THE EFFECT OF NEIGHBORHOOD QUALITY ON CHILD OVERWEIGHT STATUS

by

JOSHUA HOLMGREN

B.S., Kansas State University, 2007

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF ARTS

Department of Economics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2009

Approved by:

Major Professor
Tracy M. Turner

Abstract

Obesity is a growing problem in the United States today. This study provides an econometric analysis of the relationship between child overweight status and neighborhood quality by using nationally representative data from the Panel Study of Income Dynamics (PSID) the 2002 PSID Child Development Survey, and Census data. In this study, the probability of a child overweight status is modeled as a function of neighborhood quality, child age, race and ethnicity, and parent obesity status, income, marital status, and education level. Next, the possible endogeneity between neighborhood quality and parent health is controlled for. Auxiliary regressions, modeling neighborhood quality and parent health on factors such as parent income, education, and marital status, are used to generate predicted values for neighborhood quality and parent health, which are then substituted into the child overweight equation to control for the aforementioned endogeneity. Census track and county level factors that might affect parent health or neighborhood quality are also controlled for. Based on a sample of 1917 children, this study finds evidence that neighborhood quality affects child overweight status.

Table of Contents

List of Tables

CHAPTER 1 - Introduction

The American Obesity Society states that:

"In the past 30 years, the occurrence of overweight in children has doubled and it is now estimated that one in five children in the US is overweight. Increases in the prevalence of overweight are also being seen in younger children, including preschoolers." (American Obesity Society, 2009)

The rising incidence of obesity in the U.S. population has brought about a great deal of research by economists. Researchers have examined the impact of income, education, and living environment on the likelihood of obesity. No papers, to date, have examined the role that neighborhood quality might play. Child health status may depend on the extent to which children can play outside, ride their bicycles, and get regular exercise, activities which living in an unsafe neighborhood may curtail. Overweight children may face lifelong health challenges, lessening their lifetime work productivity, income, and standard of living. The Center for Disease control (CDC) states that obesity increases the risk of such diseases as coronary heart disease, Type 2 diabetes, cancers (endometrial, breast, and colon), hypertension (high blood pressure), dyslipidemia (for example, high total cholesterol or high levels of triglycerides), stroke, liver, and Gallbladder disease, sleep apnea and respiratory problems, osteoarthritis (a degeneration of cartilage and its underlying bone within a joint), and gynecological problems (abnormal menses, infertility) (CDC, 2009b). These factors reveal the importance of understanding the determinants of obesity.

This study provides an econometric analysis of the relationship between child overweight status and neighborhood quality by using nationally representative data from the Panel Study of Income Dynamics (PSID) the 2002 PSID Child Development Survey, and Census data. In this study the probability of a child overweight status is modeled as a function of neighborhood quality, child age, race, and ethnicity, and parent obesity status, income, marital status, and education level. Next, the possible endogeneity between neighborhood quality and parent health is controlled for. Auxiliary regressions, modeling neighborhood quality and parent health on factors such as parent income, education, and marital status, are used to generate predicted values for neighborhood quality and parent health, which are then substituted into the child overweight equation to control for the aforementioned endogeneity. Census track and county level factors that might affect parent health or neighborhood quality are also controlled for. The remainder of this paper is organized as follows. Section 2 presents an in depth review of seven key papers examining overweight and obesity issues. Section 3 describes the data used in this analysis. Section 5 displays the econometric models the data is analyzed with. Section 6 provides the empirical results found from analyzing the data with the models in section 5. Finally, Section 7 concludes.

CHAPTER 2 - Literature Review

Why does low income typically mean poor health for children? There are many channels by which income status can influence child health outcomes. First, unobservable characteristics of low socioeconomic status (SES) parents can lead to both low income and poor child health. Second the affordability of health-linked goods and living conditions is crucial to ensure good child health. A lack of these goods and living conditions may inhibit a child's ability to earn income in the future, thus causing endogeneity between child health and future income. Some unobservable low SES parent traits that may lead to low income include a tendency toward a sedentary lifestyle, an inability to delay gratification, and a preference for unhealthy foods. Trouble delaying gratification may lead to poor education, financial, and health choices that beget low SES and increased risk of poor health outcomes for parents and their children.

Some examples of healthy goods and living conditions that low SES parents may not be able to afford are: healthcare, high-quality food, access to sports clubs, leisure time to spend with their children, and green space. These topics have been examined by researchers. Government food programs, which low SES families may need to take part in, may only cover low quality foods, which lead to such negative heath outcomes as obesity (Hofferth and Curtin, 2005). The amount of leisure time a mother spends with her child may also lead to positive health outcomes (Fertig et al., 2006.) Access to green space is necessary for people to have opportunities to go outside and exercise on a regular basis, which has been proven to lessen negative health outcomes (Liu et al., 2007). Low

SES parents who are unable to afford these things will cause their children to experience more health shocks, or have a slower recovery time from sickness and injury, than high SES children. This consequence can limit the income earning potential of the low SES children causing feedback between low income and poor health (Condliffe and Link, 2008).

Importantly low SES parents may also tend to live in lower quality neighborhoods inhibiting the parents' and children's ability to safely play outside and exercise on a regular basis. No studies to date have examined the impact of neighborhood quality on child health as measure through weight status, which is the topic of this study. The remainder of this section provides a detailed review of several important papers examining the determinants of individuals' weight status.

Hofferth and Curtian (2005)

Hofferth and Curtin (2005) set out to answer two questions: do children from low income families have a greater likelihood of being overweight than children in families of other income levels and does taking part government sponsored food programs increase the rate of obesity in children? The government food programs Hofferth and Curtin are interested in are the Food Stamp Program (FSP), the National School Lunch Program (NSLP), and the School Breakfast Program (SBP). The data that Hofferth and Curtin use is from the Child Development Supplement CDS of the PSID as well as data from the National Health and Nutrition Examination Surveys (NHANES). The NHANES data is longitudinal data. Hofferth and Curtin use weights to make the NHANES dataset nationally representative. While Hofferth and Curtin analyze this data, they control for child age, sex, race, ethnicity, birth weight, family size, parent employment status, number of parents, and parent education level.

In order to answer their first question, they look at differences in the weighted mean obesity statuses of children among income levels. Hofferth and Curtin hypothesized that children from families in the near-poor to working-class range would be more likely to be obese since they have more income to spend on food than poor families yet still have access to government sponsored food programs. However, they find that there is a nonlinear relationship between child obesity and income; children from poor families tend to be less likely to be obese than those from near-poor, working-class, and moderate-income families, and so do children from high-income families. The explanation Hofferth and Curtin give for this is as follows. At low levels of real income, an increase in real income causes the quantity of food consumed to increase hence an increase in BMI. Though at some certain higher threshold level of real income, an increase in real income leads to an increase in quality of food consumed and as a result BMI decreases. However, there may be something correlated with an increase in real income, such as education, that is driving down BMI from moderate-income families to high-income families.

Hofferth and Curtin answer the second question by logistically regressing child overweight status on child participation in government food programs and control variables. They also look at differences in the mean proportion of obese children and mean BMI between income levels and partaking in government food programs. For the NSLP, data on participation in the NSLP is given in the CDS, they use public school attendance as an instrumentation variable to estimate the probability that a child eats their meals at school since it is correlated with whether or not that student eats at school but is not linked to BMI.

The authors require this instrumentation variable because there is a possible endogeneity issue between overweight status and the probability of a child eating a school lunch. It may be the case that eating a school lunch causes a child to be more likely to be overweight, but it may also be the case that overweight children are more likely to choose to eat a school lunch. Therefore, there is a correlation with the independent variable, the probability of a child eating a school lunch, and the error term. Thus an instrument variable, or other means of combating this endogeneity, is required.

Initially they hypothesize that participation in food programs will increase children's risk of obesity. They do not use the instrumentation variable to find the impact on participation in the SBP on BMI because there are not enough children in the CDS who participate in strictly the SBP and not the NSLP. Without the instrumentation variable in the regression, participating in the NSLP significantly increases the chances that a child is obese. However including the variable causes this increase in risk of obesity to disappear. A likely explanation for this could be that children who choose to eat a school lunch are already prone toward obesity beforehand. They also find that there is no relationship between a child being a part of the food stamp program and that child being obese, but the authors warn that since this result contradicts other research that more investigation on this program is required.

Fertig, Glomm, and Tchernis (2006)

Fertig *et al.* (2006) analyze the effect of mothers' employment on their children's BMI and weight status. They assume that the "production function" for a child's health is dependant on the number of hours that a mother spends with the child, and that there is a direct relationship with diminishing returns between children's health and children's time spent with the mother.

The authors believe that the production function for a child with a relatively more educated mother is above the production function, at every point except zero hours, for a child with a relatively less educated mother. The superior information education provides allows the more educated mothers to have a better child health outcomes for a given amount of time spent with their children. However, Fertig *et al.* want to know whether the slope for the child's heath production function with a more educated mother is greater than one with a less educated mother for every given amount of time spent with the child. Another difference in average child heath outcomes between more and less educated mothers may be attributed to more educated mothers choosing to work more hours than less educated mothers.

To accomplish their goals, Fertig *et al.* use the PSID CDS data. The descriptive statistics of the CDS data show that BMI of children from what the authors deem as highly educated mothers, those who have greater than 12 years of schooling, is .78 less than those from other mothers, those with less than 12 years of schooling. Highly educated mothers work four more hours per week on average than less educated mothers.

