



Graduate Theses, Dissertations, and Problem Reports

2013

Examination of Factors Associated with Obesity, Physical Activity and Income in Metropolitan Areas of the United States

Ahadu T. Tekle
West Virginia University

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

Recommended Citation

Tekle, Ahadu T., "Examination of Factors Associated with Obesity, Physical Activity and Income in Metropolitan Areas of the United States" (2013). *Graduate Theses, Dissertations, and Problem Reports*. 3668.

<https://researchrepository.wvu.edu/etd/3668>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

**Examination of Factors Associated with Obesity, Physical Activity and Income in
Metropolitan Areas of the United States**

Ahadu T. Tekle

**Dissertation Submitted to the
Davis College of Agriculture, Natural Resources and Design**

**at West Virginia University
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in
Natural Resource Economics**

**Cheryl Brown, Ph.D., Chair
Dale Colyer, Ph.D.
Tesfa Gebremedhin, Ph.D.
Michael Hendryx, Ph.D.
Tim Phipps, Ph.D.**

Department of Agricultural and Resource Economics

**Morgantown, West Virginia
2013**

**Key Words: BMI, Obesity; Physical Activity; Built Environment; Natural Amenities;
Outdoor Recreational Opportunities; Panel Data; Random Effect; Ordered Probit**

Copyright 2013 Ahadu T. Tekle

Abstract

Examination of Factors Associated with Obesity, Physical Activity and Income in Metropolitan Areas of the United States

Ahadu T. Tekle

The prevalence of obesity has increased significantly in the past three decades. Obesity and inadequate physical activity are the major contributors to health problems. In addition to understanding the contribution of biological and psychological factors in the growth of obesity rates, there is a growing interest in understanding the impact of environmental factors on obesity. This study examines the relationship between obesity, socioeconomic factors, environmental factors and physical activity. The study is focused on metropolitan areas and relied on metropolitan county level built environment, natural amenities and outdoor recreational opportunities measures. Individual level socioeconomic and location information is from the National Longitudinal Youth Survey (NLYS79). Panel data of 1768 individual observations from 1998-2008 is used. The study employed 2SLS and ordered probit estimation methods.

Not all built environment measures have significant association with obesity. Density is negatively associated with obesity, whereas, mixed land use is positively associated with obesity. No evidence is found that links street connectivity and centeredness to obesity. The association between built environment and physical activity is mixed and inconsistent. Mixed land use is found to be positively associated with the likelihood participation in regular physical activity and negatively associated with physical inactivity and occasional physical activity, whereas, street connectivity is negatively associated with regular physical activity and positively linked with no physical activity and occasional physical activity. No evidence is found that links density and centeredness with the likelihood of participating in physical activity or inactivity.

The implications of natural amenities and recreational opportunities for obesity and physical activity are mixed. Climate and winter-based recreational opportunities are found to be positively associated with obesity whereas, no significant relationship is observed between BMI and land-based and water-based recreational opportunities. Land-based and winter-based recreational opportunities are found to increase the likelihood of participating in regular physical activity and decrease the likelihood of physical inactivity and occasional physical activity. Climate and water-based recreational opportunities are not significantly associated with the likelihood of participating in physical activity or physical inactivity.

Acknowledgements

I would like to express my deepest gratitude to my committee chair Dr. Cheryl Brown for her valuable guidance, support, care and patience. I would like also to thank my committee members, Dr. Tesfa Gebremedhin, Dr. Dale Colyer, Dr. Michael Hendryx and Dr. Tim Phipps for their review and critical comments.

I wish to extend my sincere thanks to the Division of Resource Management, Davis College of Agriculture, Natural Resources, and Design, and the President's Office for Social Justice for providing research assistantship and tuition waivers for my graduate studies both in Agricultural and Resource Economics (M.Sc.) and Natural Resource Economics (Ph.D.). In addition I would like to thank Lisa A. Lewis, Melanie Jimmie, and Ellen Hartley-Smith for their administrative assistance and willingness to help at any time,

I wish to express my deepest gratitude to my parents who sacrificed in many ways to bring me up to this stage. Their constant love and support has been inspirational. I am also grateful to my brothers and sisters for their love and support throughout. My greatest thanks go to my husband Yohannes Hailu who has always believed in me and supported me to reach my goals, and to my daughter Leah who has inspired me to work hard.

