5-2014

The Effect of School and Neighborhood Environmental Factors on Childhood Obesity

Yiwei Qian
University of Arkansas, Fayetteville

The Effect of School and Neighborhood Environmental Factors on Childhood Obesity

The Effect of School and Neighborhood Environmental Factors on Childhood Obesity

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Economics

by

Yiwei Qian
Shanghai JiaoTong University
Bachelor in Finance, 2010
University of Arkansas
Master of Art in Economics, 2011

May 2014
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Dr. Rodolfo M. Nayga Jr.
Dissertation Director (Co-chair)

Dr. Arya B. Gaduh
Dissertation Director (Co-chair)

Dr. Michael R. Thomsen
Committee Member

ABSTRACT

The dissertation consists of three chapters exploring the effect of school program and neighborhood environment on childhood obesity outcome using individual panel data set of Arkansas public schoolchildren.

The first chapter (Section 2) investigates how the Fresh Fruit and Vegetable Program (FFVP), a program that provides funding for the distribution of free fresh fruits and vegetables to students in participating schools, affects childhood obesity. We combine matching methodology and difference-in-differences analysis to estimate the effect of the FFVP on childhood BMI outcomes. Estimates of the FFVP effect are sensitive to different matching methods. Methods that provide a good balance between treatment and control samples show that the FFVP program causes an economically meaningful reduction in the body mass index of participating children. Less strict matching methods yield insignificant results.

The second chapter (Section 3) measures the effect of fast-food restaurant density around the residences of Arkansas public schoolchildren on BMI outcomes. We use the distance from the child's residence to the nearest US highway or interstate highway as an instrument for the density of fast food restaurants. The results show that the exposure of fast food restaurants around the home environment does have significant and positive effects on children's BMI z-scores. Our results also indicate that some subpopulations -- children who are more affluent, rural, non-minority and female -- are disproportionately affected by fast food proximity.

Finally, the third chapter (Section 4) analyzes the effect of neighborhood parks around residences of northwestern Arkansas children on BMI outcomes. Our dataset covers the 2004 through 2007 period. To build comparative groups, we employ propensity score matching to

measure the average treatment effect on the treatment group. The results indicate that proximity of neighborhood parks from the residence have a significant and negative effect on children's BMI z-scores in both the rural and urban areas, with some heterogeneity in the effects across gender. Specifically, our results show that girls in urban and rural areas are significantly influenced by neighborhood parks. The park effect is significant for boys in rural areas but not for boys in urban areas.

ACKNOWLEDGEMENTS

Special thanks are due to my advisor, Dr. Rodolfo Nayga Jr. and Dr. Arya Gaduh and my committee member, Dr. Michael Thomsen for all their guidance to help me finish my dissertation. Also thanks to Dr. Heather Rouse who provided helpful comments.

Grant West, Diana Danforth and Stephen Lein helped me a lot to collect and manage the data. Also, special thanks to Dr. Jebaraj Asirvatham and Dr. Pedro Alviola who gave me several ideas to create statistical programs for my data analysis. I also want to thank my family members who gave their best supports.

TABLE OF CONTENTS

1. INTRODUCTION

Obesity prevalence among children and adolescents in the United States has significantly increased during the past few decades. It is now a major health problem and poses a challenge for government, public health agencies and medical communities. Approximately 13 million U.S. children and adolescents are considered obese[1], with a body mass index (BMI) at or above the 95th percentile. Ogden et al. (2010) indicated that from 1980 to 2008, obesity rates nearly tripled — from 6.5% to 19.6% — for children aged 6 to 11 and more than tripled for adolescents age 12 to 19—from 5% to 18.0%. Adding to its importance, obese adolescents have an 80% chance of becoming obese adults, which places them at greater risk for health problems throughout life (Guo and Chumlea 1999).

Among others, obesity can be caused by two ways: more calorie intake and less calorie output, which is likely to correlate with people's eating behaviors and physical activities. However, children's behaviors can be influenced by the environments they have around their homes and schools (Anderson, Butcher and Levine, 2003). Thus it is worthwhile to investigate the effect of school and environmental factors on children's obesity outcome. Our research explores three aspects which can potentially affect children's eating behaviors and physical activities: the Fresh Fruit and Vegetable Program, neighborhood fast food restaurants and neighborhood parks. We focus our study on children in the state of Arkansas. Arkansas is an interesting case to study since it has one of the highest childhood obesity rates in the US. The National Survey of Children's Health indicated that in Arkansas, about 32.9% of 10-17 year old children were either obese or overweight in 2005 and this percentage increased to 37.5% in

[1]Obesity is defined as body mass index (BMI) at or above the 95th percentile based on the 2000 Centers for Disease Control and Prevention BMI-for-age growth charts. Children with BMI between the 85th and 95th percentile are classified as overweight.

2007[2]. Additionally, Arkansas was the first state to legislatively mandate the measurement and collection of BMI for every public school student starting in 2003 (Arkansas Act 1220 of 2003). Measured annually, these data provide a unique opportunity to study child weight status and the programs and policies designed to impact BMI.

In first chapter (Section 2), we investigate how the Fresh Fruit and Vegetable Program (FFVP), a nutrition assistance program that provides funding for the distribution of free fresh fruits and vegetables to students in participating schools, affects childhood obesity. We combine matching methodology and difference-in-differences analysis to estimate the effect of the FFVP on childhood BMI outcomes. Our results suggest that FFVP effects are sensitive to the use of matching methods, but when using stricter matching methods (i.e., matching methods that produce more balance between control and treatment samples), FFVP participation reduces children's BMI measures.

In second chapter (Section 3), we measure the effect of fast-food restaurant density around the residences of Arkansas public schoolchildren on BMI outcomes. We use the distance from the child's residence to the nearest US highway as an instrument for the density of fast food restaurants. Highway proximity has been shown to exogenously increase fast food availability and has been used as an instrumental variable in studies linking body weight to fast-food availability. The results show that exposure to fast food restaurants around the home environment does have significant and positive effects on children's BMI z-scores. Our results also indicate that some subpopulations -- children who are more affluent, rural, non-minority and female -- are disproportionately affected by fast food.

[2] Source: Childhood Obesity Action Network. State Obesity Profiles, 2008.

In third chapter (Section 4), we analyze the effect of neighborhood parks around residence of northwestern Arkansas children on BMI outcomes. We use a statewide panel dataset for Arkansas covering the 2004 through 2007 period. To build comparative groups, we consider those living near a park as treatment group and others as control groups. We then employ a propensity score matching approach to measure the average treatment effect. The results indicate that the exposure of neighborhood parks and trails around the home environment does have significant and positive effects on urban children's BMI z-scores for the two-mile treatment group and rural children's BMI z-scores for the five-mile treatment group. Our results also show that both urban and rural girls are significantly influenced by neighborhood parks. The park effect is not significant for urban boys but is significant for rural boys.

REFERENCES

1. Ogden, C., M.D. Carroll, L.R. Curtin, M.M. Lamb, and K.M. Flegal. 2010. "About childhood obesity." *JAMA*, 303: 242-249.

2. Guo, S. S., and W. C. Chumlea. 1999. "Tracking of body mass index in children in relation to overweight in adulthood." *The American Journal of Clinical Nutrition*, 70(1): 145-148.

3. Anderson, P.M., K.F. Butcher, and P.B. Levine. 2003a. "Maternal employment and overweight children." *Journal of Health Economics*, 22(3): 477-504.

4. National Initiative for Children's Healthcare Quality, Child Policy Research Center, and Child and Adolescent Health Measurement Initiative. *The National Survey of Children's Health. Childhood Obesity Action Network. State Obesity Profiles, 2008&2009.*

2. THE EFFECT OF THE FRESH FRUIT AND VEGETABLE PROGRAM ON CHILDHOOD OBESITY

Authors: Yiwei Qian, Rodolfo M. Nayga Jr, Michael R. Thomsen, Heather L. Rouse

2.1 BACKGROUND AND LITERATURE REVIEW

Increasing fruit and vegetable intake would decrease high-fat/high-sugar intake for children and their parents, and could be a useful approach to preventing childhood obesity (Epstein et al., 2001). However, children and adolescents in US do not consume the recommended amounts of fruits and vegetables. The United States Department of Agriculture (USDA) guidelines recommend that children eat 6-13 serving of fruits and vegetables each day, but US children only eat 3.5 servings per day on average (Jamelske et al. 2008). Thus, strategies that encourage the consumption of healthier foods such as fruit and vegetables may be one way to address childhood obesity.

The USDA created the Fresh Fruit and Vegetable Program (FFVP) in 2002. This program is intended to increase fruit and vegetable consumption among students in the nation's poorest elementary schools by providing reimbursement to schools for offering fresh fruits and vegetables, free to students, throughout the school day and separately from lunch and breakfast meals. According to the USDA Food Nutrition Service (2010), the objectives of FFVP include: (1) to create healthier school environments by providing healthier food choices; (2) expand the variety of fruits and vegetables available to children; (3) increase children's fruit and vegetable consumption; and (4) make a positive difference in children's diets to impact their present and future health.

Arkansas schools began participating in the FFVP during the 2008-2009 school year. The FFVP is primarily administered through the Arkansas Department of Education (ADE).

Presently, for a school to participate in the FFVP the school must also participate in the National School Lunch Program (NSLP) and at least 50 percent of students must be eligible for the free and reduced lunches. This is to ensure that the program benefits low-income students who otherwise would have fewer opportunities to consume a variety of fruit and vegetables. All students in participating schools are provided fruits and vegetables. Schools are selected based on an application process and program funds are used to reimburse schools for providing fruit and vegetables as snacks at the rate of $50 to $75 per student per year (USDA Food and Nutrition Service, 2010). The average amount of funding per school during the 2008-2009 and 2009-2010 school years was $27,334 and $21,382, respectively.[3] However, nearly twice as many schools participated in the 2009-2010 school year and so the decrease in average funding is not indicative of reduced reach of the program.

There is scant literature, however, on the effectiveness of the FFVP to reduce childhood obesity. Most of the studies on the FFVP are focused on the program's impact on fruit and vegetable consumption. For example, Jamelske et al. (2008) surveyed 784 students who participated and 384 students who did not participate in the FFVP in Wisconsin and found that FFVP participants reported an increased willingness to eat fruits and vegetables compared to non-participants. Davis et al. (2009) surveyed 1,515 high school students who participated in the program and 1,377 high school students who did not participate and compared the fruit and vegetable intakes of both groups. Their results indicated that FFVP participants were more likely than non-participants to consume fruit, juice, and vegetables in amounts recommended by dietary guidelines. Ohri-Vachaspati, Turner and Chaloupka (2012) also conducted a study on 620 public elementary schools participating in the National School Lunch Program during 2009-2010 and

[3] Source: Arkansas Department of Education (ADE) Child Nutrition Unit.

found that FFVP participating schools were significantly more likely (odds ratio 2.07) to serve fresh fruit during lunch meals than FFVP non-participating schools.

Bartlett et al. (2013) evaluated the effect of FFVP on fruit and vegetable consumption and total energy intake for children. Using regression discontinuity, they estimated that the program increased average fruit and vegetable consumption of students in participating schools on FFVP days by approximately one-quarter of a cup per day. They also found no significant increase in total energy intake, which suggests that the increase in fruit and vegetable consumption replaced the consumption of other foods. Boukhris (2007) investigated FFVP participation in Texas and found that there was no significant difference between the FFVP schools and non-FFVP schools in fruit and vegetable expenditures in 2006, but in 2007 the FFVP schools had higher fruit and vegetable expenditures than non-FFVP schools.

Given the promising results of these past studies linking program participation to improvements in fruit and vegetable consumption, it would also be interesting to examine the effect of FFVP on childhood obesity. To our knowledge, no other study has evaluated this issue. In this paper, we use a unique panel dataset that includes measured body mass index (BMI) of school children in Arkansas. We employ difference-in-differences and matching methods to identify the effect of FFVP on children's BMI. Our results suggest that FFVP effects are sensitive to the use of matching methods, but when using stricter matching methods (i.e., matching methods that produce more balance between control and treatment samples), FFVP participation reduces children's BMI measures.

The next section describes the data sources and the variables used in the analysis. Section 2.3 discusses the empirical strategy we used to identify the effect of FFVP participation on children's BMI. Section 2.4 presents the results, describes their sensitivity to different matching methods and concludes.

2.2 DATA

2.2.1 DATA SOURCES

Our data come from three different sources. First, we use FFVP participation data from 2008-2010. These data were obtained from the Arkansas Department of Education (ADE) Child Nutrition Unit and include program participation status and funding information by school and year. There were 24 FFVP schools in the 2008-2009 school year and 47 FFVP schools in the 2009-2010 school year. Second, we use the Arkansas BMI dataset for 2007 to 2010. This is a unique panel dataset at the student level that includes child weight and height data collected by trained personnel in the public schools and maintained through legislative mandate at the Arkansas Center for Health Improvement (ACHI) (Justus et al. 2007). BMI is calculated as a ratio ([weight in pounds / (height in inches)^2] $\times$ 703) and then converted to age-gender specific z-scores according to the Centers for Disease Control and Prevention guidelines (CDC 2013). Measures used for this analysis included the BMI z-score[4], BMI percentile[5], gender, race and free or reduced lunch program participation status. Additionally, ACHI personnel geo referenced and interfaced these data with food store locations so that our final dataset provided measures of the food environment around the children's home and schools. Only children in even-numbered grades (kindergarten through 10th grade) were consistently measured across all years during the period of our study. For this reason, we only include students in kindergarten,

[4] BMI z-score is defined as a deviation of the value for an individual from the mean value of the reference population divided by the standard deviation for the reference population.
[5] BMI percentile is a value of a cumulative probability distribution of BMI z-score.

second, fourth, sixth and eighth grades in our study[6]. Third, we used demographic characteristics data from the American Community Survey's (ACS) 2006-2010 five-year estimates. These include data on proportion of population by race, income level, education, work status and other neighborhood measures for the census block group of the child's residence. We use these as control variables in our models.

