chosen activities, overlapping polygons were to be created, according to the previous guidelines.

Limitations on the Final Product

- For amenity types that were to be quantified by count, interpreters were instructed not to correct minor inconsistencies in the original file, so long as general location and number were accurate. For instance, if a tennis court were digitized to its surrounding fence, rather than its playing surface, this was deemed sufficient. Thus, the inclusion or non-inclusion of marginal features within a buffer will be inconsistent by a few meters in the final data.
- While the quantification of playground equipment might be refined conceptually, none of these methods was practical for aerial photograph interpretation. The footprint area of a jungle gym, for instance, might best capture the opportunity represented by it, but trials showed this to be impractical, given the presence of shadows and inadequate resolution.
- Playgrounds were to be quantified by area, but the area of playgrounds is difficult to interpret consistently. Hard-surface playgrounds are often not demarcated clearly, as they coexist not only with basketball courts and kickball fields, but with parking lots. The presence of cars on a surface may be a temporary condition at the time the photograph was taken, which does not significantly alter the recreational opportunity in a longer time frame. Playground equipment is often located in a mulched area, but this mulched area is not always consistent from one year to the next. Such changes in area, therefore, were to be disregarded, so long as the playground equipment remained.
- Even in the presence of quality controls, the quality of interpretation must vary substantially with the individual, and nine individuals contributed to the final interpretation.

Each interpreter's completed work was selected from within his or her file; the areas covered by this work were assembled into a mosaic of recreational amenities. Border features were examined and redundancies removed, and any systematic errors discovered through the comparison were corrected; errors in naming were corrected, and any features marked as unknown were examined by a second interpreter, and either classified within one of the chosen types, or discarded. Finally, in those cases in 1996 and 1998 where no photographs were available, but where the preceding and following years matched in value, either both showing present or both showing absent, that value was substituted for the missing data. The four sets of Euclidean buffers used elsewhere in the larger study were intersected with the features in the recreational-amenities file.

We next performed a merger of all intersected vector features by original amenity, so that each resulting feature represented a single polygon resulting from the intersection of one buffer with one amenity. At this point, areas in square meters were calculated for all features. For those features that were to be quantified by area, this area was substituted as the value for each year in which the original amenity was present. Two copies of the file were created; in one file, every missing value was substituted with 0, and in the other file, every missing value was substituted with –9999999. A dissolve was then performed on the intersected features in each file, preserving buffer identification but grouping by amenity type, and summing the values for each year.

The resulting values in each feature were taken, in theory, as a measurement of recreational opportunity, as available to a child living at the center of each buffer, sorted by amenity type, with a value for each year. In the file in which 0 was substituted for missing data, the final measurement would represent a minimum. In the file in which –9999999 was substituted, every measurement in which *any* component value had been missing would be negative (as a single value of –9999999 would be greater than any possible value within the largest buffer used in the study), thereby allowing identification of the uncertain quantities. The file with the minimum values was used for the regressions in this study.

Appendix B

Land Use Variables

This appendix describes the data created for a set of social and physical environmental variables for use in the child obesity research. Data are provided for quarter-mile and half-mile buffers surrounding the children's residences. The variables for the quarter-mile buffers end in *25* and the variables for the half-mile buffers end in *5*.

Census Variables

Population density and the proportion of the population African American were created from the 2000 census block data from summary file 1. The education and income variables were created from the census block group data from summary file 3. Data from the surrounding counties were included, so there are no boundary issues near the border of Marion County.

Population Density—popden25 and popden5

This is the gross population density in persons per acre. Block population density was converted to a grid theme using fifty-foot grid cells (used in all of the data creation). The values are the means of the grid cell densities in the quarter-mile and half-mile buffers.

Proportion of the Population African American—prblk25 and prblk5

Block total population density and the population density African American were converted to the grid cells, the means for the buffers were calculated, and these were divided to obtain the proportion African American. Areas with zero population could not have a proportion calculated. This affected the variable *prblk25,* which has one missing value.