Finally, I would like to extend special thanks for Dr. Mulugeta Kahsay for his valuable input and support, to all my friends, families and colleagues.

Dedication

I dedicate this work to my husband Yohannes Hailu, my daughter Leah, my parents and siblings.

Table of Contents

Abstract	ii
Acknowledgement	iii
Table of Contents	iv
List of Figures	vi
List of Tables	vi
CHAPTER 1: INTRODUCTION	
1.1.Introduction and Problem Statement	1
1.2.Objective of the Study	7
1.3.Hypothesis	7
1.4.Methodology	8
1.5.Overview of the study Area	8
1.6.Organization of the Study	10
CHAPTER 2: LITERATURE REVIEW	
2.1. Definition of the Built Environment and Urban Sprawl	11
2.1.1. Definition of the Built Environment	11
2.1.2. Definition and Measurement of Sprawl	13
2.2. The Impact of the Built Environment and Sprawl on Obesity	15
2.2.1. Sprawl Impact Travel Behavior and Obesity	17
2.3. The Relationship between Socioeconomic Factors and Obesity	18
2.4. The Labor Market and Obesity	19
2.5. Obesity and Earnings	20
2.6. Methodological Approaches of Previous Studies	21
CHAPTER 3: THEORETICAL FRAMEWORK	
3.1. Introduction	23
3.2. Implicit Demand Functions	27
3.3. Comparative Static Analysis	28
CHAPTER 4: EMPIRICAL MODELS AND DATA DESCRIPTION	
4.1. Introduction	31
4.2. Methodology	31
4.2.1. Fixed Effect and Random Effect Approaches	32
4.2.2. Ordered Probit Model	33
4.3. General Model Specification	35
4.4. Random Effect Model Specification	37
4.5. Ordered Choice Models	40
4.6. Data Sources and Variables Description	42
CHAPTER 5: RESULT AND ANALYSIS	
5.1. Empirical Estimation	55
5.2. Findings and Analysis	55
5.2.1. BMI Equation: Findings and Analysis	56

5.2.2. Physical Activity Equation: Findings and Analysis	62
5.2.3. Income Equation: Findings and Analysis	69
5.3. Revisiting the Hypotheses	71
CHAPTER 6: SUMMARY AND CONCLUSION	
6.1. Summary and Conclusions	76
6.2. Recommendations	81
6.3. Limitation of the Study	82
6.4. Future Study	84
REFERENCE	86
APPENDIX	95

List of Figures

Figure 1.1. Obesity Trends among U.S. Adults.	2
---	---

List of Tables

Table 1.1. Metropolitan Areas Used in Study	8
Table 2.1. Dimensions of the Built Environment	12
Table 4.1. Socio-economic & Demographic Variable Description and Data Sources	43
Table 4.2. Physical Activity Dummy Variables	44
Table 4.3. Built Environment Variables Description and Data Sources	46
Table 4.4. Definition and Data Sources for Natural and Outdoor Recreational Amenities	48
Table 4.5. Results of Principal Component Analysis for Land-based Outdoor Recreational Index	48
Table 4.6. Results of Principal Component Analysis for Water-based Outdoor Recreational Index	49
Table 4.7. Results of Principal Component Analysis for Winter-based Outdoor Recreational Index	50
Table 4.8. Results of Principal Component Analysis for Climate Index	50
Table 4.9. Regional Categorization of States	51
Table 4.10. Summary of Descriptive Statistics	53
Table 5.1. BMI Equation Results	58
Table 5.2. Physical Activity Equation Results	63
Table 5.3. Income Equation Results	70