2.2.2 VARIABLE DEFINITIONS AND DESCRIPTIVE STATISTICS

The choice of control variables for the matching and the regression models is an important consideration in our study. Matching is a "data hungry" technique in terms of the number of variables required to find matched groups. In our study, the control variables are based on the factors which are hypothesized to affect our outcome variable, children's BMI. Table 1 exhibits the description of the variables used in the analysis.

One important factor for obesity is income level. Wang (2001) indicates that for 10-18 year old children in the US, the obesity and overweight rate is 32.7% for low-income households, 25.5% for middle-income households and 19% for high-income households. Casey et al. (2001) also analyzed data from 5,669 children (0-17 years old) from 3,790 households. They found that children in low-income families reported a higher obesity and overweight rate (46.7%) than children in high-income families (31.5%). Singh, Siahpush and Kogan (2010a) analyzed obesity outcomes for more than 44,000 children from 2003-2007 and found that obesity prevalence for children below the poverty threshold was 27.4%, 2.7 times higher than the prevalence for children with family income exceeding 400% of the poverty threshold. One reason for the inverse relationship between obesity rates and income is that low-income communities often lack access to stores that sell fresh fruit and vegetables and have instead stores that sell foods low in

[6] Since the FFVP targets elementary schools, no 10th graders were affected by the program.

nutritional value. Haynes-Maslow et al. (2013) identified 6 major community-level barriers to accessing fruits and vegetables. These are cost, transportation, quality, variety, the food environment, and societal norms on food. Their research showed that in lower income communities, access to fresh fruit and vegetables can be difficult because of the lack of affordable transportation options. Moreover, the quality and variety of fresh fruit and vegetables can be limited in lower income areas.

To measure and control for access to healthy foods, we computed the distance between the student's residence and the nearest large grocery store that contained a fresh produce department. Grocery stores and their locations in Arkansas, by year, were obtained from Dun and Bradstreet. We adopted the low access area criteria found in the USDA/ERS Food Desert Locator[7]. That is, students living in urban census block groups were classified as having low-access to healthy foods if their residence was more than one mile from a large supermarket. Students in rural block groups were classified as low access if this distance was greater than ten miles. Food access is also affected by transportation options and so controls are included for the proportion of population that uses public transport for commutes to work and for the proportion of families with no vehicle availability.

Educational level, working status and marital status of parents are also important factors for childhood obesity. For example, Nayga (2000) has shown that schooling can influence obesity outcomes. His results also suggested that health knowledge decreases the probability of an individual becoming obese. Singh, Siahpush and Kogan (2010a) found that obesity prevalence for children with parents having less than 12 years of education was 30.4% in 2007, 3.1 times higher than the prevalence for children with parents with a college degree. Obesity prevalence

[7] http://www.ers.usda.gov/data/fooddesert/about.html.

also increased significantly among children from single-mother households from 18.9% in 2003 to 21.9% in 2007. Anderson, Butcher and Levine (2003a) investigated whether children are more or less likely to be overweight if their mothers work and their results indicated that a child is more likely to be overweight if his/her mother worked more intensively.

We do not have information about the education level, working status, and marital status of parents of the students in our sample, but we are able to measure these for the neighborhood of the child's residence using census block group data from the American Community Survey. The BMI data from ACHI also include some important individual-level control variables. These include age in months, gender, ethnicity, and free or reduced lunch participation status. Additional income controls at the census-block-group level include the proportion of population below the poverty level and median value of owner occupied housing units. These control variables are listed in Table 1.

Table 1. Description of Variables in the Study

Variables	**Description**
Outcome Variables	
BMI z-score	Individual's BMI z-score
BMI percentile	Individual's BMI percentile
Treatment Variables	
D1	Binary indicator (1 if period 2; 0 otherwise)
FFVP	Binary indicator (1 if in FFVP participating school; 0 otherwise)
DID	D1* FFVP (DID interaction term)
Control Variables	
Age	Age of student in months
Black	Binary indicator (if individual is Black then =1, 0 otherwise)
Hispanic	Binary indicator (if individual is Hispanic then =1, 0 otherwise)
Male	Binary indicator (if individual is male then =1, 0 otherwise)
Free	Binary indicator (if individual participated in free lunch then =1, 0 otherwise)
Reduced	Binary indicator (if individual participated in reduced lunch then =1, 0 otherwise)
Urban	Binary indicator (if individual lived in urban area=1, 0 otherwise)
Lowaccess	Binary indicator that describes the accessibility to large grocery stores. It takes the value of one for urban students living more than one-mile from a large grocery store and for rural students living more than 10 miles from a large grocery store.
Singlemother_prp	Proportion of families that have children under 18 with female householder with no husband present
Highschool_prp	Proportion of population with high school degree
Somecollege_prp	Proportion of population with some college or an associate's degree
Collegeplus_prp	Proportion of population with college and post-graduate degrees
Incomebelowpoverty	Proportion of population below the poverty level
Workingmother_prp	Proportion of families that have children under 18 with mother in the labor force
Novehicle_prp	Proportion of families with no vehicle availability
Medhousevalue	Median value for owner occupied housing units ($'000)

2.3 METHODOLOGY

A major concern in assessing the effect of FFVP is that FFVP participation by schools is not randomly assigned, so it is possible that schools self-selected into the program. Hence, the characteristics of FFVP participating schools could be quite different from those of non-participating schools. It is also possible that some unobserved factors could influence both FFVP participation and obesity outcomes (e.g., school health related programs and parental factors). The availability of panel data allows us to address some of these endogeneity issues. In addition, since FFVP participation started after the initial period for which we have data, we can compare measures before and after the implementation of FFVP which provides us with a quasi-experimental setting. Hence, in addition to panel data estimation, we also are able to use a difference-in differences (DID) framework. Finally, to alleviate concerns regarding the comparability of the treatment and control groups and to limit model dependence (Campbell et al. 2011; Islam 2011), we use matching techniques prior to running our DID panel models. Heckman, Ichimura, and Todd (1997) concluded that DID matching helps control for heterogeneity in initial conditions and also controls for unobserved determinants of participation. Hence, we attempt to account for potential selection biases by combining matching, DID, and panel estimation methodologies in our analysis. In the same spirit as in Angelucci and Attanasio (2013), our panel DID estimation deals with time-invariant unobserved factors while the matching rebalances the sample to deal with time-varying unobserved factors. We run our panel DID models using several different matching methods.

Our panel data include student level observations from 2007-2010. Since FFVP in Arkansas started during the 2008-2009 school year, we use the 2007-2008 school year as period 1 (or the before period) and the 2009-2010 school year, the 2nd year of the FFVP implementation, as period 2 (or the after period). We then define the treatment group as those students who participated in FFVP in both the 2008-2009 and 2009-2010 school years and the control group as those students who did not participate in FFVP from 2007-2010.

2.3.1 MATCHING

The main idea of matching is to find a group of control individuals that are similar to the treated individuals in all pre-treated characteristics. We use propensity score matching (PSM) and coarsened exact matching (CEM) to match the treated and control groups.

2.3.1.1 PROPENSITY SCORE MATCHING (PSM)

Rosenbaum and Rubin (1983, 1985) introduced Propensity Score Matching (PSM) as a matching method to construct a statistical comparison group that is based on a model of the probability of participating in the treatment conditional on observed characteristics. To get the propensity score, first we run a standard logit model where the dependent variable is the treatment variable, which is *FFVP* participation, and the independent variables are a set of control variables.

One of the most frequently used matching techniques is *nearest-neighbor matching*, where each treatment unit is matched to the comparison unit with the closest propensity score. Rosenbaum and Rubin (1985) and Becker and Ichino (2002) introduced the structure of *nearest-neighbor matching.* Denote by *C(i)* the nearest neighbor matching sets for treated unit *i.* This is defined as:

$$C(i) = \min_{j} \left| p_i - p_j \right| .$$

where *pi* is the propensity score for treated unit and *pj* is the propensity score for control unit. And *nearest-neighbor matching within n neighbors* means that for each matched treated unit, there are *n* matched control units which have the *n* closest propensity scores. In our analysis, we choose the *nearest-neighbor matching* within 2 neighbors and within 3 neighbors.

The other matching algorithm we choose is *Mahalanobis matching.* Rosenbaum and Rubin (1985) introduced the structure of PSM based on the Mahalanobis distance. The Mahalanobis distance is the distance between two *N* dimensional points scaled by the statistical variation in each component of the point. For example, if x_1 and x_2 are two points from the same distribution with covariance matrix, Σ, then the Mahalanobis distance can be expressed as:

$$D(x_1, x_2) = (x_1 - x_2)'\Sigma^{-1}(x_1 - x_2).$$

In our study, we use *Mahalanobis matching* without calipers and *Mahalanobis matching* with calipers of 0.05, 0.075, and 0.1. The use of a caliper provides stricter matches because observations are matched only if their absolute distance in propensity scores is smaller than the caliper. Hence, a treated individual will remain unmatched if the nearest observation in the control group is outside of the bound set by the caliper.

2.3.1.2 COARSENED EXACT MATCHING (CEM)

We also utilize a strict matching method called *coarsened exact matching* (CEM) (Iacus, King and Porro 2011; Iacus, King and Porro 2012; Blackwell, Iacus, King, and Porro 2009). The main motivation for CEM is that while exact matching always provides perfect balance, it typically produces few matches due to the curse-of-dimensionality. The idea of CEM is to temporarily coarsen each variable and then exact match on these coarsened data. Afterwards, the original (uncoarsened) values of the matched data are retained.

The advantage of CEM is obvious in that it generally provides stricter matching criteria compared to PSM and it also allows the analyst to add continuous variables as control covariates. For PSM, if a lot of continuous variables are used in the matching, it is possible that the matched samples have close propensity scores but not close values on these continuous variables. However, for the CEM, the value of every matching variable needs to be the same (after coarsening). Since almost half of control variables of our research are continuous measures, CEM can be a better matching strategy than PSM. In our study, we let the coarsening algorithm cut the range of the continuous variable into equal intervals of length.

To summarize, our matching strategy includes the use of the following matching procedures: *nearest-neighbor matching* within 2 neighbors, *nearest-neighbor matching* within 3 neighbors, *Mahalanobis matching* without calipers, *Mahalanobis matching* with calipers of 0.05, 0.075, and 0.1 and *coarsened exact matching.*

2.3.2 THE IMBALANCE TEST

After matching control observations to treated observations using the seven different methods discussed above, we need to test the degree of imbalance in the covariates in the two groups. The goal of measuring imbalance is to summarize the differences in the multivariate empirical distribution of the pretreatment covariates for the treatment group and matched control group. That is, we wish to assess how similar the control and treated groups are based on a given set of characteristics. In our study, we choose the imbalance test introduced by Iacus, King and Porro (2011); i.e., the $\mathcal{L}_1$ statistic as a comprehensive measure of global imbalance.

To build this measure, Iacus, King and Porro (2011) obtained two multidimensional histograms by direct cross-tabulation of the covariates in the treated and control groups, given a choice of bins for each variable. Let *H(X1)* denote the set of distinct values generated by the bins chosen for variable X_1, i.e., the set of intervals into which the support of variable *X1* has been cut. Then, the multidimensional histogram is constructed from the set of cells generated as *H(X1)*×· · ·×*H(Xk) = H(X) = H.* Set *f* and *g* as the relative empirical frequency distributions for the treated and control units, respectively and record the k-dimensional relative frequencies for the treated $f_{\ell_1 \dots \ell_k}$ and control $g_{\ell_1 \dots \ell_k}$ units. The measure of imbalance is the absolute difference over all the cell values:

$$\mathcal{L}_1(f,g) = \frac{1}{2}\sum_{\ell_1 \dots \ell_k \in H(X)} |f_{\ell_1 \dots \ell_k} - g_{\ell_1 \dots \ell_k}|.$$

The $\mathcal{L}_1$ measure offers an intuitive interpretation, for any given set of bins: if the two empirical distributions are completely separated (up to H), then $\mathcal{L}_1$= 1; if the distributions exactly coincide, which indicates perfect global balance, then $\mathcal{L}_1$= 0. In all other cases, $\mathcal{L}_1 \in (0, 1)$. If $\mathcal{L}_1$ = 0.7, then 30% of the area under the two histograms overlap. Thus, if we want to choose the best matching methodology, we need the $\mathcal{L}_1$statistic to be as low as it can be.