To answer their questions empirically, the authors try to calculate $\frac{dB}{dH}$ by calculating $\frac{dB}{dM} \cdot \frac{dM}{dH} + \frac{dT}{dH} \cdot \frac{dB}{dT}$, where B is equal to the BMI of the child, H is equal to the mother's work hours, M is equal to the number of meals the child eats per day, and T is equal to the number of hours the child watches TV per day. However, this method would have endogeneity issues since in this model number of hours a child watches TV per day and number of meals a child eats per day are dependant upon the same variables. To control for this the authors should find the predicted number of meals a child eats per day, $\hat{M}$, and the predicted number of hours a child watches TV per day, $\hat{T}$, from mothers work hours and other controls. Then they could estimate $\frac{dB}{dH}$ by multiplying $\frac{dB}{d\hat{T}}$ by $\frac{d\hat{T}}{dH}$ or multiplying $\frac{dB}{d\hat{M}}$ by $\frac{d\hat{M}}{dH}$.

The controls the authors use in their regression are: age of child, race, sex, birth order, birth weight, number of children in the household, fraction of meals eaten at restaurants, age of the mother at the child's birth, education of the mother, parental obesity status, parents marital status, parents labor income during the child's life, hours per week the father works, whether or not the mother has received food stamps in the past year, region, and whether or not the child lives in an urban area.

From the regressions, the authors find that an increase in the number meals per day a child consumes due to mother's work hours has a more significant impact on a child's BMI than the number of hours a child watches TV per day. Though the number of hours a child watches TV per day has the most positive elasticity out of all the variables in the regression, it is only significant for the children of highly educated mothers, where as the number of meals a child eats per day is statistically significant for children of mothers from both education backgrounds. The total elasticity of BMI regressed on the of mother's work hours is one percent for both highly and less educated mothers. They do find that the number of meals a child eats at a restaurant is not significantly correlated with an increase in BMI. The authors speculate that their inability to discern healthy restaurant choices from non-healthy restaurant choices is likely to be the culprit for this result.

Liu, Wilson, Qi, and Ying (2007)

Liu *et al.* (2007) analyze the link between environmental factors and the risk of obesity. The environmental factors they focus on are the distance between children's homes and different types of food outlets and the amount of vegetation around children's homes.

Liu *et al.* cite previous research that has shown that greenery in landscapes has been associated with many positive health outcomes. On average patients recuperating from surgery in hospital rooms with windows overlooking scenery with lots of vegetation tend to recover quicker and need less pain medication than those patients without. Children cognitively function better in green environments; research has shown that the symptoms of Attention Deficit Disorder in children are less pronounced during activities in green areas outside compared to similar activities elsewhere.

The effect of the distance between the places children live and fast food and convenience establishments has on children's risk of obesity may be more intuitive than the effect of vegetation around children's homes. Research has confirmed this relationship for adults, however there is little research validating this intuition for children. The distance between places adults live and supermarkets has been shown to be negatively correlated with fruit and vegetable consumption.

The authors note that population density will definitely be a factor with these effects. People shopping in densely populated urban environments pay significantly more for their groceries than people shopping in more spread out suburban environments (3 percent to 37 percent more). This fact may cause people living in these urban environments to settle for eating at the closer fast food and convenience stores where most of the food available is sugary and fatty and only available in large portions. With these results and intuition in mind, the authors hypothesize that children with homes closer to supermarkets, further away from fast food stores, and in greener areas will be less likely to be obese.

In order to test their hypotheses they study cross-sectional data from pediatric clinical data coupled with environmental data from geocoding subject's home addresses. They gathered the pediatric clinical data from the Indiana University Medical Group network of seven urban primary care clinics located in Marion County Indiana. In order to quantify the amount of vegetation around children's houses, the authors use images from Landsat Enhanced Thematic Mapper Plus satellite imagery. They adapt the pixel values to numeric measurements with the Normalized Difference Vegetation Index, which converts on the principle that plants absorb and reflect.   The distance from children's homes to food retail places was figured using network distance around street centerlines.

The authors controlled for children's age, sex, race, population density, and median neighborhood family income. The authors use data from the 2000 U.S. Census to control for population density around children's homes. The analysis of the environmental data is done with Pearson χ^2 test and cumulative logit models. The frequency of categorical and ordinal variables between population densities are analyzed using Pearson χ^2 tests. The overweight index variable is regressed on population density and other environmental factors in cumulative logit models. The mean distance from children's homes to supermarkets is larger for high population density areas, whereas the mean distance to market, convenience, and fast food stores is greater for low population density areas. With all the controls in the model, the amount of vegetation around children's homes was negatively and significantly correlated with risk of overweight status of children in high population density areas. This was the only environmental factor that had a significant impact on risk of overweight status in high population density areas. In low population density areas, the only environmental factor that had an impact on risk of over weight status in children was distance to the nearest supermarket from children's homes, which is positively correlated with risk of obesity in children. Other environmental factors that were controlled for but were not significant in either case include: distance to convenience stores, grocery stores, and fast food restaurants.

Plantinga and Bernell (2007)

Plantinga and Bernell (2007) analyze the link and causality between obesity and urban land development. Urban sprawl is intuitively connected to obesity for many reasons. Low population density, single use development, and poor street connectivity increase commute time along with making biking and walking impractical and unsafe.

People may also substitute their physical activities for the extra travel time they must commit to while living in suburbs. Thus people living in sprawling areas exercise less.

The data Plantinga and Bernell use in their analysis is from the National Longitudinal Survey of Youth 1979 (NLSY79). Other characteristics the authors use for controls in an analysis of the determinants of BMI and housing density choice include: sex, race, educational status, smoking status, marital status, income, age, region of country, number of children, urban sprawl around residence, and home county. The measure of sprawl around residences is derived from a county sprawl index, though this may not be the best measure of neighborhood sprawl, created by Ewing and McCann (2003).

The first equation that the authors estimate in their study is used to confirm that the relation between sprawl and obesity Ewing *et al*. found holds for the NLSY79 data. For this equation the authors use a nationally weighted sample of 4,700 observations. The authors regress BMI on urban sprawl using OLS, an estimate of the covariance matrix that controls for heteroskedasticity, and the controls previously discussed above; the authors suspect heteroskedasticity to be present because they suspect that individuals within the same county will have correlated error terms. Ewing *et al*.'s analysis has some differences from Plantinga and Bern ell's; Ewing *et al*. includes one more education dummy variable for "some college", dummy variables to divide people into age groups, a variable for fruit and vegetable consumption, and people of ages 18 and older included in the observations. The NLSY79 data does not have any information on people's diets therefore Plantinga and Bernell does not control for this. Most importantly, Ewing *et al*.'s model controls for county-specific effects whereas Plantinga and Bernell only have enough information in their data to control for region within the U.S. Despite the differences in the models created by Plantinga and Bernell and Ewing, both models show a statistically significant positive correlation between sprawl and BMI. The coefficient that Ewing *et al*. derives for sprawl is within a 95 percent confidence interval calculated using the coefficient that Plantinga and Bernell finds with their model.

The next model that Plantinga and Bernell estimate is used to test whether or not BMI effects people choice of residence based on the level of urban sprawl of the neighborhood. To accomplish this, the authors regress "dense", a binary variable defined as whether or not an individual moved to a county with a level of sprawl above or below an arbitrary amount from 1998 to 2000, on their BMI as of 1998 and control variables. Since the observations used in this model are spread out over the course of two years the authors are unable to weight the data. The authors use 381 observations from their data because they were limited by the number of people that moved over this time period. The only variables that significantly effect "dense" are income, pre-move BMI, and having more than two children. All of these variables are positively correlated with choosing a sprawling county. Plantinga and Bernell also find that the correlation between the dependant variable and pre-move BMI is still the same when the definition of a "dense" county is moderately changed.

Finally the authors estimate a model that shows how sprawl of a county affects BMI, while controlling for the endogeneity between these two factors. The authors use 262 observations from people who moved once between the years 1996 and 1998 and not again at least until after 2000. "BMI change" is the dependant variable that the authors use in their regression, which is defined as the difference of people's BMI in 2000 and their BMI upon their move. “Sprawl change” and other controls are the independent variables used in the regression. “Sprawl change” is, of course, the difference in county sprawl index of people’s new residence to their old one. Differences in control variables from the previous models are as follows: income and dummy variables for people’s change in education levels are omitted, marital status, smoking status, and number of children over the course of 1996 to 1998 are included. The sprawl index is derived by information from 2000, therefore, it would be possible that the sprawl index measurement of people’s neighborhoods in 1996 would be inaccurate. However, it would be unlikely that the sprawl of a county would change significantly over the course of a few years. Again, the data is not weighted due to observations taken in more than one year and the robust covariance matrix is used to fix heteroskedasticity in the model (White, 1980). From the regression, the coefficient of “change in sprawl” is negative and significant implying that moving to a less sprawling area reduces BMI.

Burkhauser and Cawley (2007)

Burkhauser and Cawley (2007) investigate whether or not body mass index (BMI) is the best measure for determining obesity and examine alternative measures of overweight status in a study of employment status. BMI is computed as weight in kilograms divided by height in meters squared. The authors feel compelled to do this study because BMI does not take into account a person’s muscle or bone mass. Thus, judging people’s obesity by their BMI will inevitably result with some false positives because people with a large amount of muscle mass or bone mass will have a large BMI but should not be considered obese. The investigation consists of analyzing alternative measures of obesity, testing if gender and racial differences in obesity are significantly affected by what measure of obesity is used, and testing if the correlation between employment status and fatness is altered when different measures of fatness are used.