Chapter 1

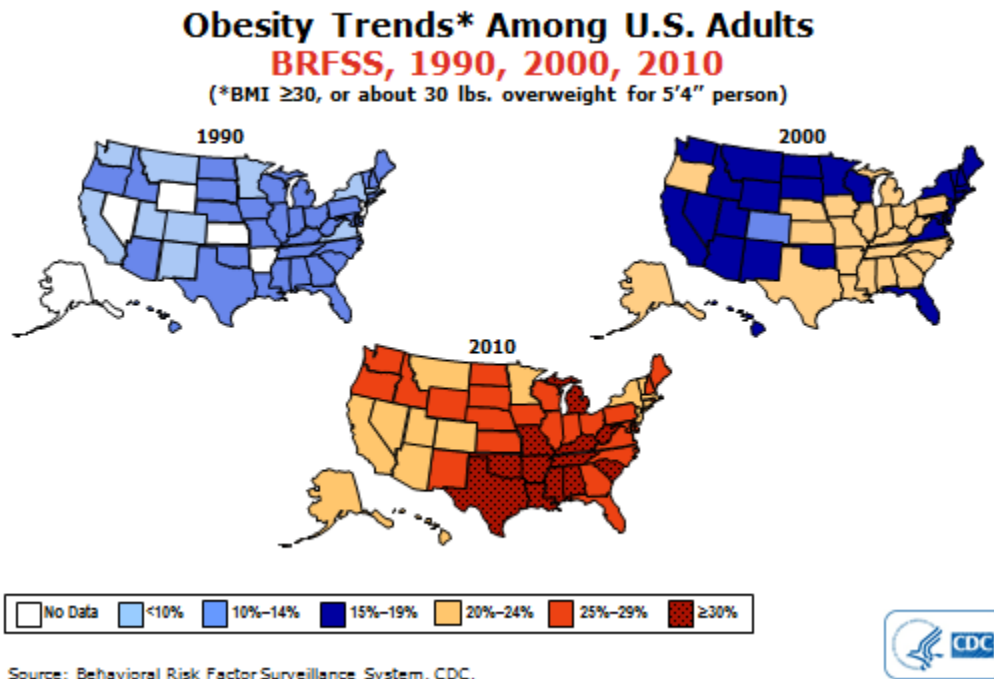
1.1. Introduction and Problem Statement

Obesity is a major public health and public policy concern in the United States. The number of obese and overweight adults has increased over time. According to the National Health and Nutrition Examination Surveys (NHANES) of 2003-2006 and 2007-2008, over two thirds of U.S. adults are considered overweight and over one third of adults are considered obese (U.S. Department of Health and Human Services, 2012). Even more alarming, the prevalence of obesity is expected to increase in the near future. Wang and Beydoun (2007) predicted by 2015, 75% of the adult population will be obese or overweight and 41% of adults will be obese.

Obesity, measured by Body Mass Index (BMI), has negative health implications. According to Rubensteinn (2005) obesity and overweight are linked to an increased risk of coronary heart disease, type 2 diabetes, cancers, hypertension, dyslipidemia, stroke, sleep apnea and respiratory problems. Obesity and overweight have also become the second leading preventable cause of disease and death and are likely to become the leading cause of disease and death in the near future (Wang and Beydoun, 2007).

The prevalence of obesity has significantly changed over the past two decades. Figure 1.1 indicates the trend of obesity among adults in 1980, 1990 and 2010. As shown in the figure, the prevalence of obesity has increased nationwide in the last two decades. For example, in 1990, the prevalence of obesity was less than 15% in all states. However, in 2000, the prevalence of obesity was over 20% nationwide. Recently, the prevalence of obesity has increased to more than 30% in eleven states (Texas, Oklahoma, Arkansas, Louisiana, Alabama, Mississippi, South Carolina, Tennessee, West Virginia, Kentucky, Missouri and Michigan) in 2010.

Figure 1.1. Obesity Trends among U.S. Adults.



Source: CDC, Behavioral Risk Factor Surveillance System (2010)

Obesity and overweight have significant direct and indirect cost implications for individuals and the economy. The direct cost for individuals is the medical expenses they pay for diagnosis and treatment of diseases resulting from obesity and overweight. Obese people face higher lifetime medical expenditures than non-obese people and they are likely to spend an additional \$10,000 throughout their lifetimes compared to non-obese individuals (Bhattacharya and Sood, 2004). The costs of obesity and overweight are not only borne by the affected individual but also by society at large, as obesity imposes a tax burden on society, increases the number of welfare recipients, and contributes to a loss of income to the individual and household (Tunceli, Kemeng, and Williams, 2006)

Medical costs increase with an increase in the prevalence of obesity and overweight. Obesity accounted for 27% of the rise in medical expense between 1988 and 2001 (Keehan,

Sisko, and Truffer, 2008). Finkelstein et al. (2009) reported the total medical cost of obesity in 2008 reached \$147 billion. The share of health care expenditures in terms of share of gross domestic product (GDP) has increased over time. For instance, the share of health care expenditures in terms of GDP increased from 7.2% in 1970 to 16% in 2006 (Catlin et al., 2008). Further, the GDP share of health care expenditures is expected to increase to 19.5% in 2017 partly due to the rise of medical problems associated with obesity and overweight (Catlin et al., 2008).