2.3.3 DIFFERENCE-IN-DIFFERENCES DESIGN

After matching, we run a difference-in-differences regression on these new matched samples. The DID equation is:

$$Y_{it} = \alpha + \beta_1 D_{1it} + \beta_2 FFVP_{it} + \beta_3 DID_{it} + \delta X'_{it} + \epsilon_{it}$$

where Y_{it} denotes the outcome variables (i.e., BMI z-score and BMI percentile) for individual i at period t; D_{1it} is a dummy variable for the different periods and takes the value of 1 if observations are from period 2 and a value of 0 otherwise ; $FFVP_{it}$ is a dummy variable that takes a value of 1 if the individual is in the treated group and a value of 0 otherwise; DID_{it} is the

DID interaction term (D_{1it}* $FFVP_{it}$); X'_{it} is a vector of control variables and ϵ_{it} is the error term. To test the robustness of the results, we run the DID regression using both fixed effects and random effects panel estimation and using matched samples.

2.4 RESULTS

Before discussing the main results, we first need to compare the estimates of imbalance test from each matching method. These are reported in Table 2. Note that the lower the $\mathcal{L}_1$ statistic, the more similar are the treatment and the control groups on average, which also indicates that the control and treatment samples are better matched. Results depicted in table 2 indicate that if we use nearest neighbor matching with 3 neighbors, the $\mathcal{L}_1$ statistic is 0.994 and the number of observations in the control group is 3,079. This provides a baseline reference for the analysis, which we can use as a point of comparison between matching solutions. As expected, the number of observations in the control group shrinks to 2,097 when using nearest neighbor matching with 2 neighbors. When we use Mahalanobis matching without caliper, the $\mathcal{L}_1$ statistic further falls to 0.903 and the number of observations in the control group declines to 811. The number of observations in the treatment group is 1,116, under each of these matching strategies.

Once we add a caliper 0.1 to the Mahalanobis matching algorithm, the matching becomes stricter. The $\mathcal{L}_1$ statistic becomes 0.620 and the number of observations in the control group falls to 206 while the number of observations in the treatment group falls to 266. This means that the algorithm could not find matches within the control group for the remaining treated observations. If we reduce the value of the caliper for Mahalanobis matching, the matching becomes even stricter and the $\mathcal{L}_1$ statistic becomes 0.605 for a caliper of 0.075 and 0.532 for a caliper of 0.05 with the number of observation being further reduced. Finally, when we use coarsened exact matching, the $\mathcal{L}_1$ statistic becomes 0.536 and the number of observations in the control and treated groups are 167 and 157, respectively.

While the use of stricter matching routines significantly decreased the number of observations in both the control and treatment groups, the resulting matches still include students widely distributed across different schools. In the sample after CEM, the individuals in the treatment group come from 13 schools (out of total of 14 FFVP participating schools) while the individuals in control group come from 78 schools. The same results are found in the matched groups using the Mahalanobis matching technique.

To further check the balance in the covariates, we also report in Table 2 the comparison of descriptive statistics of the variables for the control and treatment groups for each matching method as well as the results of tests for the differences between the control and treatment groups[8]. When less strict matching methods are used, there are important differences between the treatment and control groups in mean values for several of the individual and neighborhood controls. With the strict CEM method, the average values of these variables are much closer in the treated and control groups. Since these variables, especially the income measures are potentially important determinants of FFVP school participation, reducing the gap in these variables between the treated and control groups can also reduce selection bias issues. The results from the imbalance test suggest that the coarsened exact matching (CEM) and the Mahalanobis matching with caliper 0.05 provide the best balance between the control and treated groups. But as we mentioned before, the CEM always provides more precise matching than PSM when the list of variables used in the matching includes continuous variables. Hence, we rely

[8] We used proportion test for binary variables and t-test for other variables.

more on CEM but also report results of the DID matching panel estimates using the PSM methods for comparative purposes.

The estimates of our panel DID models are exhibited in Table 3[9]. Bertrand, Duflo, and Mullainathan (2004) showed that conventional standard errors often severely understate the standard deviation of the estimators in a DID framework. For this reason we use robust standard errors, clustered at the school level, introduced by Cameron, Gelbach and Miller (2011). DID estimations using matched samples based on the nearest neighbor and Mahalanobis matching with no caliper algorithms are positive but not statistically significant in both the fixed effects and random effects DID models. As we previously mentioned, selection bias is likely in these weakly matched samples. The estimates of the imbalance test for these three matching methods are quite close, and so it is not surprising that the coefficient of each DID interaction term is similar across these matching strategies.

In contrast, the DID estimates using the Mahalanobis matching with the caliper and from CEM are different. When using the matched samples from the Mahalanobis matching with the 0.1 caliper, the DID coefficient shows an effect on BMI z-score of -0.054 in the fixed effects model and -0.045 in the random effects model. These are, however, not significantly different from zero. With a 0.075 or 0.05 caliper, the DID coefficient is still negative, larger in magnitude, and insignificant. There was an important change in the coefficients from positive values to negative values with the reduction of the $\mathcal{L}_1$ statistic. With the CEM sample, the coefficients become -0.15 in the fixed effects model and -0.139 in the random effects model and both are now statistically significant at the 0.05 level.

[9] The complete set of estimates is available from the authors upon request.

Table 2. Comparison of Descriptive Statistics (Mean Values Crossing Periods) of Control and Treatment Group for Different Matching Methods

	Nearest-neighbor(3)		**Nearest-neighbor (2)**	**Mahalanobis matching**	**Mahalanobis matching with caliper 0.1**		**Mahalanobis matching with caliper 0.075**		**Mahalanobis matching with caliper 0.05**		**Coarsened exact matching**	
Balance Test ($\mathcal{L}_1$ statistic)	0.994		0.994	0.903	0.620		0.605		0.532		0.536	
Variable	*Control*	*Treated*	*Control*	*Control*	*Control*	*Treated*	*Control*	*Treated*	*Control*	*Treated*	*Control*	*Treated*
Number of Observations	N=3,079	N=1,116	N=2,097	N=811	N=206	N= 266	N=197	N= 254	N=180	N=229	N=167	N=157
Age	102.4**	100.9	102.5**	107.3***	104.5	103.9	104.5	104.1	104.7	104.2	102.8	104.1
Black	0.162	0.148	0.156	0.136	0.039	0.034	0.041	0.035	0.039	0.030	0.041	0.038
Hispanic	0.100**	0.120	0.099*	0.093*	0.027	0.017	0.028	0.018	0.030	0.020	0.029	0.022
Male	0.533	0.541	0.543	0.532	0.553	0.571	0.553	0.566	0.527	0.537	0.544	0.541
Free	0.519	0.510	0.525	0.477	0.367	0.333	0.360	0.329	0.350	0.316	0.323	0.324
Reduced	0.095*	0.108	0.092**	0.124	0.073	0.070	0.071	0.067	0.067	0.058	0.060	0.057
Urban	0.563	0.572	0.563	0.560	0.436*	0.388	0.431*	0.377	0.433*	0.362	0.440*	0.370
Lowaccess	0.204	0.200	0.200	0.239**	0.211	0.171	0.205	0.165	0.205	0.168.	0.221	0.184
Singlemother_prp	0.260	0.256	0.257	0.261	0.208	0.195	0.201	0.188	0.200	0.185	0.204	0.186
Highschool_prp	0.368	0.368	0.369	0.369	0.372	0.364	0.371	0.364	0.370	0.364	0.374	0.368
Somecollege_prp	0.268	0.266	0.268	0.272**	0.270	0.269	0.269	0.268	0.269	0.267	0.273	0.267
Collegeplus_prp	0.168	0.167	0.167	0.169	0.182*	0.199	0.183	0.197	0.187	0.201	0.179	0.195
Incomebelowpoverty	0.201***	0.212	0.202*	0.195***	0.169	0.162	0.165	0.162	0.168	0.162	0.168	0.166
Workingmother_prp	0.258	0.251	0.255	0.244	0.207	0.193	0.202	0.188	0.201	0.186	0.201	0.185
Novehicle_prp	0.061	0.060	0.061	0.061	0.055	0.054.	0.055	0.054	0.056	0.054	0.051	0.052
Medhousevalue	89.63***	86.28	89.45**	93.97***	100.96	100.13	100.83	99.90	101.89	101.56	101.6	101.4

Note: The samples of observations in treatment group for nearest-neighbor matching with 3 and 2 neighbors and Mahalanobis matching without caliper are same so we just report it once. *, **, *** denote that there are significant differences between control and treatment group for mean values of variables by t-test at the 0.1, 0.05, and 0.01 levels, respectively

Table 3. The Comparison of Results among Different Matching Methods

Matching Method	**Coefficient of DID (Fixed effects)**		**Coefficient of DID (Random effects)**	
	Bmi z-score	*Bmi percentile*	*Bmi z-score*	*Bmi percentile*
Nearest-neighbor matching with 3 neighbors	0.150	0.034	0.146	0.033
Nearest-neighbor matching with 2 neighbors	0.146	0.034	0.143	0.032
Mahalanobis matching without caliper	0.131	0.028	0.146	0.031
Mahalanobis matching with caliper 0.1	-0.054	-0.019	-0.045	-0.016
Mahalanobis matching with caliper 0.075	-0.064	-0.021	-0.055	-0.019
Mahalanobis matching with caliper 0.05	-0.082	-0.026	-0.072	-0.024
Coarsened exact matching	-0.150**	-0.038**	-0.139**	-0.037*

Note: *, **, *** denote significance at the 0.1, 0.05, and 0.01 levels, respectively

To test the robustness of our findings, we also ran our models using BMI percentiles as the outcome measure, in addition to the BMI z-score (also in Table 3). Results are similar to those discussed above. When using matched samples from nearest neighbor matching and Mahalanobis matching without caliper, the DID coefficients of the FFVP effect are always positive and not significant. However, when using matched samples from Mahalanobis matching with caliper 0.1, 0.075, and 0.05, the DID coefficient is negative but not statistically significant. When using the CEM matched sample, however, the coefficient is negative and significant at the 0.05 level for fixed effects and 0.10 level for random effects. The magnitude of the FFVP effect is robust across the fixed effects and random effects DID models (i.e., -0.038 for fixed effects and -0.037 for random effects), which suggests that those who participate in the FFVP have 3.8% (for fixed effects) and 3.7% (for random effects) lower BMI percentiles than those who do not participate in the FFVP. Given this finding and those of other past studies discussed previously suggesting the generally positive effects of FFVP participation on students' fruit and vegetable consumption, the FFVP program seems like a promising way of improving the diet and reducing childhood obesity among elementary school children, especially considering that the cost for each student in participating schools has been estimated to be only 50-75 dollars per year

REFERENCES

1. Anderson, P.M., K.F. Butcher, and P.B. Levine. 2003a. "Maternal employment and overweight children." *Journal of Health Economics,* 22(3): 477-504.

2. Angelucci, M., and O. Attanasio. 2013. "The demand for food of poor urban Mexican households: understanding policy impacts using structural models". *American Economic Journal: Economic Policy*, 5(1):146-178.

3. Bartlett S., L. Olsho, J. Klerman, K.L. Patlan, M. Blocklin, P. Connor, K. Webb, L. Ritchie, P. Wakimoto, P. Crawford. 2013. *Evaluation of the fresh fruit and vegetable program (FFVP).* Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service. Available at: http://www.fns.usda.gov/Ora/menu/Published/CNP/FILES/FFVP.pdf

4. Becker, S.O., and A. Ichino. 2002. "Estimation of average treatment effects based on propensity scores." *The Stata Journal*, 2(4): 358-377.

5. Bertrand, M., E. Duflo, and S. Mullainathan. 2004. "How much should we trust differences-in-differences estimates?", *The Quarterly Journal of Economics*, 119(1): 249-275.

6. Blackwell, M., S. M. Iacus, G. King, and G. Porro. 2009. "CEM: coarsened exact matching in stata." *Stata Journal*, 9(4): 524-546.

7. Boukhris, T. T. 2007. “A public response to childhood obesity: evaluating the fresh fruit and vegetable program in Texas schools.” PhD dissertation, Texas State University.

8. Cameron, A. C., J.B. Gelbach, and D. L. Miller. 2011. "Robust inference with multi-way clustering." *Journal of Business & Economic Statistics*, 29(2): 238-249.

9. Campbell, B.L., R. M. Nayga, J. L. Park, and A. Silva. 2011."Does the national school lunch program improve children’s dietary outcomes?" *American Journal of Agricultural Economics*, 93(4): 1099-1130.

10. Casey, P.H., K. Szeto, S. Lensing, M. Bogle, and J. Weber. 2001."Children in food-insufficient, low-income families: prevalence, health, and nutrition status." *Archives of Pediatrics & Adolescent Medicine*, 155(4): 508.

11. Center for Disease Control and Prevention. 2013. *About BMI for children and teens.* Available at: http://www.cdc.gov/healthyweight/assessing/bmi/childrens_bmi/about_childrens_bmi.html

12. Davis, E.M., K.W. Cullen, K.B. Watson, M. Konarik, and J. Radcliffe. 2009. "A fresh fruit and vegetable program improves high school students' consumption of fresh produce." *J Am Diet Assoc*, 109(7): 1227-31.