The data the authors use in their study is the NHANES III data because it has detailed and numerous measures of fatness for many of the people observed in it. This data was collected from 1988 to 1994. In the authors' analyzation of the data, they use the sample weights provided in the NHANES III. The alternative measures of obesity the authors analyze are total fat, percent body fat, waist circumference, and waist-to-hip ratio, which are reported in the NHANES III. Increased risk of diseases associated with obesity, such as type II diabetes and cardiovascular disease, is caused by a combination of a person's quantity of fat, the amount of nonfat body mass, and where the fat is located. Thus, each of these alternative measures of fatness has their strengths and weaknesses in indicating if a person has an increased risk of such a disease.

In order to test the magnitude of the correlation between true obesity status and BMI, the authors first analyze obesity rate differences between males and females with obesity defined by BMI and percent body fat. Using the percent body fate definition of obesity, more people are obese than using the BMI definition. However this does not necessarily indicate which one is a better measure. To get around this "threshold" difference in obesity definitions they redefine the cutoffs for obesity under the percent body fat definition such that the population has the same obesity rates under both definitions. Assuming that the percent body fat definition of obesity, which, the authors assert, has a more substantial theoretical background than BMI, is correct, the BMI definition yields a false positive rate of 22.25 percent for women and 42.64 percent for men. It also yields a false negative rate of 6.44 percent for women and 9.93 percent for men. The authors argue that it is intuitive that the BMI definition gives more false positives for men than women, since men tend to have more muscle mass than women.

The BMI definition also produces false positives that vary by race. Like before the authors once again alter the percent body fat cutoffs to account for threshold differences and assume that the percent body fat definition is true. The BMI definition gives a false positive rate of 4.43 percent for white females, 10.56 percent for African American females, 7.77 percent for white males, and 10.40 percent for African American males. The differences in false positives across races, for people of the same gender, are statistically significant. This difference in false positive rates is due to the difference in fat free body mass between races. The average African American female has 3.56 kg more fat free body mass than the average white female, and the average African American male has 1.33 kg more fat free body mass than the average white male.

Burkhauser and Cawley also investigate whether or not the correlation between fatness and employment status is sensitive to the type of fatness measure used where employment status indicates whether someone is employed or unemployed. To accomplish this they estimate twelve equations, each of which have self reported employment status as the dependant variable and controls for age, age squared, education by dummy variables for completed high school and greater than high school education, and marital status. Each equation includes a measure of fatness: obesity defined by BMI, obesity defined by percent body fat, BMI, and total body fat with fat free mass simultaneously. For every fatness measure, the authors estimate an equation by race, either white or African American, and gender combinations. The authors may have endogeneity issues in their models with this approach since one would expect feedback between their model, employment status as a function of a measure of fatness and other controls, and a model consisting of a measure of fatness as a function of employment status and other controls.

Unsurprisingly, the marginal effects of obesity defined by BMI and obesity defined by percent body fat on employment for women are closer to being the same than those for men. For white males the marginal effects of obesity defined by BMI is not even statistically significant at the ninety percent confidence level. The difference between these two marginal effects is greater for African Americans. The marginal effect of BMI on employment, a negative correlation, was only significant for white females. There was a significant negative correlation between total body fat and employment for white people; there was no significant correlation between total body fat and employment status for African Americans. Finally, the authors did not find a significant correlation between fat free mass and employment status.

Ewing, Schmid, Killingsworth, Zlot, and Raudenbush (2003)

Ewing *et al.* (2003) analyze the correlation between urban sprawl, adult health outcomes, and adult health related behaviors. The Data they use for their analysis is from years 1998 through 2000 of the Behavioral Risk Factor Surveillance System (BRFSS,). The authors use maximum likelihood estimation so that the more observations a county contributes, the more weight is put on the observations from that county, when calculating parameters for the econometric models. Ewing et al. pools observations from each year to increase statistical power of their analysis.

Ewing et al. create three leisure-time physical-activity variables: one for whether or not a person gets any leisure-time physical activity during the previous month, one for whether or not a person gets the recommended amount of physical activity during the previous month, and one for amount of minutes the person walks during the previous month. The weight related variables the authors use are BMI. Other negative health outcome variables include whether the respondent has coronary heart disease, diabetes, or hypertension. The authors control for covariates of health outcomes and health related behaviors, such as age, gender, race/ethnicity, education, smoking status, and fruit and vegetable consumption, by using Hierarchical linear and nonlinear modeling (HLM) methods. Ewing et al. use the *metropolitan sprawl* (density) *index*, made available by the group Smart Growth America (SGA), to measure urban sprawl at the metropolitan level where the index takes on higher values as density increases. The authors also derive a simpler measure of urban sprawl (density), called the *county sprawl index*, to measure density at the county level, which is the smallest geographic unit reported in the BRFSS data.

The authors use "HLM 5" a statistical software program, which adjusts for the heteroskedasticity apparently commonly present in BRFSS data, to estimate relationships between urban sprawl and leisure time physical activity levels, BMI, obesity, hypertension, diabetes, and CHD. The HLM software first models observations' health status or behavior within each location as a function of observations' characteristics plus an error term. Next, the location-specific intercept and coefficients are treated as outcomes and modeled by their corresponding location characteristics plus an error term. The binary health outcome and health related behavior variables are modeled linearly, while the continuous variables were modeled nonlinearly.

From their analysis Ewing et al. found that males, younger people, white non-Hispanic people, and more educated people are more likely to exercise during leisure time than other people. The groups more likely to meet their recommended requirement of exercise are the same. However, people of ages 65 and older tend to be more likely to get their recommended amount of exercise than young adults due to the amount of leisure walking they do. Females, older people (until they reach age 75), and more educated people tend to spend more time walking per month than other people. With controls, Ewing et al. found that getting any exercise during the previous month was not related to the county density, however getting the recommended amount of exercise per week and the number of minutes a person walks per month is positively correlated with the county density. The authors find that out of all the health outcomes and health related behaviors only minutes walked has a significant (and positive) correlation with metropolitan density.

Ewing et al. also finds that males, African and Hispanic Americans, older people (until they reach age 45), less educated people, nonsmokers, and people who do not eat three or more servings of fruit and vegetables per day tend to have a higher BMI than other groups of people. The authors find that both BMI and obesity status variables are negatively correlated with the county density. Ewing et al. analyze the indirect effects of urban sprawl on BMI and obesity by treating minutes walked as an independent variable and predicting its marginal effects on obesity and then on BMI. Unsurprisingly the authors found a significant negative correlation in both cases. Ewing et al. find that older people are at higher risk of having hypertension, CHD and diabetes than others, males have a higher risk of diabetes and CHD than females, and less educated people are more likely to have diabetes and hypertension. The only health outcome that has a statistically significant correlation (a negative one) with the county density was hypertension.

Condliffe and Link (2008)

Condliffe and Link (2008) investigate the relationship between socioeconomic status (SES) on child health. In particular, they mimic the research done by Case, Lubotsky, and Paxson (2002) and Currie and Stabile (2003) using CDS data from the PSID and data from the Medical Expenditure Panel Survey (MEPS). Case et al. finds that there is a positive relationship between SES and child health and that, for children of low SES, this relationship becomes more pronounced as the children age.

Currie and Stabile as discussed in Condliffe and Link propose three possibilities: low SES Children have the same number of health shocks as high SES children but due to lack of income, lack of information, or delayed treatment of conditions are less likely to recover from the shocks, low SES children have more health shocks than high SES children due factors related to dangerous living conditions or dangerous lifestyles, or low SES children both experience more health shocks and are less likely to recover from them than high SES children. Currie and Stabile find that the increasing, positive relationship between SES and children's heath is due to low SES children having more shocks. Using panel data of Canadian children from years 1994 to 1998, Currie and Stable find that children with chronic health conditions in 1994 have the same likelihood of being in poor health whether they are of high or low SES. However, Condliffe and Link find something else: that each possibility is valid.

Condliffe and Link concentrate on children in the PSID that have chronic conditions such as, epilepsy, asthma, diabetes, retardation, vision problems etc., in 1997 and those children's health status in 2002. The MEPS data the author uses has information on children's family's SES and on children's health status. The authors first replicate the study by Case et al. by estimating in a cross-sectional context the effects of the natural log of family income (the average of the 1997 and 2001 family incomes of a given child), the mother's education (a high school education dummy variable), the child's age, and other socioeconomic factors on a variable (a categorical variable for a probit model with a value of one signifying excellent health and five for poor health and a dummy variable for a linear probability model) for the child's poor health status. In the regressions, like The authors find significant negative correlations between family income and children's poor health status for both the MEPS and PSID data and with and without the mother's education variable, although the effects of income fall slightly without this variable. These models also confirm that the disparity in health statuses between high and low SES children gets more negative as children age because the coefficients of the probit and linear probability models get more negative the older the age group the model is estimated with.

Next the authors estimate a few models to test whether or not low and high SES children recover from health shocks in the same manner with the PSID data. These models measure the effects of a chronic condition present in 1997, the natural log of family income interacted with a chronic condition present in 1997, and other socioeconomic factors on a dummy variable that is equal to one if a child's health status is good, fair, or poor in 2002. When chronic conditions in 1997 and the natural log of family income are included in the model, both coefficients have their expected signs (positive and negative, respectively) and are statistically significant. When an interaction term between these two variables is added, its coefficient is negative and statistically significant, and the other coefficients keep their signs and significance.