Proportion Graduated From High School—prhs25 and prhs5

This is the proportion of the population aged twenty-five and over who have graduated from high school. The densities of the population aged twenty-five and over and the numbers graduated from high school were converted to the grid cells, the means for the buffers were calculated, and these were divided to obtain the proportion graduated from high school. Areas with zero population aged twenty-five and over could not have a proportion calculated. This affected the variable *prhs25,* which has one missing value.

Proportion Graduate From College—prcoll25 and prcoll5

This is the proportion of the population aged twenty-five and over who have graduated from college. The densities of the population aged twenty-five and over and the numbers graduated from college were converted to the grid cells, the means for the buffers were calculated, and these were divided to obtain the proportion graduated from college. Areas with zero population aged twenty-five and over could not have a proportion calculated. This affected the variable *prcoll25,* which has one missing value.

Median Family Income—faminc25 and faminc5

This is an estimate of the median family income for the buffers. The block group median family income was converted to the grid cells, and the means for the buffers were calculated. Areas with no families and no median family income reported did not have values. This affected the variable *faminc25,* which has one missing value.

Road Network Variables

The planning literature suggests that greater density and interconnectedness of the road network (indicated by the density of intersections or nodes) should be associated with greater pedestrian use and physical activity. Data creation begins using the Etak road network for 2000. This was selected because it represented the network during the middle of the period for the obesity data, which seemed more reasonable than using the current road network. Limited-access highways and road segments associated with the interchanges were deleted from the network as these would not contribute to pedestrian activity. Data from the surrounding counties were included, so there are no boundary issues near the border of Marion County.

Road Network Density—rdlen25 and rdlen5

This is the sum of the length in miles of the road segments with their centroids within the buffers. The road segments were converted to a point layer with the line centroids, this was converted to the grid cells, and the results were summed for the buffers.

Number of Nodes—nodes25 and nodes5

The layer of road features was converted to a point layer of nodes. Dangling nodes and pseudonodes were deleted from this layer, leaving those nodes that represent intersections between roads. This layer was converted to the grid cells, and the count of the number of nodes in the buffers was obtained by summing those results.

Land Use Variables

The planning literature suggests that mixed land use, especially the presence of commercial land uses, should be associated with greater pedestrian use and physical activity. A parcel-based layer of land use in Marion County in 2002 from the Indianapolis Department of Metropolitan Development was used. Areas of streets and roads were not included in the delineation of land use. This data set covered only Marion County, so the proportions of land use near the boundaries reflect only land use within Marion County.

Proportion Land Use Commercial

This is the proportion of the classified areas of land use (not including areas of roads) that were classified in one of the commercial (retail and office) land use categories. The land use data were converted to the grid cells with values of 1 if commercial, 0 if other land use, and no data if road area. The means of these values were determined for the buffers to provide the proportion commercial.

Proportion Land Use Residential

This is the proportion of the classified areas of land use (not including areas of roads) that were classified in one of the residential categories. The land use data were converted to the grid cells with values of 1 if residential, 0 if other land use, and no data if road area. The means of these values were determined for the buffers to provide the proportion residential.

References

Anderson, M., and D. A. Matsa. 2007. Are restaurants really supersizing America? University of California Berkeley. Working Paper.

Booth, K. M., M. M. Pinkston, and W. S. Carlos Poston. 2005. Obesity and the built environment. *Journal of the American Dietetic Association* 105:S110–17.

Centers for Disease Control and Prevention. 2000. A SAS program for the CDC growth charts. Available at: www.cdc.gov/nccdphp/dnpa/growthcharts/resources/sas.htm.

Chou, S., M. Grossman, and H. Saffer. 2004. An economic analysis of adult obesity: Results from the behavioral risk factor surveillance system. *Journal of Health Economics* 23 (3): 565–87.

Chou, S.-Y., I. Rashad, and M. Grossman. 2008. Fast-food restaurant advertising on television and its influence on childhood obesity. *Journal of Law and Economics* 51 (4): 599–618.

Committee on Environmental Health. 2009. The built environment: Designing communities to promote physical activity in children. *Pediatrics* 123 (6): 1591–98.