13. Epstein, L.H., C.C. Gordy, H.A. Raynor, M. Beddome, C. K. Kilanowski, and R. Paluch. 2001. "Increasing fruit and vegetable intake and decreasing fat and sugar intake in families at risk for childhood obesity." *Obesity Research*, 9(3): 171-178.

14. Haynes-Maslow, L.E. Parsons, S.B. Wheeler, L.A. Leone, 2013. "A qaualitative study of perceived barriers to fruit and vegetable consumption among low-income populations, North Carolina, 2011." *Preventing Chronic Disease* 10.

15. Heckman, J.J., H. Ichimura, and P.E. Todd. 1997. "Matching as an econometric evaluation estimator: evidence from evaluating a job training programme." *The Review of Economic Studies*, 64(4): 605-654.

16. Iacus, S.M., G. King, and G. Porro. 2012. "Causal inference without balance checking: coarsened exact matching." *Political Analysis*, 20(1): 1-24.

17. Iacus, S.M., G. King, and G. Porro. 2011. "Multivariate matching methods that are monotonic imbalance bounding." *Journal of the American Statistical Association*, 106(493): 345-361.

18. Islam, A. 2011. "Medium-and long-term participation in microcredit: An evaluation using a new panel dataset from Bangladesh." *American Journal of Agricultural Economics*, 93(3): 847-866.

19. Jamelske, E., L. A. Bica, D. J. McCarty, and A. Meinen. 2008. "Preliminary findings from an evaluation of the USDA fresh fruit and vegetable program in Wisconsin schools." *Wisconsin Medical Journal (WMJ)*, 107(5): 225.

20. Justus, M.B., K.W. Ryan, J. Rockenbach, C. Katterapalli, and P. Card-Higginson. 2007. "Lessons learned while implementing a legislated school policy: body mass index assessments among Arkansas's public school students." *Journal of School Health*, 77(10): 706-713.

21. Nayga, Rodolfo M. 2000. "Schooling, health knowledge and obesity." *Applied Economics*, 32(7): 815-822.

22. Ohri-Vachaspati, P., L. Turner, and F.J. Chaloupka. 2012. "Fresh fruit and vegetable program participation in elementary schools in the United States and availability of fruits and vegetables in school lunch meals." *Journal of the Academy of Nutrition and Dietetics*, 112(6): 921-926.

23. Rosenbaum, P.R., and D.B. Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika,* 70(1): 41-55.

52. Rosenbaum, P.R., and D.B. Rubin. 1985. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score." *The American Statistician,* 39(1): 33-38.

24. Singh, G.K., M. Siahpush, and M.D. Kogan. 2010a. "Rising social inequalities in US childhood obesity, 2003–2007." *Annals of Epidemiology,* 20(1): 40-52

25. Singh, G.K., M. Siahpush, and M.D. Kogan. 2010b. "Neighborhood socioeconomic conditions, built environments, and childhood obesity." *Health Affairs*, 29(3): 503-512.

26. United States Department of Agriculture, Food and Nutrition Service. 2010. *Handbook for fresh fruit and vegetable program.* Available at: http://www.fns.usda.gov/cnd/ffvp/handbook.pdf.

27. Wang, Y. 2001. "Cross-national comparison of childhood obesity: the epidemic and the relationship between obesity and socioeconomic status." *International Journal of Epidemiology,* 30(5): 1129-1136.

APPENDIX 1: THE STATEMENT OF RESEARCH SPONSORSHIP

This research is funded by the Agriculture and Food Research Initiative of the USDA National Institute of Food and Agriculture, grant number 2011-68001-30014. This work is also partly supported by the National Research Foundation of Korea (NRF-2011-330-B00074). Dr. Rodolfo M. Nayga Jr. is also affiliated with Department of Food and Resource Economics, Korea University and Norwegian Agricultural Economics Research Institute

UNIVERSITY OF ARKANSAS
DIVISION OF AGRICULTURE

Department of
Agricultural Economics an Agribusiness

UNIVERSITY OF ARKANSAS
DALE BUMPERS COLLEGE
OF AGRICULTURAL, FOOD & LIFE SCIENCES

217 Agriculture Building, University of Arkansas, Fayetteville, AR 72701-1201
479-575-2256 • Fax 479-575-5306

THE STATEMENT OF MUTIPLE AUTHORS

The chapter titled "The Effect of The Fresh Fruit and Vegetable Program on Childhood Obesity" has multiple authors. The student, Yiwei Qian, is the first author and completed at least 51% of the work for this chapter.

Dr. Rodolfo M. Nayga Jr. Dissertation Director (Co-chair)	Date
Dr. Arya B. Gaduh Dissertation Director (Co-chair)	Date

3. THE EFFECT OF NEIGHBORHOOD FAST FOOD ON CHILDHOOD OBESITY

Authors: Yiwei Qian, Michael R. Thomsen, Rodolfo M. Nayga Jr, Heather L. Rouse

3.1 BACKGROUND AND LITERATURE REVIEW

The consumption of fast food has been discussed as one of the main factors contributing to the increasing rates of childhood obesity. According to Prentice and Jebb (2003), average energy density of menu items at fast food restaurants can be more than twice that recommended for healthy diets. Jeffery et al. (2006) surveyed 1033 Minnesota residents about their body height, weight and frequency of eating at restaurants and found that frequency of eating at fast food restaurants was positively associated with Body Mass Index (BMI). Similar results were reported in Pereira et al. (2003) and Duffey et al. (2007) for 18-30 year old adults.

Children and adolescents can be more vulnerable to the high energy content of fast food than adults because they have not developed cognitive dietary restraint habits. Eating fast food more frequently could reduce children's dietary quality in many ways (French et al., 2001) and could increase the risk of obesity (Niemeier et al, 2006). French et al. (2001) surveyed 4,746 students (7-12 grades) in Minnesota and found that the frequency of fast food restaurant use was positively related to intake of total energy, percent energy from fat, daily serving of soft drinks, cheeseburgers, french fries and pizza. It was negatively related to daily servings of fruit, vegetables and milk. They also found that fast food frequency is positively related to the availability of unhealthy foods in the home and negatively related to the student's own and perceived maternal and peer concerns about healthy eating. Bowman et al. (2004) investigated over 6,000 children aged 4 to 19 years and found similar results. Niemeier et al (2006) tracked nearly 10,000 adolescents into adulthood and found that the greater number of days of fast food consumption when aged 11-21 years, the higher was the BMI z-score at ages 18-27 years.

Thompson et al. (2003) investigated 101 healthy eight to twelve year old girls and found that those who ate fast food twice a week or more were likely to increase their relative BMI over time.

Considering the potential harm that fast food consumption could render to children's dietary quality and obesity level, our main goal in this paper is to determine whether the availability of neighborhood fast food restaurants is a significant driver of childhood obesity outcomes. A number of studies have attempted to estimate the causal effect of the density of fast food restaurants on obesity outcomes among children. For example, Currie et al. (2010) found that a fast food restaurant within 0.1 miles of a school resulted in 5.2 percent increase in obesity rates among ninth graders. Davis and Carpenter (2009) investigated geocoded data on over 500,000 youths and found that students with fast food restaurant near (within 1.5 miles) their schools consumed fewer fruits and vegetables and more soda, and were more likely to be overweight or obese than those who had no exposure to fast food restaurants.

Some other recent studies have acknowledged that fast food availability is endogenous with obesity outcomes (Chen et al., 2009; Dunn, 2010; Dunn, Sharkey and Horel, 2012; Anderson and Matsa, 2011; Alviola et al., 2013). The concern is that the distribution of fast food restaurants and consumers' choice of residential location is determined by preferences and behaviors that also affect obesity outcomes. Therefore, these studies used instrumental variable models to solve the endogeneity problem. Chen, Florax and Snyder (2013) examined the effect of the density of neighborhood fast food restaurants and grocery stores surrounding residents of Marion County, Indiana on individual BMI. They used the amount of land that is zoned non-residential and arterial roads as instrumental variables and found that BMI was positively related to fast food density.

Dunn (2010), Anderson and Matsa (2011), Dunn, Sharkey and Horel (2012) and Alviola et al. (2013), use highway proximity as an instrumental variable to assess the effect of fast food restaurants. The argument is that fast food restaurants tend to cluster near highways to capture demand from travelers and so the presence of highways substantially increases the accessibility of fast food. Anderson and Matsa (2011) found no evidence linking fast food restaurants to obesity level among adults. However, Dunn (2010) found significant and positive effect of fast food proximity on BMI but only among female and minority subgroups. Similarly, Dunn, Sharkey and Horel (2012) found that obesity rate of minorities are more likely to be affected by fast food availability. Alviola et al. (2013) focused on the effect of fast food restaurants surrounding a school on school level obesity rates in Arkansas and found that an additional fast food restaurant within a mile from a school resulted in an increase of 1.23 percentile points in school obesity rates.

Like Alviola et al. (2013), we examine Arkansas public schoolchildren. However, our study differs in that we examine individual-level BMI z-scores as opposed to aggregate school-level obesity rates. Our study is similar to Anderson and Matsa (2011) and Dunn (2010) in that we examine the role of fast foods around the residences but different in that our focus is on children. Another important difference is that the Arkansas BMI data provide rooftop level geographic precision and so we are able to measure fast food restaurant counts in the microenvironment surrounding the children's actual residences. We follow each of these earlier studies by instrumenting fast food density by a measure of highway proximity.

The next section discusses the empirical strategy we used to identify the regression model. Section 3.3 describes the data sources and the variables used in the analysis. Section 3.4 discusses the validity of instrumental variables and section 5 presents the results and concludes.

3.2 MODEL SPECIFICATION

The difference in studying childhood obesity versus adult obesity is that children's food choices and preferences are largely dependent on parental decisions (Anderson, Butcher and Levine, 2003b), while those of adults are not. Parents, based on their preferences and work status among others, can choose where to live (e.g., in areas with lower or higher fast food density), and so their children's BMI could be influenced by this decision. Fast food restaurants may also geographically position themselves based on characteristics of nearby consumers, which can also be correlated with obesity outcome (Dunn, 2010; Anderson and Matsa, 2011).

Dunn (2010) argued that there may be multiple directions of the endogeneity bias. On one hand, fast food establishments could choose to locate in areas where consumers do not generally care about dietary health, and hence obesity rates may already be higher among this group of consumers. On the other hand, fast-food restaurants may tend to target those who have high opportunity cost of food preparation at home. These consumers may also have higher incomes, which have been shown to be associated with low obesity rates (Casey et al., 2001; Singh, Siahpush and Kogan, 2010a). In our case, children who have lower BMI outcomes may have parents who care about their health and tend to avoid fast food meals. These preferences could conceivably affect the choice of residential location. Therefore, since numerous unobservable factors could be correlated with BMI, residential location and fast food distribution, directly conducting a regression between BMI outcome and the density of fast food could render biased estimates. Given the endogeneity issues involved, we opted to use an instrumental variable approach and panel data estimation to tackle this problem.

Our basic model can be represented as:

$$Y_{it} = \beta_1 FF_{it} + \beta_2 X'_{it} + \lambda_i + \epsilon_{it}$$

where Y_{it} denotes the BMI z-score for child i at period t; FF_{it} denotes the counts of fast food restaurants within a given radial distance of the child's residence (i.e., within 0.5 miles and 1 mile radius); X'_{it} is a vector of control variables; λ_i is the latent time-invariant variables and ϵ_{it} is the error term.

Since we are concerned that the density of fast food restaurants is endogenous, we estimated the first stage equation involving an instrumental variable. The equation is:

$$FF_{it} = \gamma_1 X'_{it} + \gamma_2 Z_{it} + \mu_{it}$$

where Z_{it} denotes the instrument we are using, which is the distance between the child's residence and the nearest US or interstate highway and μ_{it} is the error term. We discuss the validity of this instrument in a subsequent section.

3.3 DATA

Our data come from several different sources. First, we use the Arkansas BMI data from 2004 to 2010. This is from same dataset from the Arkansas Center for Health Improvement (ACHI) in Section 2. In addition to the BMI z-scores, the Arkansas BMI data provides information on the child's gender, race, ethnicity and whether the child was eligible for the free or reduced lunch program[1]. Our dataset include all grades from kindergarten to 10th grade for 2004-2007. Again, after 2007, BMI screenings switched to a biennial schedule with only the

[1] Free and reduced lunch status is used to define income level of households. To receive a free meal, household income must be below 130% of the Federal poverty threshold, and to receive a reduced-price meal, household income must be below 185% of the Federal poverty threshold, as defined by the U.S. Office of Management and Budget.

even grades being measured in any given year. For this reason, we have only even-grade measurements for 2008-2010 because ACHI only measured students on even grades after 2007.