Next the authors model poor health on the presence of asthma, they choose this condition in particular because of its severity and its high prevalence in children, in 1997 and the natural log of family income. Both of the coefficients in the model are significant. From this model, the authors see that the presence of asthma in 1997 has a positive correlation with poor health in 2002, while family income has a negative correlation with poor health in 2002 as expected. Finally the authors find the effects of asthma in 1997, the natural log of family income, and an interaction term between asthma in 1997 with family income on the poor health in 2002, however only family income's coefficient is significant.

To find the short run effects of the full model the authors replace the variable for presence of a chronic condition in 1997 with the presence of a chronic condition in 2002, being hospitalized in the last year, and the presence of asthma in 2002 along with each variables interaction with family income in three separate models and use the MEPS data instead of the PSID data. In each of these models the coefficient on these variables are positive and statistically significant, and the coefficient of each variable interacted with family income is negative and significant, thus low SES children are more likely to be in poor health in the short run if they have a chronic condition in the current year.

The authors also test whether or not low SES children gain more chronic conditions as they age compared to high SES children by estimating the effects of the natural log of family income, the natural log of family income interacted with children's age, previous chronic conditions, and other socioeconomic factors on the number of new chronic conditions children have gained since the previous period. To accomplish this they create several models: some bivariate and some multivariate. In all of the models the coefficients for family income and age interacted with family income are positive and negative respectively, however these coefficients are only significant using the MEPS data. When controls for asthma are present in the previous period or other chronic conditions present in the previous period entered the model their coefficients are positive and significant in both data sets. Putting these controls in the model did decrease the age-income interaction.

Next the authors estimate these models using the PSID data while restricting age groups to 10-11, 12-14, and 15-18 years old. The coefficient estimates for income on the presence of a new chronic condition since the initial time period are negative whether the model is bivariate or multivariate and for all age groups. These coefficient estimates are only significant when the models are estimated with age groups 15-18 (in both the bivariate and multivariate models) and 10-11 (in the multivariate model only.)

CHAPTER 3 - Data Section

The data used in this study is from the Panel Study of Income Dynamics (PSID), the PSID Child Development Supplement (CDS), and the Summary File 3 (SF 3) from the 2000 U.S. Census. A list of the variables used from each data source can be found in Table 2.[1] The CDS is a special supplement to the PSID devoted to accumulating data on the development of children. The CDS data on the same children at two points in time: 1997 and 2002. The 2002 CDS has data for 2,019 households containing 2,907 children The CDS includes information on child health, emotional well-being, intellect, academics, relationships with the child's family and friends, and other demographical information. In particular from the CDS, this analysis uses child weight status, age, and neighborhood quality data. Following Fertig et al. 2006, child age is included as a dependant variable in the child OW equation. As in Fertig et al.'s study child age is expected to be negatively correlated with OW status (Fertig, p11).

Measuring Child Overweight Status

This study follows the criteria for measuring child overweight status that is recommended by the Center for Disease Control (CDC) (CDC, 2009c). The CDC commissioned an expert committee to create a growth chart to classify children's weight

[1] PSID confidential location indicators were used to merge the county-level census data to the PSID households. The location indicators are from the PSID GEOCODE data files and obtained from the University of Michigan under special contract. Individuals wishing to use the GEOCODE data may contact the University of Michigan for permission.

status. The most recent growth chart was created in 2000. The CDC classify a child as underweight if that child has a BMI less than the 5^{th} percentile for their sex and age, as of normal weight if the child has a BMI greater than the 5^{th} percentile but less than the 85^{th} for their sex and age, as at risk of overweight if the child has a BMI greater than the 85^{th} percentile but less than the 95^{th} for their sex and age, or as overweight if the child has a BMI greater than the 95^{th} percentile for their sex and age. The PSID CDS has information about child weight status in both 2002 and 1997; in both 1997 and 2002 weight status in the PSID CDS is defined by the 2000 CDC growth charts. Using this information, a variable is created indicating overweight status for each child, using the 85^{th} percentile cut off. In this analysis a child is considered overweight (OW) if his or her BMI exceeds the 85^{th} percentile for his or her sex and age and not OW if his or her BMI is less than the 85^{th} percentile.

In statistics specificity and sensitivity are qualities of binary specification tests; specificity measures the fraction of true negatives correctly identified by the test, while sensitivity measures the fraction of the true positives correctly identified by the test. Regarding the CDC cutoff points, they are chosen because they maximize the "specificity" of testing whether or not a child is OW, meaning they minimize the proportion of children incorrectly considered OW. Alternatively one could maximize "sensitivity" of this test which would minimize the proportion of children incorrectly considered not OW (Himes and Dietz, p309). The cutoff of the 85^{th} percentile has a sensitivity of 29 percent and 23 percent and a specificity of 99 percent and 100 percent for boys and girls, respectively, when this cutoff is used to determine if a child has an amount of body fat in excess of the 90th percentile of body fat amounts for that child's weight, age, and sex.

The 2000 CDC growth chart has been updated from the 1977 CDC growth chart with a nationally representative reference population from years 1963-1994. This data was collected by the NHANES. The committee deciding these cutoffs was not aware of any other data suggesting different cutoff points, such as other BMI percentile cutoff points or levels of body fat or BMI, more useful in predicting children's risk of negative health outcomes from excess body fat. Only slight differences are observed when comparing the prevalence of obesity in children in the NHANES III data by using the old and new chart.

Neighborhood Quality Measures

Two measures of neighborhood quality are created using the survey data reported in the CDS. The first measure indicates whether the family lives in a neighborhood that is good for raising children. Specifically the CDS asks the parent, "How would you rate your neighborhood as a place to raise children?" Response options include: "excellent", "very good", "good", "fair", and "poor". The variable called "bad neighborhood (parent rating) " is equal to one if the parent responded with poor and is equal to zero if the response is excellent, very good, good, or fair. The second measure of neighborhood quality is formed based on the CDS question "How likely is it that a neighbor would do something if someone was trying to sell drugs to your children in plain sight?" The response options are very likely, likely, unlikely, and very unlikely, and similarly The variable "bad neighborhood (drugs)" is defined to be one if the parent responds with unlikely or very unlikely and is equal to zero if the parent responded with likely or very likely.

Parent Obesity Status

The PSID core survey is a longitudinal, nationally representative survey of roughly 7,000 families, which provides extensive household and person level data. The PSID, started in 1968, has annual observations until 1997, when the survey became biannual. The CDC also defines the cut off point for adult overweight status as a BMI of greater than or equal to 25 and less than 30 and obesity status as BMI greater than equal to 30 (CDC, 2009a). For this analysis a variable called "parent obese" is created which is equal to one if at least one of the parents in the household is obese, and zero if neither the parent is obese. Parent obesity is analyzed instead of instead of overweight status because this will control for the influence of parent genetic factors on child overweight status. Controlling for parent obesity is more likely to measure unobservable characteristics passed on from the parent to the child such as a genetic predisposition toward overweight status or preferences toward fatty foods and sedentary activities, since it is more severe than being overweight and more likely to be due to factors other than environment.

Census Data

The SF 3 contains data on population and housing characteristics for the entire U.S. The smallest geographical unit that the SF 3 contains data on is the census block, which is defined as "an area bounded on all sides by visible and/or non-visible features shown on a map prepared by the Census Bureau" by the Technical Documentation file of the SF 3. The data of the SF 3 is used to control for the availability of safe and affordable neighborhoods in a county, urban sprawl, the percentage of the people in a county that bike to work, the percentage of the people in the county who walk to work, and region.

Measures of the supply of safe and affordable neighborhoods are necessary to predict neighborhood quality and identify the system of equations described in the econometric model section of this study. It is expected that these measures of safe and affordable neighborhoods are negatively correlated with neighborhood quality. Urban sprawl is included to predict parent obesity status, following the research of Plantinga and Bernell (Plantinga and Bernell, p858). Urban sprawl is, likely, positively correlated with parent obesity. The variables created from the percentage of people in a county that bike or walk to work is included in this study to predict the obesity status of parents and to identify the system of equations described in the econometric model section. Regional controls in the parent health and neighborhood quality equations, as described in the economic section, are included, following the approach of Fertig *et al.* 2006, since there may be regional differences in preferences for unhealthy foods or sedentary activities that would affect the likelihood of parent obesity status and since there may be regional differences in the supplies of safe and affordable neighborhoods (Fertig, p13).

Two measures of the supply of safe housing within the state where the household lives are formed by creating a high poverty rate variable and a high unemployment rate variable. The high poverty rate (and high unemployment rate) variable is equal to the percentage of counties in the state where the household lives that have a poverty rate (or unemployment rate) above the mean poverty rate (or unemployment rate) of all the U.S. counties. A state level measure of the availability of affordable housing within the county that the household lives is formed in a similar fashion to the measures of availability of safe housing: by taking the percentage of counties in the state where the household lives that has an affordability ratio above the mean affordability ratio for all U.S. counties. An affordability ratio is formed by taking the median house value of the county and dividing it by the median household income for the county. Census tract level measures of neighborhood safety and affordability are formed that are made in an exactly analogous fashion to the county level measures. When models with census tract measures are compared to those with county level controls, the models with county level controls had a higher log likelihood value, and therefore these measures are not controlled for at the census tract level.

To control for urban sprawl a population density variable is created from the Census 2000 data by dividing the total population within a census tract by the total land area in that census tract. Other variables that are created from the Census 2000 data that may influence parent obesity status are the percent of people within a county that bike to work and the percent of people within a county that walk to work. The bike and walk to work variables only include people of ages 16 or older who commute to their place of work.