Currie, J., S. DellaVigna, E. Moretti, and V. Pathania. 2009. The effect of fast food restaurants on obesity and weight gain. NBER Working Paper no. 14721. Cambridge, MA: National Bureau of Economic Research, February.

Estabrooks, P. A., and S. Shetterly. 2007. The prevalence and health care use of overweight children in an integrated health care system. *Archives of Pediatric Adolescent Medicine* 161:222–7.

Ewing, R., R. C. Brownson, and D. Berrigan. 2006. Relationship between urban sprawl and weight of United States youth. *American Journal of Preventative Medicine* 31:464–74.

Galvez, M. P., M. Pearl, and I. H. Yen. 2010. Childhood obesity and the built environment. *Current Opinion in Pediatrics* 22 (2): 202–7.

Gordon-Larsen, P., M. C. Nelson, P. Page, and B. M. Popkin. 2006. Inequality in the built environment underlies key health disparities in physical activity and obesity. *Pediatrics* 117 (2): 417–24.

Jeffery, R. W., A. Drewnowski, L. H. Epstein, A. J. Stunkard, G. T. Wilson, R. R. Wing, and D. R. Hill. Long-term maintenance of weight loss: Current status. *Health Psychology* 19 (1 suppl): 5–16.

King, A. C., R. W. Jeffery, F. Fridinger, L. Dusenbury, S. Provence, S. Hedlund, and K. Spangler. 1995. Environmental and policy approaches to cardiovascular disease prevention through physical activity: Issues and opportunities. *Health Education Quarterly* 22:499–511.

Los Angeles Times. 2008. Available at: www.latimes.com/news/local/politics/cal/la-me-fastfood30-2008jul30,0,7844906.

McDonald, N. C. 2007. Active transportation to school: Trends among U.S. schoolchildren, 1969–2001. *American Journal of Preventive Medicine* 32 (6): 509–16.

Margetts, B. 2004. WHO global strategy on diet, physical activity and health. Editorial. *Public Health and Nutrition* 7:361–3.

Merom, D., C. Tudor-Locke, A. Bauman, and C. Rissel. 2006. Active commuting to school among NSW primary school children: Implications for public health. *Health and Place* 12:678–87.

Papas, M. A., A. J. Alberg, R. Ewing, K. J. Helzlsover, T. L. Gary, and A. C. Klassen. 2007. The built environment and obesity. *Epidemiologic Reviews* 29 (1): 129–43.

Plantinga, A. J. and S. Bernell. 2007. The association between urban sprawl and obesity: Is it a two-way street? *Journal of Regional Science* 47 (5): 857–79.

Quigley, F. 2003. Indianapolis is missing the bus: IndyGo is worst-funded system in

Midwest. Available at: www.nuvo.net/archive/2003/12/24/indianapolis_is_missing_the_bus.html.
Sallis, J. F., A. Bauman, and M. Pratt. 1998. Environmental and policy interventions to promote physical activity. *American Journal of Preventive Medicine* 15:379–97.
Schlossberg, M., J. Greene, P. P. Phillips, B. Johnson, and B. Parker. 2006. School trips: Effects of urban form and distance on travel mode. *Journal of the American Planning Association* 72:337–46.
Singh, G. K., M. Siahpush, and M. D. Kogan. 2010. Neighborhood socioeconomic conditions, built environments, and childhood obesity. *Health Affairs* 29 (3): 508–15.
Tester, J., and R. Baker. 2009. Making the playing fields even: Evaluating the impact of an environmental intervention on park use and physical activity. *Preventative Medicine* 48 (4): 316–20.
Timperio, A., D. Crawford, A. Telford, and J. Salmon. 2004. Perceptions about local neighborhood and walking and cycling among children. *Preventive Medicine* 38:39–47.
Tobler, W. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46 (2): 234–40.
Wilfley, D. E., R. I. Stein, B. E. Saelens, D. S. Mockus, G. E. Matt, H. A. Hayden-Wade, R. R. Welch, K. B. Schechtman, P. A. Thompson, and L. H. Epstein. 2007. Efficacy of maintenance treatment approaches for childhood overweight: A randomized controlled trial. *Journal of the American Medical Association* 298 (14): 1661–73.