Second, we purchased geo-coded business lists from Dun and Bradstreet (D&B). To define fast-food restaurants, we started with all establishments with a standard industrial classification (SIC) code of 5812 "Eating Places" and then removed full-service restaurants based on six and eight-digit SIC codes, if available. Otherwise, we identified fast-food restaurants using the company name or, in the case of chain or franchise restaurants, the trade name. When the type of establishment remained in doubt, we used internet searches and identified fast food restaurants based on website information (e.g., menus), customer ratings, or street-view images in the Google search engine. Fast-food restaurants, as used in our study, include the major hamburger chains and drive-in restaurants (e.g. McDonalds, Burger King, Wendy's), dairy stores with large fast-food menus (e.g., Dairy Queen), quick-service taco formats (e.g., Taco Bell), and fried chicken restaurants (e.g., KFC, Chick-Fil-A). Our definition of fast-food establishments excludes specialty stores such as ice-cream parlors not selling other fast foods (e.g., Baskin-Robbins), coffee shops (e.g. Starbucks), and donut shops (e.g. Krispy Kream). We obtained archival business lists from D&B so that our restaurant data represent establishments as of December for each year for which we have BMI data.

We overlaid the fast food restaurant coordinates onto the residential coordinates of students in the BMI data. By doing this, we were then able to count the number of fast food restaurants within 0.5 miles and 1 mile for each student's residence. To control for other dimensions of the commercial food environment in our models, we also counted the number of convenience stores within a 1 mile radius from the child's residence and developed a "low healthy-store access" variable to indicate whether the neighborhood where the child resides has low access or not to large grocery stores with fresh produce and other healthier food options. Finally, we overlaid residential coordinates onto highway maps from the Arkansas Transportation Department to calculate the distance from a child's residence to the nearest US highway.

Table 1 presents the variable names and definitions used in our study and the descriptive statistics. The total number of observations in our panel data is 1,246,949. The average BMI z-score is 0.707, while the average number of fast food restaurant within a half mile radius is 0.444 and within a one-mile radius is 1.675. Additionally, 21.2% and 45.3% of children reside in areas that have at least one fast food restaurant within half a mile and one mile, respectively. The average distance between residence and nearest highway is 2.105 miles.

3.3 INSTRUMENT VALIDITY

As mentioned above, we used the distance of residence from the nearest US highway as the instrumental variable to identify our model. As previously mentioned, a number of studies have employed measures of proximity to highways to identify the relationship between fast food establishments and obesity outcomes (i.e., Dunn, 2010; Anderson and Matsa, 2011; Alviola et al., 2013). The rationale for the use of this instrument is that fast food restaurants tend to cluster near highways to target travelling customers. Hence, the presence of highways substantially increases the accessibility of fast food. Both Dunn (2010) and Anderson and Matsa (2011) assessed the effect of residing close to highways on individuals' behaviors; for example, physical activity, fruit and vegetable consumption and other BMI related factors. They concluded that the effect is insignificant or, in the case of Dunn (2010), that the effect is statistically significant but of very small magnitude. Additionally, Anderson and Matsa (2011) evaluated the nature of the

Table 1. Description and Descriptive Statistics of Variables in the Study

Variables	**Description**	**Mean**
Outcome Variable		
BMI z-score	Individuals' BMI z-score	0.707 (1.060)
Instrumental Variable		
Nearest highway	The distance between individuals ‘ residence and nearest major highway (in miles)	2.105 (3.249)
Control Variables		
Individual level		
Fastfood_Half	Number of fast food restaurants within 0.5 mile radius from individual's' residence	0.444 (1.050)
Fastfood_One	Number of fast food restaurants within 1 mile radius from individual's' residence	1.675 (2.384)
Age	Age of student in months	129.9 (37.55)
White	Binary indicator (if individual is White then =1, 0 otherwise)	0.674 (0.468)
Black	Binary indicator (if individual is Black then =1, 0 otherwise)	0.225 (0.417)
Hispanic	Binary indicator (if individual is Hispanic then =1, 0 otherwise)	0.078 (0.267)
Female	Binary indicator (if individual is female then =1, 0 otherwise)	0.486 (0.500)
Free	Binary indicator (if individual participated in free lunch then =1, 0 otherwise)	0.441 (0.496)
Reduced	Binary indicator (if individual participated in reduced lunch then =1, 0 otherwise)	0.097 (0.296)
Urban	Binary indicator (if individual lived in urban area =1, 0 otherwise)	0.617 (0.486)
Lowaccess*	Binary indicator that describes the accessibility to large grocery stores. It takes the value of one for urban students living more than one-mile from a large grocery store and for rural students living more than 10 miles from a large grocery store.	0.313 (0.464)
Convenient_one	Number of convenience stores within 1 mile radius from individual's' residence	2.10 (2.68)
Year 2005-2010	Binary variables for the year of BMI measurement(Year 2004	

	is baseline)

Note: Standard errors appear in parenthesis.
* The low access area criteria follow the USDA/ERS Food Desert Locator, website: http://www.ers.usda.gov/data/fooddesert/about.html

connection between highway proximity and fast food availability using a survey and showed that highway proximity is highly correlated with frequency of fast food restaurant consumption.

We used distance to the nearest major highway to instrument fast food density (Dunn, Sharkey & Horel, 2012; Alviola et al., 2013). By major highway, we specifically mean US highways or interstate highways. Alviola et al. (2013) reported that in Arkansas, the interstate highway system does not serve many portions of the state as US highways. According to their 2008 data, only a few fast food restaurants were located close to interstate highway while many more were located close to US highways. In the case of Arkansas, US highways are important connecting routes and have a much more significant linkage with fast food availability.

We also used an OLS balancing test following Dunn, Sharkey and Horel (2012) and Alviola et al. (2013) to further assess the validity of our instrument by testing whether our IV estimates are driven by the difference of characteristics of individuals and the neighborhoods where they reside. Table 2 presents the results of this balancing test where we regressed the explanatory variables with the instrumental variable. While the control variables show statistically significant association with the instrument, the magnitudes of the coefficients are too small to explain the preceding results for income level and grocery store availability .However, for urban status and the number of convenience stores within one mile radius, the magnitudes of the coefficients are large, which is logical because highways connect cities and many convenience stores with gas stations are located near highways (Dunn, 2010). Dunn (2010) also

Table 2. OLS Balancing Test Regressions (N=1,246,949)

Dependent Variables	Nearest highway
Free	-0.011***
	(0.0001)
Reduced	0.0023***
	(0.0006)
Urban	-0.096***
	(0.0001)
Lowaccess	0.0007***
	(0.0001)
Convenient_one	-0.259***
	(0.001)

Note: Each estimate is from a different OLS regression. The explanatory variable is the distance from the child's residence to nearest highway. Free, Reduced, Urban and Lowaccess are binary variables so we report the marginal effect from a logistic regression. For Convenient_one, we report the OLS estimate.
Robust standard errors appear in parenthesis.
*** Statistical significance at the 0.01 level

argued that the association between convenience stores and interstate exits does not explain the positive association between obesity and fast-food availability.

To further assess instrument validity, we checked whether there is an association between current BMI outcome and future distances to the nearest highway. If BMI is one of the determinants of households' location and households can choose how close to live to a highway, it is possible that BMI may be correlated with future distance from the residence to the nearest highway. Otherwise, we expect that future households' location should not be affected by BMI outcomes. Table 3 presents the results of a regression of the distance from residence to nearest highway at period *t+2* on the BMI z-score at period *t* for entire sample and the subsample of children that moved household location between period *t* and *t+2*. Both coefficients are statistically insignificant.

3.4 RESULTS

Table 4 presents the first stage estimation results. As expected, the density of fast food restaurants decreases as the distance to nearest highway increases and this coefficient is significant at the 1% level for both the half mile and one mile distances. The first stage F statistics of excluded instrument are large and signify that our instrument is not weak.

Table 5 reports the results of the estimates from fixed effects and random effects models without the instrumental variable along with the results from the IV-fixed effects model. For restaurant densities within half a mile, the coefficient of fast food availability in the fixed effects model is insignificant while the corresponding coefficient in the random effects model is significant at the 5% level and positive. Importantly, the effect of number of fast food restaurants within half a mile from residence on BMI z-score is positive and significant at the 1% level in the IV fixed effects model. Moreover, the IV model shows a much larger effect (0.091) in comparison to the other models (0.001). This result suggests that an additional fast food restaurant within a half a mile of the residence increases a child's BMI *z*-score by 0.091 standard deviations[2]. To place this in context, a child located at the sample mean BMI z-score would increase by 2.7 BMI percentile points[3].

---

[2] Considering that the IV fixed effects estimates could have been driven by the number of convenience stores based on the balancing test results, we re-did all the analyses after dropping

Table 3. Results of Instrumental Invalidity Test

	BMI z-score
Nearest highway 2 years ahead	-.0030
(N=513,585)	(.005)
Nearest highway 2 years ahead for movers	-.0053
(N=239,349)	(.0106)

Note: Each estimate is from a different model. The dependent variable is the future distance from the residence to nearest highway. The first estimate is from the entire sample that was observed during both period t and t+2. The second is for a subsample that moved between period t and t + 2. The explanatory variables include BMI z-score (reported) and other variables mentioned in Table 1 (not reported). Robust standard errors appear in parenthesis.

For fast food availability within one mile, the results are similar, but the importance of one additional restaurant within this larger radius is smaller. Both the fixed effects model and the random effects model show positive and significant fast food effects, and again the coefficient in the IV fixed effects model is much larger. The IV fixed effects model estimate suggests that if the number of fast food restaurants within one mile increased by one, children's BMI z-score will increase by 0.035 standard deviations, which equals 1.13 BMI percentile point increase for the children located at the mean BMI z-score. The fact that the magnitude of the effect is smaller in one mile versus the case of half a mile is intuitive and reasonable because accessing fast food restaurants at longer distances could lead to higher transportation costs, which can then reduce consumption demand. Furthermore an additional restaurant within a mile radius is likely a less prominent feature of the built environment surrounding the residence than a restaurant within a half mile.

---

the number of convenience stores variable. The results show that the magnitudes of effect of fast food availability slightly change but the statistical significances are very robust.

[3] BMI percentile is a value of a cumulative probability distribution of BMI z-score. In our study, average BMI percentile is 0.758.

Table 4. First Stage Estimates (N=1,246,949)

	Fastfood_Half	Fastfood_One
	-0.022***	-0.058***
Nearest highway	(0.0008)	(0.0017)
	0.001	0.001
Age	(0.0004)	(0.001)
	0.011***	0.027***
Free	(0.003)	(0.006)
	0.01***	0.035***
Reduced	(0.003)	(0.006)
	0.411***	1.525***
Urban	(0.006)	(0.012)
Lowaccess	-0.321***	-1.071***
	(0.004)	(0.008)
	0.113***	0.409***
Convenient_one	(0.001)	(0.001)
	-0.004	0.002
Year 2005	(0.004)	(0.008)
	-0.023***	-0.045***
Year 2006	(0.008)	(0.016)
	0.018	0.106***
Year 2007	(0.013)	(0.024)
	0.041**	0.198***
Year 2008	(0.017)	(0.032)
	0.023	0.129***
Year 2009	(0.022)	(0.041)
	-0.010	0.014
Year 2010	(0.027)	(0.050)
F test of excluded instruments (1, 878909)	709.89	1117.34

Note: Robust standard errors appear in parenthesis.
*** Statistical significance at the 0.01 level
** Statistical significance at the 0.05 level
* Statistical significance at the 0.1 level

To check if there is heterogeneity in the results on the effect of fast food availability across demographic groups, we re-estimated the IV fixed effects model across different sub-groups of children. Table 6 presents the IV fixed effects estimates for several subpopulation groups by gender, free and reduced lunch program status, student grade level, urban or rural residential status, and ethnicity. Again, all of the first stage F statistics for excluded instruments are large and are well above 10. The results show that density of fast food has significant and positive effect on females' BMI z-scores but not on males' BMI z-scores. Interestingly, fast food availability has no significant effect on BMI z-scores of children from lower-income families, i.e., those who were eligible for free or reduced lunch programs during each year for which we observe a BMI screening.[4] However, the effect is significant and positive for the higher-income children who were not eligible for the free and reduced lunch programs. The effect is significant and positive for white children but not for minority groups. The analysis by grade level shows that the effect is consistently significant and positive for students in the fourth through seventh grades.

Table 7 presents the IV fixed effects of individuals living in urban area and rural area. Since the distribution of location of households and fast food restaurants are quite different for urban and rural areas given that children living in urban areas tend to have more chance to get fast food than children in rural areas within a certain distance, we analyzed the effects of fast food availability within a quarter mile, half mile, and one mile for the urban group, and fast food availability within one mile, two miles and five miles for the rural group. The effect is significant and positive for those who live in rural areas but not for those who live in urban areas and these results are robust for the different distances examined.

[4] The subsamples by lunch status consist of children who did not change status over the study period. For example, these subsamples exclude children who qualified for free lunch in one year but not in another. Similarly, children that moved between urban and rural residences during the study period are excluded from the urban and rural subsamples.