CHAPTER 4 - Econometric Model

This section describes the approach this study takes to examine the impact of neighborhood quality on child overweight status. The probability that a child is overweight is modeled as:

$$prob(Y_i) = \Phi(C_i'\alpha_1 + Z_i'\alpha_2 + \alpha_3 N_i + \alpha_4 H_i + \epsilon_{1i}), \qquad (1)$$

where, *prob*(Y_i) is equal to one if child *i* is overweight is equal to zero if child *i* is not overweight, $\Phi(\cdot)$ is the standard normal cumulative distribution, C_i is a vector of exogenous child characteristics, Z_i is a vector of exogenous parent characteristics, N_i is a measure of neighborhood quality, H_i is a measure of parent health status, and ϵ_{1i} is a random error term.

The variables that represent C_i are child age, child race, and child ethnicity. The age variable is equal to the age of the child at the time of the 2000 CDS interview. The variable, child black, indicates if the child is black and not Hispanic. This study includes the variables black and Hispanic following Fertig *et al.* 2006 (Fertig, p11). It is expected that these two variables will be positively correlated with childhood obesity. The variables that represent Z_i are log (income), high school diploma, some college, and college degree or higher. Income enters this model in this study following the approach of Fertig *et al.* 2006, and as in Fertig *et al.*'s 2006 study it is expected to negatively correlated with OW status (Fertig, p4). Three measures of family income are formed from the PSID data: total family income for the year of 2002 in tens of thousands of dollars, total family income for the year 2002 in tens of thousands of dollars squared, and the natural log of total family income for the year 2002, log (income). In this model income is chosen to represent with log (income) by comparing the log likelihood of models with log (income), income, and income with income squared as income controls. The excluded education indicator variable is no high school diploma. Each education variable indicates what the highest education level is amongst the parents in the household. Educational status is divided into levels in this model that mimic the levels Hofferth and Curtin 2005 choose for their models (Hofferth and Curtin, p719). It is expected that, all else equal, a child with parents of a particular education level will have a lower likelihood of being OW than a child with parents of a lower education level. For reasons described in the empirical results section, between the two choices for measures of neighborhood quality, the more interesting variable seems to be neighborhood bad (drugs); therefore this is chosen to represent N_i. H_i is measured with parent obesity status. N_i, H_i, and Z_i may be correlated; thus, to control for this possible endogeneity a two stage probit approach is used.

Predicted Neighborhood Quality

To estimate N_i for two stage probit approach the following model is estimated:

$$\mathit{prob}(N_i) = \Phi(Z_i'\beta_1 + W_i'\beta_2 + \epsilon_{2i}). \qquad (2)$$

prob(N_i) is equal to one if the neighborhood child *i* lives in is of poor quality and is equal to zero if the neighborhood where child *i* lives is of good quality. $\Phi(\cdot)$ is the standard normal cumulative distribution. Again, Z_i is a vector of parent characteristics of child *i*. W_i, is added, which is a vector of exogenous factors that influences the quality of neighborhood a family lives in but is independent of $Z_{i,}$ to predict N_i and identify the system of equations. ϵ_{2i} is a random error term.

The variables included in Z_i are log (income), wife unemployed and married, not married, high school diploma, some college, and college degree or higher. Variables representing the family's total family income and highest education level enter this model because these variables are expected to be negatively correlated with the quality of the neighborhood that a family lives in. Wife unemployed and married is equal to one if wife of the household is married, looking for work, and not working at the time of the 2001 PSID interview. Wife unemployed and married is equal to zero if the wife of the household is not looking for work or the head of household is not married. The not married and wife unemployed and marred variables are included in this model because it is expected that these variables may decrease a family's ability to afford to live in a quality neighborhood. The variables included in W_i are Northeastern region, Western region, Southern region, low affordability rate, high poverty rate, and high unemployment rate. Northeast region, Western region, and Southern region are indicator variables that designate whether the household lives in one of the states that the U.S. Census categorizes in their respective region. The excluded region variable in this analysis is Midwestern region.

Predicted Parent Health

To acquire a prediction of H_i for two stage least squares the following equation is estimated:

$$prob(H_i) = \Phi(Z_i'\delta_1 + V_i'\delta_2 + \epsilon_{3i}). \tag{3}$$

prob(H_i) is equal to one if the parent obese variable is equal to one for child *i* and equal to zero if the parent obese variable is equal to zero for child *i*. $\Phi(\cdot)$ is the standard normal cumulative distribution. Z_i is a vector of parent characteristics. V_i is a vector of exogenous factors that influence parent health but are independent of all the variables included in Z_i. ϵ_{3i} is a random error term. The vector Z_i includes variables income, income squared, single female parent, high school diploma, some college, and college degree or higher. Plantinga and Bernell 2007 note that income and education effect adult BMI, which is why income and educational attainment are included to predict the obesity status of the parents in this equation (Plantinga and Bernell, p862). Income and education are expected to be negatively correlated with obesity status. Single female parent is equal to one if the head of household is female and not married. Stress from being a single female parent may lead to an increased likelihood of obesity, which is why this variable is included in this model. This variable is zero if the head of household is married or male. The variables included in V_i are Northern region, Western region, Southern region, percent that bike to work, percent that walk to work, and population density.

CHAPTER 5 - Empirical Results

Table 1 examines child overweight status by neighborhood and income status. Income status in 2002 is divided into high and low income groups according to the sample the average annual total family income for that year. Notice that 22.3 percent of OW children live in bad neighborhoods as measured by the bad neighborhood (drugs) variable, whereas only 17.8 percent of OW children live in bad neighborhoods. Using a pooled two sample t-test shows that these means are statistically different. Moreover, the difference exists only for low income children. That is, when the means are examined by income status, it becomes clear that, for only lower income households, OW children are more likely to live in good neighborhoods than NOT OW children. There is no difference in neighborhood quality by weight status among children living in higher income households.

Table 3 inspects the means, by child overweight status and by neighborhood quality status, of the variables used in the probit analysis of the data. The first part of Table 3 analyzes only variables produced from PSID data. Many of these variables, when broken down by OW status and neighborhood quality status have means that are significantly different. Using the pooled two sample t-test shows that, OW children are more likely to live in households with a lower income. There is also a significant difference in total family income between households living in good neighborhoods and bad neighborhoods; intuitively, the mean total family income of households living in bad neighborhoods is significantly lower than the mean total family income of families living in good neighborhoods. These means also demonstrate that there is a significantly higher proportion of single female parents living in bad neighborhoods than good ones. OW children who live in two parent households are more likely to live in households where the wife is unemployed and looking for work at the time of the 2001 PSID interview. The number of unemployed wives is also considerably higher in bad neighborhoods than in good neighborhoods. The t-tests show, that OW children are more likely to live in single parent households than not OW children. A significantly larger, fraction of these single parent households live in bad neighborhoods as well. OW children are more likely to live with parents who have a maximum education level less than a high school diploma than NOT OW children. There is also a larger proportion of OW children that live with parents who have a maximum education level of a high school diploma than not OW children. A larger proportion of parents with an education level less than or equal to a high school diploma live in a poor quality neighborhood than in a good quality neighborhood. For parents with a highest education of some college, the difference in means exists only when broken down by neighborhood quality. A higher proportion of these households live in good neighborhoods. Unsurprisingly a larger proportion of Not OW children that live with parents with a maximum education level of a college degree or higher than OW children. Table 3 also shows that a higher proportion of the children that live in low quality neighborhoods, as defined by the neighborhood bad (drugs) variable, are OW than not OW.

The second page of Table 3 shows strictly variables constructed from Census 2000 data. Of the census variables only Northeastern region, percent that walk to work, high poverty rate, and high unemployment rate have means that are significantly different when broken down. There is a bigger fraction of the children who live in the Northeastern region who are NOT OW than OW. Intuitively the average number of people that walk to work in a county is higher for households in good neighborhoods than for those in bad neighborhoods. Households that live in bad neighborhoods are more likely to have high state unemployment rates than households that live in good neighborhoods. OW children are more likely to live in a state with a high state level of unemployment than not OW children. Households living in neighborhoods with poor neighborhood quality are more likely to live in states with high levels of poverty than households living in good neighborhoods.

The final part of Table 3 shows the means for N_i, predicted N_i, H_i and predicted Hi, where the predicted N_i and the predicted H_i are the fitted values of N_i and H_i calculated from the predicted neighborhood quality and predicted parent health equations, respectively. N_i, predicted N_i, and predicted H_i have significantly different means when broken down by child OW status. A larger percentage of OW children live in households that suffer from the negative outcomes these three variables represent. Significantly more of the households in bad neighborhoods than of the households that live in good neighborhoods have at least one parent in the household who is obese and is predicted to have one parent that is obese. The predicted neighborhood quality variable correctly predicts that families living in bad neighborhoods actually live in bad neighborhoods.

Table 4 reports the estimates for the child OW equation. Several models are shown that estimate equation (1). As Table 4 illustrates, the base model that maximizes the log likelihood is model (3), which includes both N_i and H_i. In this model the only variables that have a significant influence on child overweight status are: college degree or higher, age of child, child Hispanic, and parent obese.