Table 5.1. Fixed Effects, Random Effects and IV-Fixed Effects Estimates (N= 1,246,949) for Fast Food Availability within a Half Mile

	Fixed Effects	Random Effects	IV Fixed Effects
Fastfood_Half	0.001	0.001*	0.091***
	(0.001)	(0.001)	(0.021)
Age	0.00001	0.001***	-0.00002
	(0.0002)	(0.00003)	(0.0002)
Female		-0.048***	
		(0.003)	
White		0.030***	
		(0.007)	
Black		0.158***	
		(0.007)	
Hispanic		0.243***	
		(0.008)	
Free	0.004**	0.029***	0.004**
	(0.002)	(0.002)	(0.002)
Reduced	0.005**	0.023***	0.004**
	(0.002)	(0.002)	(0.002)
Urban	0.013***	-0.009***	-0.029***
	(0.003)	(0.002)	(0.010)
Lowaccess	-0.002	-0.002	0.027***
	(0.002)	(0.002)	(0.007)
Convenient_one	0.001**	0.002***	-0.010***
	(0.0004)	(0.0004)	(0.003)
Year 2005	0.013***	-0.001	0.013***
	(0.003)	(0.001)	(0.003)
Year 2006	0.020***	-0.010***	0.022***
	(0.005)	(0.001)	(0.005)
Year 2007	0.042***	-0.005***	0.040***
	(0.008)	(0.001)	(0.008)
Year 2008	0.058***	-0.006***	0.054***
	(0.011)	(0.002)	(0.011)
Year 2009	0.077***	-0.005**	0.074***
	(0.014)	(0.002)	(0.014)
Year 2010	0.103***	0.003	0.103***
	(0.017)	(0.002)	(0.017)
Intercept	0.654***	0.467***	
	(0.024)	(0.008)	

Note: Robust standard errors appear in parenthesis.
*** Statistical significance at the 0.01 level
** Statistical significance at the 0.05 level
* Statistical significance at the 0.1 level

Table 5.2 Fixed Effects, Random Effects and IV-Fixed Effects Estimates (N= 1,246,949) for Fast Food Availability within One Mile

	Fixed Effects	Random Effects	IV Fixed Effects
Fastfood_One	0.002***	0.001***	0.035***
	(0.0004)	(0.004)	(0.008)
Age	0.00001	0.001***	-0.00001
	(0.0002)	(0.00003)	(0.0002)
Female		-0.048***	
		(0.003)	
White		0.031***	
		(0.007)	
Black		0.158***	
		(0.007)	
Hispanic		0.244***	
		(0.008)	
Free	0.004**	0.029***	0.004**
	(0.002)	(0.002)	(0.002)
Reduced	0.005**	0.023***	0.004*
	(0.002)	(0.002)	(0.002)
Urban	0.011***	-0.010***	-0.045***
	(0.003)	(0.002)	(0.014)
Lowaccess	-0.001	-0.001	0.036***
	(0.002)	(0.002)	(0.010)
Convenient_one	0.0002	0.002	-0.014***
	(0.0004)	(0.0004)	(0.004)
Year 2005	0.013***	-0.001	0.013***
	(0.003)	(0.001)	(0.003)
Year 2006	0.020***	-0.009***	0.022***
	(0.005)	(0.001)	(0.005)
Year 2007	0.042***	-0.003***	0.038***
	(0.008)	(0.001)	(0.008)
Year 2008	0.058***	-0.005***	0.051***
	(0.011)	(0.002)	(0.011)
Year 2009	0.077***	-0.005**	0.072***
	(0.014)	(0.002)	(0.014)
Year 2010	0.103***	0.003	0.102***
	(0.017)	(0.002)	(0.017)
Intercept	0.654***	0.467***	
	(0.024)	(0.007)	

Note: Robust standard errors appear in parenthesis.
*** Statistical significance at the 0.01 level
** Statistical significance at the 0.05 level
* Statistical significance at the 0.1 level

Table 6. Estimates of IV Fixed Effects for Different Groups

	Fastfood_Half	Fastfood_One
Male (N= 640,767)	0.046	0.019
	(0.033)	(0.013)
Female(N= 606,181)	0.140***	0.052***
	(0.031)	(0.011)
Free or Reduced[a] (N= 402,480)	0.037	0.019
	(0.029)	(0.015)
Non Free and Reduced[b] (N= 441,991)	0.113**	0.028**
	(0.051)	(0.013)
Kindergraton-3 grades (N= 411,788)	0.093**	0.039**
	(0.040)	(0.016)
4-7 grades(N= 371,748)	0.143***	0.054***
	(0.041)	(0.015)
8-10 grades(N= 246,561)	0.077*	0.032*
	(0.042)	(0.018)
White(N= 841,908)	0.1146***	0.047***
	(0.037)	(0.012)
Black (N= 279,711)	-0.013	-0.007
	(0.018)	(0.009)
Hispanic (N= 91,237)	-0.005	-0.002
	(0.041)	(0.020)

Note: Each coefficient is from a different IV fixed effect model. F statistics of all first stage estimation are larger than 10. Robust standard errors appear in parenthesis.

*** Statistical significance at the 0.01 level

** Statistical significance at the 0.05 level

* Statistical significance at the 0.1 level

a. Included those who always participated free or reduced lunch program

b. Included those who never participated free nor reduced lunch program

Table 7. Estimates of IV Fixed Effects for Urban and Rural Groups

	Fastfood_Qtr	Fastfood_Half	Fastfood_One	Fastfood_Two	Fastfood_Five
Urban[a]	0.027	0.007	0.003		
(N= 696,593)	(0.033)	(0.008)	(0.003)		
Rural[b]			0.101**	0.029**	0.005**
(N= 405,896)			(0.051)	(0.015)	(0.002)

Note: Each column is a different IV fixed effect model. F statistics of all first stage estimation are larger than 10. Fastfood_Qtr, Fastfood_Half, Fastfood_One, Fastfood_Two, Fastfood_Five denote the number of fast food restaurants within quarter miles, half miles, one mile, two miles and five miles. Robust standard errors appear in parenthesis.

** Statistical significance at the 0.05 level

a. Included those who always lived in urban area

b. Included those who always lived in rural area

REFERENCES

1. Alviola, P.A., R.M. Nayga, M.R. Thomsen, D. Danforth, and J. Smartt. 2014. "The effect of fast-food restaurants on childhood obesity: a school level analysis." *Economics & Human Biology*, 12: 110-119

2. Anderson, P., Butcher, K., Levine, P., 2003b. Economic perspectives on childhood obesity. *Economic Perspectives*, 27(3): 30–34.

3. Anderson, M. L., and D. A. Matsa. 2011. "Are Restaurants Really Supersizing America?"*American Economic Journal: Applied Economics*, 3(1): 152-88.

4. Binkley, J. K., J. Eales, and M. Jekanowski. 2000. "The relation between dietary change and rising US obesity." *International journal of obesity*, 24(8): 1032-1039.

5. Bowman, S.A., S.L. Gortmaker, C.B. Ebbeling, M.A. Pereira, and D. S. Ludwig. 2004. "Effects of fast-food consumption on energy intake and diet quality among children in a national household survey." *Pediatrics*, 113(1): 112-118.

6. Chen, S., R. Florax, and S. Snyder. 2013. "Obesity and fast food in urban market: a new approach using geo-referenced Micro data", *Health Economics*, 22: 835-856

7. Currie, J., S. D Vigna, E Moretti, and V Pathania. 2010. "The effect of fast food restaurants on obesity and weight gain." *American Economic Journal: Economic Policy*: 32-63.

8. Davis, B., and C. Carpenter. 2009. "Proximity of fast-food restaurants to schools and adolescent obesity." *Journal Information* 99(3).

9. Duffey, K. J., P. Gordon-Larsen, D. R. Jacobs, O. D. Williams, and B. M. Popkin. 2007. "Differential associations of fast food and restaurant food consumption with 3-y change

in body mass index: the Coronary Artery Risk Development in Young Adults Study." *The American Journal of Clinical Nutrition,* 85(1): 201-208.

10. Dunn, R.A. 2010. "The effect of fast-food availability on obesity: an analysis by gender, race, and residential location." *American Journal of Agricultural Economics* 92(4): 1149-1164.

11. Dunn, R.A., J.R. Sharkey, and S. Horel. 2012. "The effect of fast-food availability on fast-food consumption and obesity among rural residents: an analysis by race/ethnicity." *Economics & Human Biology,* 10(1): 1-13.

12. French, S. A., L. Harnack, and R. W. Jeffery. 2000. "Fast food restaurant use among women in the Pound of Prevention study: dietary, behavioral and demographic correlates." *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity,* 24(10): 1353-1359.

13. French, S.A., M. Story, D. Neumark-Sztainer, J. A. Fulkerson, and P. Hannan. 2001. "Fast food restaurant use among adolescents: associations with nutrient intake, food choices and behavioral and psychosocial variables." *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity,* 25(12): 1823-1833.

14. Jeffery, R.W., J. Baxter, M. McGuire, and J. Linde. 2006. "Are fast food restaurants an environmental risk factor for obesity?." *International Journal of Behavioral Nutrition and Physical Activity,* 3(1): 2.

15. Niemeier, H.M., H.A. Raynor, E.E. Lloyd-Richardson, M.L. Rogers, and R.R. Wing. 2006. "Fast food consumption and breakfast skipping: predictors of weight gain from adolescence to adulthood in a nationally representative sample." *Journal of Adolescent Health,* 39(6): 842-849.

16. Pereira, M.A., A.I. Kartashov, C.B. Ebbeling, L.V. Horn, M. L. Slattery, D. R. Jacobs Jr, and D.S. Ludwig. 2005. "Fast-food habits, weight gain, and insulin resistance (the CARDIA study): 15-year prospective analysis." *The Lancet,* 365 (9453): 36-42.

17. Prentice, A.M., and S.A. Jebb. 2003. "Fast foods, energy density and obesity: a possible mechanistic link." *Obesity reviews,* 4(4):187-194.

18. Singh, G.K., M. Siahpush, and M.D. Kogan. 2010a. "Rising social inequalities in US childhood obesity, 2003–2007." *Annals of Epidemiology,* 20(1): 40-52

19. Singh, G.K., M. Siahpush, and M.D. Kogan. 2010b. "Neighborhood socioeconomic conditions, built environments, and childhood obesity." *Health Affairs,* 29(3): 503-512.

20. Thompson, O.M., C. Ballew, K. Resnicow, A. Must, L. G. Bandini, H. D. W. H. Cyr, and W. H. Dietz. 2003. "Food purchased away from home as a predictor of change in BMI z-score among girls." *International journal of obesity,* 28(2): 282-289

APPENDIX 1: THE STATEMENT OF RESEARCH SPONSORSHIP

This research is funded by the Agriculture and Food Research Initiative of the USDA National Institute of Food and Agriculture, grant number 2011-68001-30014. This work is also partly supported by the National Research Foundation of Korea (NRF-2011-330-B00074). Dr. Rodolfo M. Nayga Jr. is also affiliated with Department of Food and Resource Economics, Korea University and Norwegian Agricultural Economics Research Institute

UNIVERSITY OF ARKANSAS
DIVISION OF AGRICULTURE

Department of
Agricultural Economics an Agribusiness

UNIVERSITY OF ARKANSAS
DALE BUMPERS COLLEGE
OF AGRICULTURAL, FOOD & LIFE SCIENCES

217 Agriculture Building, University of Arkansas, Fayetteville, AR 72701-1201
479-575-2256 • Fax 479-575-5306

THE STATEMENT OF MUTIPLE AUTHORS

The chapter titled "The Effect of Neighborhood Fast Food on Childhood Obesity" has multiple authors. The student, Yiwei Qian, is the first author and completed at least 51% of the work for this chapter.

Dr. Rodolfo M. Nayga Jr. Dissertation Director (Co-chair)	Date
Dr. Arya B. Gaduh Dissertation Director (Co-chair)	Date

4. PROXIMITY TO PARKS AND CHILDHOOD OBESITY: A MATCHED INDIVIDUAL LEVEL ANALYSIS

Authors: Yiwei Qian, Rodolfo M. Nayga Jr, Michael R. Thomsen, Arya B. Gaduh, Heather L. Rouse

4.1 BACKGROUND AND LITERATURE REVIEW

The lack of physical activity can be blamed as one of the main reasons for the increasing rates of obesity. Lakdawalla and Philipson (2009) found that much of the increase in body weight throughout the 20th century is due to decreased physical activity. Hence, insufficient physical activity can be one of the reasons for the relatively high rates of childhood obesity (Anderson et al 1998; Nemet et al. 2005) in the US.

Parks and playgrounds are important spaces where children can have physical activity (Blanck, et.al, 2012). A number of studies have discussed the effect of neighborhood parks and playgrounds on childhood obesity. Potwarka, Kaczynski and Flack (2008) found that children with access to a park/playground within 1 km from their residence are almost five times more likely to be classified as being of a healthy weight than children without playgrounds or parks within 1 km from residence. Their results also suggest that availability of certain park facilities (e.g., a playground) may play a more important role in promoting physical activity and healthy weight status among children than availability of park space in general. Singh, Siahpush and Kogan (2010) found that overweight and obesity rates are higher among children in neighborhoods with no access to parks and playgrounds and that the effects are greater for females and younger children. For example, girls ages 10–11 who live in neighborhoods without access to parks and playgrounds were found to be two to four times more likely than their counterparts from neighborhoods with parks and playgrounds to be overweight or obese. Wolch, et al. (2011) followed 3,173 children aged 9-10 from 12 communities in Southern California in

1993 and 1996 for eight years and found that the children with access to parks which are within 500 meters from their homes had lower BMI at age 18 and the effects were larger for boys than for girls. Using the Children's Lifestyle and School Performance Study of Canada, Veugelers et al. (2008) find that children in neighborhoods with good access to playgrounds, parks, and recreational facilities are less likely to be overweight or obese.

Fan and Jin (2014) found a statistically and economically significant effect of neighborhood parks and playgrounds on childhood obesity based on covariate matching estimators by using the 2007 National Survey of Children's Health data. Their results suggested that adding neighborhood parks can reduce childhood obesity rates and make children more fit, but they cautioned that relevant interventions must consider the socioeconomic status of the targeted children as well as other neighborhood amenities. On average, the causal impact they estimated is greater among girls than boys; the treatment effect is greater among the young cohort aged 10–13 compared with those aged 14–17; Non-Hispanic white youth benefit from neighborhood parks and playgrounds much more than blacks and Hispanics; Children living above 133% of the federal poverty level are also more likely to benefit from neighborhood parks and playgrounds.

The focus of our study is similar to that of Fan and Jin (2014). Our study, however, differs from their study in many respects. First, we use a unique panel dataset that includes measured BMI of children and geocoded parks and trails in northwestern Arkansas to analyze the effect of parks and trails on children's BMI. Fan and Jin (2014) used a cross-sectional dataset from the 2007 National Survey of Children's Health. Second, their data on parks locations are also from self-reports while our park data are actual parks that were geo-coded. Third, we focus on a region of Arkansas which has relatively high childhood obesity rate compared with other states and conduct our analysis separately for urban and rural areas, while Fan and Jin (2014) used a random sample of households in each of the 50 states and the District of Columbia. Fourth, we use propensity score matching (PSM) instead of covariate matching (CVM) to estimate the average treatment effects. The idea of PSM and CVM is similar in that they both impute counterfactual outcomes for treated individuals using the untreated individuals with similar values for the covariates. Fan and Jin (2014) used CVM because it allowed exact matching for some of their crucial variables. In our research, we use the PSM instead of the CVM method since most of our covariates are binary indicator variables which can be easily exactly matched.

Our results suggest that neighborhood parks or trails have a significant and negative effect on children's BMI z-score. Our results also show that girls in urban areas are more likely to be influenced by neighborhood parks than urban boys. In contrast, the effect of parks and trails is greater among boys than girls in rural areas.

The next section describes the data sources and the variables used in the analysis. Section 4.3 discusses the empirical strategy we used to identify the effect of neighborhood parks on children's BMI. Section 4.4 discusses the matching quality and presents the results.

4.2 DATA

Our data come from several sources. First, we use the Arkansas BMI data from 2004 to 2010. These data are from the Arkansas Center for Health Improvement (ACHI), as discussed in Section 2. In addition to the BMI z-scores, the Arkansas BMI dataset includes demographic information about the child's gender, race, ethnicity and whether the child was eligible for the free or reduced lunch program. After 2007, BMI assessments switched to a biennial schedule with only children in the even grades being measured in any given year. So our dataset only includes kindergarten through 12th grade BMI screenings for the 2004-2007 period.

Second, we contacted the Park and Recreation Departments of several cities in northwest Arkansas to get the locations of the parks and trails in each city. These cities include Bella Vista, Bentonville, Bethel Heights, Farmington, Fayetteville, Lowell, Rogers, Siloam Springs, Springdale, Van Buren, and Fort Smith. Our definition of park includes all those with playgrounds and trails. There are 191 parks in our data; 16 of them were built during 2004 to 2007. After acquiring the parks data, we then overlaid the parks' coordinates onto the residential coordinates of students in the BMI data for each year. By doing this, we were able to get the network distance[14] from residence to the nearest park for each individual. To create a comparable sample, we generated dummy variables to describe the park environment around children's residence. These dummy variables represent the presence of a park within half, one, two and five miles of a child's residence. To control for other dimensions of the food environment in our models, we developed a "low healthy-store access" variable to indicate whether the neighborhood where the child resides has low access to large grocery stores with fresh produce and other healthier food options.

Table 1 presents the variable names and definitions used in our study and the descriptive statistics. The total number of observations in our panel data is 17,022. 12,438 students always lived in an urban area and 4,584 students always lived in rural area during the 2004-2007 period of our study. For those who lived in urban places, 10.8% have at least one park or trail within a half mile of their residence, 32.2% have access to parks or trails within one mile of their

[14] Network distance is defined as the shortest distance from one point to another using road networks instead of a straight line.

residence, and76.5% have at least one park or trail within two miles of their residence. For those who lived in rural places, about 29% of students have at least one park or trail within five miles of their residence. 66% of urban children and 91.7% of rural children are white. 23.6% of urban children and 4.5% of rural children are Hispanic. About 29.4% of urban children and 16.2% of rural children always participated in the free or reduced lunch program, and 50.2% of urban children and 62.3% of rural children never participated in the free or reduced lunch program. 45.4% of urban children and 4% of rural children have low access to grocery stores during the period of study.

4.3 METHODOLOGY

We discussed the concept of the matching methodology in Section 2. We apply the same methodology in this study given that proximity of parks for individuals is a non-random event. New parks or trail can be built near the residence and the decision of building parks can be affected by regional obesity rates. However, in our case, many of the parks in our study area are relatively old; in fact only 8.4% of parks were built during 2004-2007. On the other hand, a household could move to a new place which has a different park environment and so the decision to move can be affected by some unobservable factors which may correlate with children's BMI. It is possible that a household self -selects to move to a new location which is closer to a park to improve their children's health status.

To tackle this problem of potential endogeneity, we employ propensity score matching to measure the average treatment effect of neighborhood parks on children's BMI z-score. The main idea of matching is to find a group of control individuals that are similar to the treated individuals in all pre-treated characteristics and then measure the average treatment effect on the treated group for these groups with similar characteristics. The matching techniques we used are *nearest-neighbor matching* with one neighbor and *Mahalanobis matching* (Rosenbaum and Rubin 1985). Table 1 presents the covariates used in the matching which include child's age, gender, ethnicity, income status and food environment status.[15] The average treatment effect on treated (ATT) can be written as follows:

$$ATT = E[Y_{1i} - Y_{0i} | K_i = 1]$$

where Y_{1i} denotes the potential outcome variable for individual i in treatment group and Y_{0i} denotes the potential outcome variable for individual i in the control group; K_i denotes the neighborhood park status for individual i. The treatment group includes only the students who always have park accessibility within certain distance from their residence from 2004-2007 while the control group includes only those who did not have access to parks within certain distance from their residence from 2004-2007. The outcome variable is children's BMI z-score during 2007.

We employ matching for urban and rural children separately because park proximity is quite different in urban and rural areas. Only 8.9% of rural children have at least one park within two miles of their residence. Thus instead of half, one or two miles, we used a treatment group that has park accessibility within five miles for rural children. We then also match on male and female subsamples separately because males and females experience substantially different metabolic processes and types of body development when they are teenagers and adolescents (Tarnopolsky 1999; Fan and Jin 2014) and neighborhood amenities may affect males and females differently (Gomez et al. 2004; Fan and Jin 2014).

[15] Section 2 has already discussed the idea of different matching strategies and the principles in choosing control variables.

Table 1. Description and Descriptive Statistics of Variables in the Study (N=17,022)

Variables	**Description**	**Mean**	
		Urban** (N=12,438)	Rural (N=4,584)
Outcome Variable			
BMI z-score	Individuals' BMI z-score in 2007	0.568 (1.018)	0.523 (1.017)
Treatment Variable			
Park_Half	If the distance between individual's residence and nearest park is less or equal to 0.5 miles, then =1, 0 otherwise	0.108 (0.311)	0.001 (0.036)
Park_One	If the distance between individual's residence and nearest park is less or equal to 1 mile, then =1, 0 otherwise	0.322 (0.467)	0.006 (0.079)
Park_Two	If the distance between individual's residence and nearest park is less or equal to 2 miles, then =1, 0 otherwise	0.765 (0.423)	0.089 (0.285)
Park_Five	If the distance between individual's residence and nearest park is less or equal to 5 miles, then =1, 0 otherwise	0.883 (0.321)	0.290 (0.453)
Control Variables			
Individual level			
Age	Age of student in months	118.51 (35.53)	122.8 (36.11)
White	Binary indicator (if individual is White then =1, 0 otherwise)	0.660 (0.498)	0.917 (0.274)
Hispanic	Binary indicator (if individual is Hispanic then =1, 0 otherwise)	0.236 (0.424)	0.045 (0.208)
Female	Binary indicator (if individual is female then =1, 0 otherwise)	0.463 (0.498)	0.461 (0.498)
Free or Reduced	Binary indicator (if individual always participated in free and reduced lunch from 2004-2007 then =1, 0 otherwise)	0.294 (0.455)	0.162 (0.369)
Non Free and Reduced	Binary indicator (if individual never participated in free and reduced lunch from 2004-2007 then =1, 0 otherwise)	0.502 (0.500)	0.623 (0.484)

Lowaccess*	Binary indicator that describes the accessibility to large grocery stores. "Low access" status is defined for urban students living more than one-mile from a large grocery store and for rural students living more than 10 miles from a large grocery store. This variable equals to one if children always stay in "Low access" status	0.454 (0.497)	0.040 (0.194)

Note: Standard errors appear in parenthesis.
* The low access area criteria follow the USDA/ERS Food Desert Locator, website: http://www.ers.usda.gov/data/fooddesert/about.html
** Urban (Rural) group are defined as individuals that always live in urban (rural) area.

4.4 RESULTS AND DISCUSSION

To assess the quality of the estimated treatment effects, we present the results of the balance tests in Table 2, which measure how similar the matching variables are between the treated and untreated groups before and after matching. While there are significant differences on age, ethnicity, income level and food environment between the treatment group and the control group before matching, the differences shrink after using the nearest neighbor matching. The mean difference is still statistically significant however. The differences almost disappear (except for the continuous variable) after using Mahalanobis matching[16].

Table 3 presents the results of the average treatment effect on treated of neighborhood parks on children's BMI z-score. For urban children, proximity to a park within half a mile and one mile from the residence has no significant effect on BMI z-score. In contrast, we see a significant and negative effect on BMI z-score of proximity to park within two miles from the residence using the Mahalanobis matching procedure. The estimate suggests that for urban

[16] Matching with male and female separately have similar results for balance test.

children who have park accessibility within two miles from their residence, their BMI z-scores are

Table 2. Balance Test of Matching Covariates

Variables		**Difference[a]** **Urban(N=12,438)**			**Rural (N=4,584)**
		Park_Half	Park_One	Park_Two	Park_Five
Age	Unmatched	-2.525***	-2.509***	2.480***	-0.396
	Nearest-neighbor	-1.38**	0.06***	-0.037***	-0.72
	Mahalanobis	-0.05**	-0.04***	-0.22***	-0.08
White	Unmatched	0.016	0.068***	0.157***	0.028***
	Nearest-neighbor	-0.001*	0	0.002***	0.006*
	Mahalanobis	0	0	0	0
Hispanic	Unmatched	0.236	-0.053***	-0.128***	-0.050***
	Nearest-neighbor	0.004	-0.005***	0.001***	0
	Mahalanobis	0	0	0	0
Female	Unmatched	-0.018	-0.014	-0.015	-0.006
	Nearest-neighbor	-0.012	0.003	-0.025	-0.023
	Mahalanobis	0	0	0.001	0
Free and Reduced	Unmatched	-0.023*	-0.072***	-0.124***	-0.133***
	Nearest-neighbor	-0.016	-0.16***	-0.001***	0
	Mahalanobis	0	0	0	0
Non Free and Reduced	Unmatched	-0.025*	0.045***	0.144***	0.067***
	Nearest-neighbor	0.015**	0.004**	0.008***	-0.006***
	Mahalanobis	0	0	0	0
Lowaccess	Unmatched	0.094***	0.143***	0.219***	0.036***
	Nearest-neighbor	0.001***	0.004***	-0.001***	0
	Mahalanobis	0	0	0	0

Note: *,**,***denote significance at the 0.1, 0.05, 0.01 level, respectively
a. Mean differences of each matching covariate between those in the untreated group and those in the treated group. We used proportion test for binary variables and t-test for student's age.

0.078 lower than children in urban areas who do not have access to a park within two miles from their residence. Based on average BMI z-score, this suggests about a 2.78 difference in BMI percentile points. For rural children, the results suggest that those who have park accessibility within five miles from their residence have lower BMI (i.e., 0.195 standard deviation, which equates to 6.9 BMI percentile points) than those who do not have access to parks within 5 miles from their residence. This finding is consistent across the use of nearest neighbor and Mahalanobis matching procedures.

Table 3 also presents the results for males and females separately. For urban boys, there are no significant effects of neighborhood parks on BMI z-scores across all three distances of half mile, one mile, and two miles from residence. For urban girls, the result using the Mahalanobis matching procedure suggests that there is a significant and negative effect. The estimate suggests that for urban girls who have park accessibility within two miles from their residence, their BMI z-scores are 0.114 (4 BMI percentile points) lower than children in urban areas who do not have access to a park within two miles from their residence. These results are consistent with those of Fan and Jin (2014) where they found that girls are more likely to be affected by neighborhood parks than boys. Singh, Siahpush and Kogan (2010) also suggested that girls are more vulnerable than boys to less favorable neighborhood built environmental conditions with respect to obesity outcomes. However, in our study, for children (both boys and girls) in rural areas, the average treatment effects on the treated are negative and significant, with magnitudes of 0.209 and 0.146 standard deviation when we use the nearest neighbor matching. The results are robust with the use of Mahalanobis matching. The effect on boys is larger than on girls, which suggests that rural boys could benefit more from neighborhood parks than rural girls.