Note that, for a probit estimation, the marginal effect of an independent variable on y, denoted $\frac{dy}{dx}$, is $\frac{dy}{dx} = \Phi'(X'\beta)\cdot\beta$ where $\Phi'(X'\beta)$ is equal to the normal probability distribution evaluated at $X'\beta$ or, equivalently, $\bar{y}$. In the analysis, $\bar{y}$ is the fraction of children who are OW. It is equal to 0.348 as shown in Table 3. The marginal effect of the parent having a college degree or higher relative to having no high school diploma is therefore equal to (0.348)·(-0.299), which is equal to -0.104. Thus, the effect of the parent having a college degree or higher relative to no high school diploma decreases the probability of the child being overweight by 10.4 percentage points. Since the sample average OW rate is 34.8 percent, a 10.4 percentage point decrease is a 29.9 percent decline in the probability of the child being overweight ($100\cdot\frac{10.4}{34.8}$). The marginal effect of the age of child variable is computed similarly; it is equal to (0.348)·(-0.016), which is equal to -0.006. Hence, every passing year decreases the probability of the child being overweight by 0.6 percentage points. Since the sample average OW rate is 34.8 percent, a 0.6 percentage point decrease is a 1.7 percent decline in the probability of the child being overweight. The marginal effect of the child Hispanic variable is equal to (0.348)·(0.360), which is equal to 0.12. Thus, the probability of a Hispanic child having OW status, relative to a non-Hispanic child, is 12.0 percentage points higher. Since the sample average OW rate is 34.8 percent, a 12 percentage point increase is a 35 percent rise in the probability of the child being OW. That is, all else equal, Hispanic children are 35 percent more likely to be OW than non-Hispanic children. The marginal effect of the parent obese variable is equal to (0.348)·(0.479), which is equal to 0.167, a very large effect. Thus, the effect of having at least one parent that is obese increases the probability of the child being OW by 16.7 percentage points. This implies that having an obese parent increases the likelihood of the child being OW by 48 percent ($100\cdot\frac{16.7}{38.4}$).

Not that a Wald test statistic is computed by $\lambda = (R\beta - r)'(R\hat{V}R')^{-1}(R\beta - r) \sim \chi^2(J)$, where β is an $(N \times 1)$ column vector of the coefficients estimated in the model, $\hat{V}$ is the $(N \times N)$ estimated variance matrix, and J is the number of hypotheses to be jointly tested. The null hypotheses to be jointly tested are in the form $R\beta = r$, thus R is an $(J \times N)$ matrix that selects the specific coefficients in the particular linear combination from β to form the hypotheses to be tested and r is a $(J \times 1)$ column vector that is composed of elements that each linear combination of the coefficients to be tested is equal to (Griffiths, Hill, and Judge, p454). When a joint test of significance test is done, using a Wald test, on the neighborhood bad (drugs) and parent obese variables, the null hypothesis is rejected. This means that these variables are jointly significant when predicting child OW status. Referring to model (1), A 50 percent increase in the likelihood of living in an unsafe neighborhood increases the likelihood of being OW by $0.50\cdot\frac{dy}{dx}$ = 0.50·(0.348)·0.152 = 2.6 percentage points which is a 6, $100\cdot\frac{2.6}{34.8}$, percent increase in the likelihood of OW status.[2]

[2] When an interaction term, low income interacted with neighborhood bad (drugs,) is added to the child OW equation, the interaction term is not significant, and the significance of the other variables in the model does not change. Low income is equal to one if the household has a total family income less than or equal to the mean total family income of the sample, and zero otherwise. When this interaction term enters while

The auxiliary equations that generate the predicted values for neighborhood quality and parent health status are reported in Tables 5 and 6, respectively. Referring to Table 5, the log likelihood is maximized by model (6), which has the census controls for the availability of safe and affordable neighborhoods in the model. In this model the variables that have significant influence over neighborhood quality are: wife unemployed status, marital status, high school diploma, and low affordability. As this model is estimated as a probit, the marginal effects of the independent variables are calculated in a similar fashion as the marginal effects for equation (1). For this equation the dependant variable, N_i, has a mean of 0.17 as shown in Table 3. Hence, the marginal effect of the household having an unemployed wife is equal to $(0.17)\cdot(0.698)$, which is equal to 0.119. The effect of the household having an unemployed wife thus increases the probability of the family living in a poor quality neighborhood by 11.9 percentage points, which is a huge effect. Specifically, since the sample average of families living in bad neighborhoods is 17.0 percent, an 11.9 percentage point rise is a 70.0 percent, $(100\cdot\frac{11.9}{17})$, increase in the probability that the family lives in a bad neighborhood. The marginal effect of being single is equal to $(0.17)\cdot(0.465)$, which is equal to 0.08. The effect of the household being a single parent household thus increases the probability of the family living in a poor quality neighborhood by 8.0 percentage points, also a large impact. An 8.0 percentage point rise corresponds to a 47.19 percent, $(100\cdot\frac{8}{17})$, increase in the probability that the family lives in a bad neighborhood. The marginal effect of the parent having a high school diploma relative to having no high school diploma is equal to

dropping the neighborhood control, all variables maintain their significance except child black, which loses it's significance.

$(0.17)\cdot(0.337)$, which is equal to 0.057. Thus, the effect of the parent having a high school diploma relative to no high school diploma increases the probability of the household living in a poor quality neighborhood by 5.7 percentage points. Since the sample average of families living in bad neighborhoods is 17.0 percent, a 5.7 percentage point increase is a 33.5 percent rise in the probability of the family living in a poor quality neighborhood. To find the marginal effect for the low affordability variable consider a 10 percent increase in the variable, which means that the number of unaffordable counties in a state increase by 10 percent. The marginal effect in this case would be $0.1\cdot\frac{dN}{dx} = 0.1\cdot(0.17)\cdot(0.401)$, which is equal to 0.007. Thus, the effect of a 10 percent increase in the number of low affordability counties increases the probability of the household living in a poor quality neighborhood by 0.7 percentage points. Since the sample average of families living in bad neighborhoods is 17.0 percent, a 0.7 percentage point increase is a 4.1 percent, $(100\cdot\frac{.7}{17})$, rise in the probability of the child being overweight.

Referring to Table 6, which reports the parent health estimations, the model that maximizes the log likelihood is model (6), which includes the census controls that influence parent health. Significant variables in this model include: income, income squared, single female parent, college degree or higher, and percent that bike to work. This model is a probit and therefore the marginal effects are obtained in the previously discussed fashion. The mean of the dependent variable, H_i, is 0.354 as shown in Table 3.

However, the marginal effect of income is calculated slightly differently, since income squared enters this model. The marginal effect, $\frac{dH}{dx}$ in this case, of an independent variable plus its square, on the dependent variable in a probit model is equal to:

$$\frac{dH}{dx} = \Phi'(X'\beta)\cdot(-0.022 + 2\cdot 0.0001\cdot x).$$

In this case -0.022 is the coefficient estimate for income, and 0.0001 is the coefficient estimate for income squared. $\Phi'(X'\beta)$ is estimated by using the mean of H_i, and the sample mean of income for X, which is equal to 7.71. Therefore, the marginal effect of income on H_i is equal to $(0.354)\cdot[-0.022+2\cdot(7.71)\cdot(0.0001)]$, which is equal to -0.007. Thus, for every additional $10,000 of yearly income the probability of the family having least one parent classified as obese decreases by 0.7 percentage points. Since the sample average H_i is 35.4 percent, a 0.8 percentage point decrease is a 2.0 percent, $(100\cdot\frac{0.8}{35.4})$, decline in the probability of having one parent classified as obese.

The marginal effect of the household having a single mom is equal to $(0.354)\cdot(0.234)$, which is equal to 0.083. The effect of the household having a single mother thus increases the probability that the family has at least one parent classified as obese by 8.3 percentage points. Since the sample average of families with at least one parent classified as obese is 35.4 percent, an 8.3 percentage point rise is a 23.4 percent, $(100\cdot\frac{8.3}{35.4})$, increase in the probability that the family has at least one parent that is classified as obese. The marginal effect of the parent having a college degree or higher relative to having no high school diploma is equal to $(0.354)\cdot(-0.321)$, which is equal to -0.114. Thus, the effect of the parent having a college degree or higher relative to no high school diploma decreases the probability that the family has at least one parent classified as obese by 11.4 percentage points. Since the sample average of families with at least one parent classified as obese is 35.4 percent, an 11.4 percentage point decrease is a 32.2 percent, $(100\cdot\frac{11.4}{35.4})$, decline in the probability that the family has at least one parent who is classified as obese.

To find the marginal effect for the percent that bike to work variable consider a one percent increase in the variable, which means that one percent more people in the county bike to work on a regular basis. The marginal effect, denoted $\frac{dH}{dx}$, in this case would be $0.01\cdot\frac{dH}{dx} = 0.01\cdot (0.354)\cdot(-17.80)$, which is equal to -0.063. Thus, the effect of a one percent increase in the percent that bike to work variable decreases the probability that at least one parent is obese in the household by 6.3 percentage points. Since the sample average of families with at least one parent classified as obese is 35.4 percent, a 6.3 percentage point decrease is a 17.8 percent, $(100\cdot\frac{6.3}{35.4})$, decline in the probability that at least one parent in the household is classified as obese.

CHAPTER 6 - Conclusion

Intuition suggests that neighborhood safety influences child health. Children living in unsafe neighborhoods likely spend less time outside participating in non-sedentary activities that children from safe neighborhoods can take part in. The mental stress of living in a bad neighborhood might also increase the likelihood of child overweight status. This study uses the PSID, PSID CDS, and Census 2000 SF 3 data to form a nationally representative study of the effects of neighborhood quality on childhood overweight status, controlling for parent health. The data includes observations on 1917 children, their families and the neighborhoods they live in during 2002.