Table 3. Average Treatment Effect on Treated of the Neighborhood Parks on Childhood Obesity with Different Matching Methods

	Urban			Rural		
	Entire Group	Male	Female	Entire Group	Male	Female
Number of Observations	12,438	6,678	5,760	4,584	2,468	2,116
Treatment Variable: Park_Half						
Number of Treated Observations	1,227	645	582			
Nearest-neighbor	-0.027 (0.044)	-0.078 (0.064)	-0.020 (0.061)			
Mahalanobis	-0.034 (0.044)	-0.039 (0.064)	-0.008 (0.062)			
Treatment Variable: Park_One						
Number of Treated Observations	3,623	1,927	1,696			
Nearest-neighbor	0.010 (0.030)	-0.053 (0.043)	-0.054 (0.041)			
Mahalanobis	-0.032 (0.030)	-0.023 (0.043)	-0.052 (0.042)			
Treatment Variable: Park_Two						
Number of Treated Observations	8,669	4,647	4,022			
Nearest-neighbor	-0.049 (0.033)	-0.023 (0.047)	-0.075 (0.047)			
Mahalanobis	-0.078** (0.034)	-0.049 (0.047)	-0.114** (0.048)			
Treatment Variable: Park_Five						
Number of Treated Observations				1,321	706	615
Nearest-neighbor				-0.195** (0.047)	-0.209** 0.067	-0.146** (0.067)
Mahalanobis				-0.209** (0.048)	-0.215** (0.067)	-0.149** (0.067)

Note: Standard errors appear in parenthesis.
**denote significance at the 0.05 level, respectively

There is one possible explanation for the difference in the results between urban and rural children. Since boys are more likely to engage in outdoor physical activities than girls and there are more options in urban areas for outdoor recreational activities, proximity to parks for urban boys may not be important because they can easily find other places for physical activities. But for rural children, the options for outdoor activity is less than for urban children and so proximity to parks can likely make a greater impact on boys than girls, who are less likely to engage in physical activities.

REFERENCES

1. Andersen, R.E., C.J. Crespo, S.J. Bartlett, L.J. Cheskin, and M. Pratt. 1998."Relationship of physical activity and television watching with body weight and level of fatness among children." *JAMA: the journal of the American Medical Association,* 279(12): 938-942.

2. Blanck, H.M., D. Allen, Z. Bashir, N. Gordon, A. Goodman, D. Merriam & C. Rutt. 2012. “Let's go to the park today: the role of parks in obesity prevention and improving the public's health.” *Childhood Obesity (Formerly Obesity and Weight Management), 8*(5): 423-428.

3. Fan, M., and Y. Jin. 2014."Do neighborhood parks and playgrounds reduce childhood obesity?" *American Journal of Agricultural Economics,* 96(1): 26-42.

4. Gomez, J. E., B. A. Johnson, , M. Selva, &J. F. Sallis. 2004. “Violent crime and outdoor physical activity among inner-city youth” *Preventive medicine,39*(5): 876-881.

5. Lakdawalla, D., and T. Philipson. 2009. “The Growth of Obesity and Technological Change: A Theoretical and Empirical Examination.” *Economic and Human Biology,* 7(3): 283–293.

6. Nemet, D., S. Barkan, Y. Epstein, O. Friedland, G. Kowen, and A. Eliakim. 2005. "Short-and long-term beneficial effects of a combined dietary–behavioral–physical activity intervention for the treatment of childhood obesity." *Pediatrics,* 115(4): 443-449.

7. Potwarka, L. R., A.T. Kaczynski, and A.L. Flack. 2008."Places to play: association of park space and facilities with healthy weight status among children." *Journal of community health,* 33(5): 344-350.

8. Singh, G.K., M. Siahpush, and M.D. Kogan. 2010b. "Neighborhood socioeconomic conditions, built environments, and childhood obesity." *Health Affairs,* 29(3): 503-512.

9. Veugelers, P., F. Sithole, S. Zhang, and N. Muhajarine. 2008. Neighborhood Characteristics in Relation to Diet, Physical Activity and Overweight of Canadian Children. *International Journal of Pediatric Obesity.* 3(3):152–159.

10.Wolch, J., M. Jerrett, K. Reynolds, R. McConnell, R. Chang, N.Dahmann &K. Berhane. 2011. "Childhood obesity and proximity to urban parks and recreational resources: A longitudinal cohort study." *Health & place*, *17*(1): 207-214

APPENDIX 1: THE STATEMENT OF RESEARCH SPONSORSHIP

This research is funded by the Agriculture and Food Research Initiative of the USDA National Institute of Food and Agriculture, grant number 2011-68001-30014. This work is also partly supported by the National Research Foundation of Korea (NRF-2011-330-B00074). Dr. Rodolfo M. Nayga Jr. is also affiliated with Department of Food and Resource Economics, Korea University and Norwegian Agricultural Economics Research Institute

UNIVERSITY OF ARKANSAS
DIVISION OF AGRICULTURE

Department of
Agricultural Economics an Agribusiness

UNIVERSITY OF ARKANSAS
DALE BUMPERS COLLEGE
OF AGRICULTURAL, FOOD & LIFE SCIENCES

217 Agriculture Building, University of Arkansas, Fayetteville, AR 72701-1201
479-575-2256 • Fax 479-575-5306

THE STATEMENT OF MUTIPLE AUTHORS

The chapter titled "Proximity to Parks and Childhood Obesity: A Matched Individual Level Analysis" has multiple authors. The student, Yiwei Qian, is the first author and completed at least 51% of the work for this chapter.

Dr. Rodolfo M. Nayga Jr. Dissertation Director (Co-chair)	Date
Dr. Arya B. Gaduh Dissertation Director (Co-chair)	Date

5. CONCLUSION

This dissertation explores the effect of school and environmental factors on childhood obesity using an individual panel data set of Arkansas public schoolchildren.

In first chapter, we used a relatively unique panel dataset with measured BMI of schoolchildren in Arkansas and a panel difference-in-difference estimation procedure to examine the effect of FFVP participation on students' BMI z-scores and percentiles. Before the panel DID estimation, however, we used several matching methods such as Propensity Score Matching and Coarsened Exact Matching to match FFVP participants to non-participants. We then estimated both fixed effects and random effects DID models using the matched samples. In addition to being the first to examine the effect of FFVP participation on childhood obesity, another contribution of this chapter is the investigation of the sensitivity of the estimated effects to the use of different matching techniques.

Our results show that while the FFVP effects on BMI are positive and not statistically significant using matched samples with less balance on the covariates, they are negative and significant when using stricter matching techniques such as the CEM, which provided more balance in characteristics between the treated and control groups. Specifically, our panel DID results using matched samples from these two techniques suggest that FFVP participation can reduce BMI percentile by 3.8 percentile points, ceteris paribus. Given this finding and those of other past studies discussed previously suggesting the generally positive effects of FFVP participation on students' fruit and vegetable consumption, the FFVP program seems like a promising way of improving the diet and reducing childhood obesity among elementary school children, especially considering that the cost for each student in participating schools has been estimated to be only 50-75 dollars per year.

However, given that the FFVP has only been implemented in Arkansas since 2008, more research is needed to draw more definitive conclusions. For instance, future research should test the robustness of our findings when more data become available (i.e., more years of implementation). To our knowledge, this study represents a first attempt at examining the childhood obesity effects of the FFVP program in a state with a relatively high obesity rate such as Arkansas. Hence, it would also be important to examine whether our findings will hold true in other states that have implemented the FFVP program in schools.

In second chapter, we investigated the effect of density of fast food restaurants near children's residence on their BMI. We used an instrumental variable, the distance between residence and nearest US highway to identify our model following Dunn (2010), Anderson and Matsa (2011) and Alviola et al. (2014). The results suggest that increasing the density of neighborhood fast food restaurants can significantly increase children's BMI z-score. Specifically, one more fast food establishment within half a mile near children's residence will cause BMI z-score to increase by 0.091. This would be equivalent to a 2.7 BMI percentile point increase based on the average BMI level of our sample. Moreover, for every additional fast food establishment within one mile from a child's residence, the BMI z-score will increase by 0.035, which would be equivalent to a 1.13 BMI percentile point increase.

We also found significant differences in the effects of fast food density on BMI z-score across different sub-groups of children. Interestingly, our results indicate that the fast food effects are positive and significant for females but not for males. Previous studies (Binkley et al., 2000; French et al., 2000; Dunn, 2010) have also found significant evidence regarding the positive relationship between fast food proximity and adult females' BMI outcome but there is scant information or discussion in the literature on the reasons why girls are more susceptible than males to increases in fast food exposures. Singh, Siahpush and Kogan (2010) suggested that girls' BMI are more likely to be influenced by unfavorable aspects of the built environment. In contrast to Dunn (2010) and Anderson and Matsa (2011) who found that greater fast food proximity has little impact on obesity outcomes of white adults in rural areas and to Dunn (2010) who reported that fast food availability effect is significant for minority adults, our finding shows that white children, children living in rural areas, and children from non-low-income households are more likely to be impacted by fast food availability. Given that our study was focused on children in Arkansas, it would be interesting to test the robustness of our findings in other states. As mentioned above, children's dietary quality depends on their parents' food choice. One explanation can be that for higher-income families, the opportunity cost of food preparation for parents is relatively high, which could then effectively increase the likelihood of substituting home cooked meals with fast food. And for families living in rural areas, it is possible that there are fewer restaurant options and so an increase in the number of fast food restaurants in a rural area would represent a major change in the food environment in comparison to a similar increase in an urban area.

In third chapter, we analyze the effect of neighborhood parks around residence of northwestern Arkansas children on BMI outcomes. We use a statewide panel dataset for Arkansas covering the 2004 through 2007 period. To build comparative groups, we consider those living near a park as the treatment group (i.e., within certain distances from the residence). We then employ propensity score matching approach to measure the average treatment effect on the treated. The results indicate that the proximity of neighborhood parks and trails from the residence can have significant and negative effects on children's BMI z-scores within two miles in an urban area and within five miles in a rural area. Our results also show that both urban and rural girls are significantly influenced by neighborhood parks. The park effect is not significant for urban boys but is significant for rural boys. These results are consistent with the Fan and Jin (2014) and Singh, Siahpush and Kogan (2010) findings that indicate that girls are more likely to be influenced by neighborhood parks than boys. However, neither of them analyzed urban and rural group separately. Interestingly, our results suggest that in rural areas, boys are more impacted by neighborhood parks than rural girls. This is not the case in urban areas where proximity to parks does not significantly affect boys' BMI. One possible explanation is that boys are more likely to engage in outdoor physical activities than girls and since there are more options in urban areas for outdoor recreational activities for boys, proximity to parks in these places may not be as important as in rural areas where there are more limited options.

One suggestion for policy intervention is to build more parks and trails in rural areas to encourage children to do more outdoor physical activity. For further research, we can measure the effect for different races, age group and income level group to explore the park effect on different groups of children. Moreover, to further explore the potential effects of parks in neighborhoods without particular amenity, it is worthwhile to measure the average treatment effects on untreated group (ATU) as well. Additional measurements such as the effect of number and size (i.e., acreage) of parks within certain distances from residence of the child could also be used to test the robustness of our findings. We also need to conduct falsification tests given the possibility that some unobservable characteristics are driving our results.

REFERENCES

1. Alviola, P.A., R.M. Nayga, M.R. Thomsen, D. Danforth, and J. Smartt. 2014. "The effect of fast-food restaurants on childhood obesity: a school level analysis." *Economics & Human Biology*, 12: 110-119

2. Anderson, M. L., and D. A. Matsa. 2011. "Are Restaurants Really Supersizing America?"*American Economic Journal: Applied Economics*, 3(1): 152-88.

3. Binkley, J. K., J. Eales, and M. Jekanowski. 2000. "The relation between dietary change and rising US obesity." *International journal of obesity,* 24(8): 1032-1039.

4. Dunn, R.A. 2010. "The effect of fast-food availability on obesity: an analysis by gender, race, and residential location." *American Journal of Agricultural Economics* 92(4): 1149-1164.

5. Fan, M., and Y. Jin. 2014."Do neighborhood parks and playgrounds reduce childhood obesity?" *American Journal of Agricultural Economics,* 96(1): 26-42.

6. French, S. A., L. Harnack, and R. W. Jeffery. 2000. "Fast food restaurant use among women in the Pound of Prevention study: dietary, behavioral and demographic correlates." *International journal of obesity and related metabolic disorders: journal of the International Association for the Study of Obesity*, 24(10): 1353-1359.

7. Singh, G.K., M. Siahpush, and M.D. Kogan. 2010. "Neighborhood socioeconomic conditions, built environments, and childhood obesity." *Health Affairs,* 29(3): 503-512.

.