Child overweight status is modeled to be a function of parent characteristics, child characteristics, and neighborhood quality. This model shows that the factors that influence child overweight status include: parent education level, child age, child race, and child ethnicity. Neighborhood quality and parent obesity status may be endogenous to parent factors such as education and income; a two stage probit approach is used to circumvent this potential issue. Specifically, this study generates instruments for neighborhood quality and parent health status by estimating separate probit models for the likelihood of children living in bad neighborhoods and having parents who are obese. In the neighborhood models marital status, employment status of the wife, educational status of the parents, and the availability of affordable neighborhoods significantly influence neighborhood quality. In the parent health models income, the sex of the head of household in single parent homes, educational status of the parents, and the percent of people who bike to work in the county where the child lives significantly predict parent obesity status.

Examining the sample means suggests that neighborhood quality influences overweight status for low income children. The probit analysis suggests that unsafe neighborhoods increase the likelihood of a child being overweight. A 50 percent increase in the likelihood of living in an unsafe neighborhood increases the likelihood of being OW by 7.4 percent. This effect lessens when parent obesity status is controlled for, but a joint hypothesis test shows that parent health and neighborhood quality are jointly significant in determining the likelihood of child OW status. This study concludes that parent educational status, parent obesity status, child race, child ethnicity, and neighborhood quality all impact child overweight status.

References

American Obesity Association. 2009. “Childhood Overweight.” The Obesity Scene. 2009 <http://www.obesity.org/information/childhood_overweight.asp> 22 Apr. 2009.

Burkhauser, Richard and John Cawley. 2007. “Beyond BMI: The value of more accurate measures of fatness and obesity in social science research.” *Journal of Health Economics,* 27, 519-529.

Centers for Disease Control and Prevention. 18 Jan. 2009. “Defining Overweight and Obesity.” Overweight and Obesity. <http://www.cdc.gov/nccdphp/dnpa/obesity/defining.htm> 22 Apr. 2009

Centers for Disease Control and Prevention. 27 Jan. 2009. “About BMI for Children and Teens.” Healthy Weight - it's not a diet, it's a lifestyle! < http://www.cdc.gov/healthyweight/effects/index.html > Apr. 22 2009

Centers for Disease Control and Prevention. 24 Mar. 2009. “Introduction." Overweight and Obesity. <http://www.cdc.gov/healthyweight/assessing/bmi/childrens_BMI/about_childrens_BMI.html> 22 Apr. 2009.

Condliffe, Simon and Charles Link. 2008. “The Relationship between Economic Status and Child Health: Evidence from the United States.” *The American Economic Review*, 98(4), 1605-1618.

Ewing, Reid Tom Schmid, Richard Killingsworth, Amy Zlot, and Stephen Raudenbush 2007. “Disparities in obesity rates: Analysis by ZIP code area.” *Social Science and Medicine*, 65(12), 2458-2463.

Fertig, Angela, Gerhard Glom, and Rusty Tchernis. 6 Dec. 2006. “The Connection between Maternal Employment and Childhood Obesity: Inspecting the Mechanisms.” Working Paper. Indiana University Bloomington. Center for Applied Economics and Policy Research. <http://ssrn.com/abstract=949855>

Griffiths, William E., R. Carter Hill, George G. Judge. 1993. <u>Learning and Practicing Econometrics</u>, Canada: John Wiley & Sons, Inc. 454.

Himes JH, Dietz WH. 1994. "Guidelines for overweight in adolescent preventive services: Recommendations from an expert committee." *American Journal of Clinical Nutrition*, 59, 307–316.

Hofferth, Sandra and Sally Curtin. 2005. “Poverty, Food Programs, and Childhood Obesity.” *Journal of Policy Analysis and Management,* 24(4), 703-726.

Liu, Gilbert, Jeffrey Wilson, Rong Qi, Jun Ying. 2006. “Green Neighborhoods, Food Retail and Childhood Overweight: Differences by Population Density.” *The Science of Health Promotion*, 21(4), 317-325.

Platinga, Andrew and Stephanie Bernell. 2007. “The Association Between Urban Sprawl and Obesity: is It A Two-Way Street?” *Journal of Regional Science,* 47(5), 857-879.

White, Halbert. 1980. “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity.” *Econometrica*, 48(4), 817-838.

Appendix A - Tables

Table 1. Percent of Children Overweight (OW) by Neighborhood and Income Status. PSID CDS 2002.

	All Children		Low-income children		High-income children	
	OW	Not OW	OW	Not OW	OW	Not OW
Bad neighborhood (parent rating)	0.026	2.86	3.55	3.96	1.21	1.89
	(15.57)	(17.04)	(17.6)	(19.02)	(6.89)	(11.41)
Bad neighborhood (drugs)	22.27*	17.83*	25.71*	20.54*	18.77	16.18
	(40.66)	(39.17)	(41.48)	(39.46)	(37.13)	(38.06)

Notes: (i) Data are weighted using the PSID 2002 child level weight. (ii) Income groups are divided by the mean family total annual income for the sample. (iii) Child overweight status is determined by the CDC (2008) criteria. (iv) The numbers in parentheses are standard deviations. (v) * denotes that means are significantly different at the 99% confidence level.

Table 2. Data Sources

Data Source	**Variables**			
PSID 1997-2003	Log (income)	Income	Income squared	Parent obese
	Parent no high school diploma	Parent high school diploma	Parent some college	Parent college degree or higher
	Single female parent	Not married	Wife unemployed and married	
PSID CDS 2002	Child overweight	Age of child	Child Hispanic	Child black
	Bad neighborhood (parent rating)	Bad neighborhood (drugs)		
Census SF 3 2000	Northeastern region	Western Region	Southern Region	Midwestern region
	Low affordability	High poverty rate	High unemployment rate	
	Population density	Percent that bike to work	Percent that walk to work	

Table 3. Weighted Sample Means, Standard Deviations in Parentheses, and Test for Difference in Means by Child Overweight (OW) and Neighborhood Status.

	Full sample	OW	Not OW	Neighborhood bad (drugs)	Neighborhood good (drugs)
Child overweight	0.348 (0.485)	1	0	0.427*** (0.026)	0.329*** (0.012)
Income	7.741 (8.864)	7.062** (7.782)	8.103** (9.438)	6.592*** (9.680)	8.021*** (8.640)
Income squared	135.6 (1201)	112.8 (613)	147.8 (1444)	131.2 (976)	136.7 (1248)
Log (income)	10.95 (0.847)	10.83*** (0.843)	11.01*** (0.842)	10.72*** (0.876)	11*** (0.831)
Single female parent	0.095 (0.299)	0.109 (0.306)	0.088 (0.295)	0.156*** (0.375)	0.081*** (0.276)
Wife unemployed	0.019 (0.144)	0.028* (0.166)	0.015* (0.128)	0.045*** (0.22)	0.013*** (0.120)
Not married	0.098 (0.305)	0.081*** (0.32)	0.169*** (0.294)	0.169*** (0.39)	0.081*** (0.278)
Parent no high school diploma	0.118 (0.328)	0.159*** (0.359)	0.095*** (0.307)	0.202*** (0.416)	0.096*** (0.3)
Parent high school diploma	0.263 (0.448)	0.305*** (0.489)	0.240*** (0.496)	0.343*** (0.514)	0.243*** (0.485)
Parent some college	0.185 (0.396)	0.168 (0.398)	0.194 (0.452)	0.154* (0.414)	0.193* (0.437)
Parent college degree or higher	0.188 (0.398)	0.117*** (0.419)	0.226*** (0.479)	0.108*** (0.41)	0.207*** (0.467)

Table 3. *Continued.* Weighted Sample Means, Standard Deviations in Parentheses, and Test for Difference in Means by Child Overweight (OW) and Neighborhood Status.

	Full sample	OW	Not OW	Neighborhood bad (drugs)	Neighborhood good (drugs)
Northeastern region	0.169 (0.379)	0.144** (0.340)	0.183** (0.400)	0.161 (0.370)	0.172 (0.380)
Southern region	0.344 (0.480)	0.408 (0.467)	0.359 (0.487)	0.341 (0.482)	0.345 (0.479)
Western region	0.243 (0.433)	0.236 (0.410)	0.247 (0.447)	0.269 (0.452)	0.237 (0.429)
Midwestern region	0.243 (0.433)	0.251 (0.419)	0.239 (0.442)	0.229 (0.428)	0.246 (0.434)
Percent that bike to work	0.004 (0.005)	0.0038 (0.004)	0.0042 (0.005)	0.004 (0.005)	0.004 (0.005)
Percent that walk to work	0.029 (0.02)	0.029 (0.021)	0.029 (0.01)	0.027** (0.017)	0.03** (0.021)
Population density	15.56 (44.21)	17.57 (41.74)	14.50 (45.63)	18.45 (32.52)	14.87 (46.53)
Low affordability	0.411 (0.382)	0.395 (0.363)	0.420 (0.393)	0.432 (0.401)	0.406 (0.377)
High poverty rate	0.396 (0.258)	0.410 (0.249)	0.389 (0.263)	0.430*** (0.236)	0.388*** (0.263)
High unemployment rate	0.395 (0.231)	0.411** (0.208)	0.387** (0.243)	0.428*** (0.222)	0.387*** (0.232)

Table 3. *Continued.* Weighted Sample Means, Standard Deviations in Parentheses, and Test for Difference in Means by Child Overweight (OW) and Neighborhood Status.

	Full sample	OW	Not OW	Neighborhood bad (drugs)	Neighborhood good (drugs)
Neighborhood bad (drugs)	0.196 (0.404)	0.241*** (0.419)	0.172*** (0.393)	1	0
Predicted neighborhood bad (drugs)	0.170 (0.075)	0.175** (0.072)	0.168** (0.076)	0.186*** (0.090)	0.167*** (0.070)
Parent obesity status	0.350 (0.486)	0.479 (0.490)	0.281 (0.468)	0.435*** (0.512)	0.329*** (0.477)
Predicted parent obesity Status	0.354 (0.104)	0.368*** (0.100)	0.346*** (0.105)	0.378*** (0.108)	0.347*** (0.102)

Notes: (i) The tests for differences in means are done by overweight status and neighborhood status. (ii) Sample means are weighted using the CDS 2002 child level weights. (iii) The sample size is 1917 for each variable except “Northeastern region,” “Southern region,” “Western region,” “Midwestern region,” “high poverty rate” and “high unemployment rate,” which all have sample sizes of 1663. (iv) ***, **, and * denote that means are significantly different at the 99, 95, and 90 confidence levels, respectively. (v) Dollars values are expressed in 2003 dollars (vii) Child overweight status is determined by the CDC (2008) criteria as discussed in the text.

Table 4. Probit Estimates of Child Overweight Status. PSID Data and 2000 Census Data.

Model	***(1)***	***(2)***	***(3)***	***(4)***	***(5)***	***(6)***
Intercept	-0.167 (0.449)	-0.368 (0.452)	-0.409 (0.453)	-0.326 (0.541)	0.026 (0.459)	-0.190 (0.550)
Log (income)	-0.0099 (0.041)	-0.0053 (0.041)	-0.0026 (0.041)	-0.0032 (0.043)	-0.017 (0.041)	-0.0062 (0.043)
Parent high school diploma	0.072 (0.073)	0.087 (0.074)	0.080 (0.074)	0.080 (0.073)	0.117 (0.077)	0.116 (0.077)
Parent some college	-0.056 (0.082)	-0.052 (0.083)	-0.052 (0.083)	-0.044 (0.084)	-0.071 (0.083)	-0.060 (0.084)
Parent college degree or higher	-0.355*** (0.09)	-0.301*** (0.091)	-0.299*** (0.091)	-0.331*** (0.098)	-0.381*** (0.091)	-0.352*** (0.099)
Age of child	-0.015* (0.0078)	-0.016** (0.008)	-0.016** (0.0079)	-0.015* (0.0078)	-0.014* (0.0078)	-0.014* (0.0078)

Table 4. *Continued.* Probit Estimates of Child Overweight Status. PSID Data and 2000 Census Data.

Model	***(1)***	***(2)***	***(3)***	***(4)***	***(5)***	***(6)***
Child black	0.225* (0.09)	0.158* (0.092)	0.147 (0.092)	0.228** (0.092)	0.239*** (0.09)	0.227** (0.092)
Child Hispanic	0.349*** (0.089)	0.380*** (0.089)	0.360*** (0.09)	0.382*** (0.089)	0.395*** (0.089)	0.403*** (0.09)
Neighborhood bad (drugs)	0.152** (0.075)		0.117 (0.076)			
Parent obese		0.485*** (0.062)	0.479*** (0.062)			
Predicted neighborhood bad (drugs)					-0.659 (0.462)	-0.662 (0.462)
Predicted parent obese				0.266 (0.38)		0.271 (0.38)
Log Likelihood	*-1247.3*	-1218.5	-1217.3	-1249.1	-1248.4	-1248.1

Notes: (i) The dependent variable in this analysis indicates the probability that the child is overweight as measured by the "child overweight" variable. (ii) The sample size for this analysis is 1917 PSID households. (iii) *, **, *** indicate significance at the 90, 95, and 99 percent confidence levels, respectively. (iii) The sample is weighted using the CDS 2002 child level weights. (iv) Dollars values are expressed in 2000 dollars.

Table 5. Probit Estimates of Neighborhood Quality. PSID Data and 2000 Census Data.

Model	***(1)***	***(2)***	***(3)***	***(4)***	***(5)***	***(6)***
Intercept	-0.585 (0.66)	-0.581 (0.689)	-0.56 (0.689)	-0.687 (0.708)	-0.588 (0.708)	-0.636 (0.713)
Log (income)	-0.041 (0.06)	-0.043 (0.062)	-0.05 (0.062)	-0.044 (0.0624)	-0.049 (0.062)	-0.046 (0.063)
Wife unemployed and married	0.709*** (0.268)	0.705*** (0.271)	0.701*** (0.271)	0.701*** (0.271)	0.701*** (0.271)	0.698** (0.271)
Not married	0.509** (0.208)	0.508** (0.209)	0.484** (0.21)	0.474** (0.211)	0.484** (0.21)	0.465** (0.211)
Parent high school diploma	0.324*** (0.106)	0.325*** (0.106)	0.333*** (0.107)	0.334*** (0.107)	0.331*** (0.107)	0.337*** (0.107)
Parent some college	-0.104 (0.12)	-0.103 (0.12)	-0.104 (0.12)	-0.102 (0.12)	-0.103 (0.12)	-0.103 (0.12)
Parent college degree or higher	-0.11 (0.126)	-0.112 (0.126)	-0.12 (0.127)	-0.117 (0.127)	-0.119 (0.127)	-0.117 (0.127)

Table 5. *Continued.* Probit Estimates of Neighborhood Quality. PSID Data and 2000 Census Data.

Model	***(1)***	***(2)***	***(3)***	***(4)***	***(5)***	***(6)***
Northeastern region		0.031 (0.135)	-0.101 (0.158)	-0.089 (0.158)	-0.097 (0.156)	-0.103 (0.160)
Western Region		0.0173 (0.124)	-0.234 (0.197)	-0.276 (0.203)	-0.235 (0.197)	-0.312 (0.209)
Southern Region		-0.0079 (0.117)	-0.0017 (0.118)	-0.061 (0.138)	-0.0040 (0.118)	-0.115 (0.155)
Low affordability			0.347 (0.211)	0.359* (0.211)	0.343 (0.212)	0.401* (0.218)
High poverty rate				0.193 (0.236)		0.424 (0.385)
High unemployment rate					0.035 (0.208)	-0.259 (0.338)
Log Likelihood	*-546.4*	*-546.3*	*-544.9*	*-544.6*	*-544.9*	*-544.3*

Notes: (i) The dependent variable indicates the probability that the household lives in a "bad" neighborhood as measured by the "neighborhood bad (drugs)" variable. (ii) The sample size for this analysis is 1112 PSID households. (iii) *, **, *** indicate significance at the 90, 95, and 99 percent confidence levels, respectively. (iii) The sample is weighted using the CDS 2002 child level weights. (iv) Dollars values are expressed in 2000 dollars. (v) Data applies to the household in the year 2000.

Table 6. Probit Estimates of Parent Obesity Status. PSID Data and 2000 Census Data.

Model	***(1)***	***(2)***	***(3)***	***(4)***	***(5)***	***(6)***
Intercept	-0.208*** (0.077)	-0.154 (0.103)	-0.086 (0.107)	-0.034 (0.121)	-0.024 (0.121)	-0.022 (0.121)
Income	-0.021** (0.0081)	-0.021** (0.0082)	-0.021** (0.0082)	-0.022*** (0.0082)	-0.022*** (0.0082)	-0.022** (0.0082)
Income squared	0.0001* (0.0001)	0.0001* (0.0001)	0.0001* (0.0001)	0.0001* (0.0001)	0.0001* (0.0001)	0.0001* (0.0001)
Single female parent	0.267** (0.115)	0.241** (0.116)	0.238** (0.116)	0.235** (0.116)	0.235** (0.116)	0.234** (0.116)
High school diploma	0.042 (0.087)	0.011 (0.088)	0.0052 (0.088)	0.0059 (0.088)	0.0031 (0.088)	0.0032 (0.088)
Some college	-0.079 (0.099)	-0.094 (0.099)	-0.093 (0.099)	-0.104 (0.099)	-0.100 (0.100)	-0.100 (0.100)
College degree or higher	-0.327*** (0.108)	-0.323*** (0.108)	-0.321*** (0.108)	-0.323*** (0.108)	-0.321*** (0.108)	-0.321*** (0.108)

Table 6. *Continued.* Probit Estimates of Parent Obesity Status. PSID Data and 2000 Census Data.

Model	***(1)***	***(2)***	***(3)***	***(4)***	***(5)***	***(6)***
Northeastern region		-0.112 (0.113)	-0.117 (0.113)	-0.076 (0.114)	-0.094 (0.115)	-0.095 (0.115)
Southern region		0.105 (0.094)	0.087 (0.094)	0.075 (0.095)	0.072 (0.095)	0.072 (0.095)
Western region		-0.244 (0.105)	-0.127 (0.115)	-0.241 (0.105)	-0.146 (0.116)	-0.147 (0.117)
Percent that bike to work			-21.77** (9.133)		-17.92* (9.684)	-17.80* (9.715)
Percent that walk to work				-3.639* (1.940)	-2.244* (2.050)	-2.41 (2.324)
Population density						0.0001 (0.001)
Log Likelihood	-847.8	-840.7	-837.7	-838.8	-837.1	-837.0

Notes: (i) The dependent variable in this analysis indicates the probability that one of the parents in the household is obese as measured by the "parent obese" variable. (ii) The sample size for this analysis is 1291 PSID households. (iii) *, **, *** indicate significance at the 90, 95, and 99 percent confidence levels, respectively. (iii) The sample is weighted using the CDS 2002 child level weights. (iv) Dollars values are expressed in 2000 dollars.