The indirect cost of obesity and overweight includes loss of productivity and human capital (Hammond and Levine, 2010). The indirect effects of obesity can be examined by looking at its impacts on the labor market. Obesity and overweight impact both the demand and supply side of the labor market. Obesity contributes to a lower level of labor supply as obese people tend to have more sick days, higher rates of absenteeism, and lower motivation due to weight-related health and personal problems in the work environment (Tunceli, Kemeng, and Williams, 2006). In addition, obese and overweight people may be marginalized in the job market and are less likely to find a job due to a high level of prejudice and discrimination (Puhl and Brownell, 2001).

The Centers for Disease Control and Prevention (CDC) identified three major contributing factors for the prevalence of obesity: genetics, calorie balance and physical environment. Genetic makeup of an individual is one of the three factors that the CDC considers an important contributor to obesity. At an individual level, biology and behavior may explain the risk of being obese or overweight; however, these factors alone may not fully determine the existing prevalence of obesity (Huang and Glass, 2008). In addition, it is difficult to understand

the contribution of genetics to the recent obesity epidemic given the fact that genetic makeup of the society changes slowly through time (Hill and Trowbridge, 1998).

Individuals gain weight when calorie intake is in excess of calories the body needs. The accumulation of excess calories over time ultimately results in weight gain. The average calorie intake in the U.S. has increased over the past four decades (National Center for Health Statistics, 2004). Alarming, the consumption of energy dense and less nutritious food items that contribute to weight gain has increased through time (Wells and Buzby, 2008). The increase in calories, however, has not been balanced by an increase in physical activity or exercise, which could burn calories and help maintain normal weight. The physical environment is the third contributing factor for the prevalence of obesity and overweight. Physical environment is defined as objective and perceived characteristics of the physical context in which people spend their time (e.g. home, neighborhood, school) including aspects of urban design (e.g. presence and absence of sidewalks) traffic density and speed, distance to and design of venues for physical activity (e.g. playgrounds, parks and school yard) crime, safety and weather conditions (Davison and Lawson, 2006). Physical environment includes both built environment and natural environment. Built environment includes buildings, transportation systems and open space (Northridge, Sclar, and Biswas, 2003). Particular characteristics of the physical environment can either provide opportunities for or hinder physical activity.

Numerous studies documented the contribution of environmental factors, such as the built environment, to physical activity levels and obesity and overweight (Lee, Ewing and Sesso, 2009; Garden and Jalaludin, 2008; Eid et al., 2008; Ewing, Brownson, and Berrigan, 2006). Environmental factors may contribute to one's decision and ability to be physically active and adopt a healthy lifestyle. High density or compact neighborhoods encourage people to walk or

ride bicycles and to be less dependent on cars for daily activities compared to low-density neighborhoods (Saelens et al., 2003; Giles-Corti and Donovan, 2002). Mixed neighborhoods, where houses are mixed with a variety of businesses and workplaces, encourage people to walk more than neighborhoods that are not mixed and where land is specifically designated for a single purpose, for example, strictly commercial or strictly residential (Saelens et al., 2003). In addition, neighborhoods with limited or no outdoor recreational opportunities, such as parks, trails, and playgrounds, result in people becoming less physically active (Plantinga and Bernell, 2007).

In examining the implication of sprawl on obesity and physical activity, certain aspects of sprawl have been considered, namely density and mixed land use. For example, Eid et al. (2008) investigated only residential sprawl and mixed use. Though important, the effect of other sprawl measures on BMI and physical activity, such as street connectivity and centeredness, are often ignored. Thus, examining the implication of different dimensions of sprawl will be one of the contributions of this study to the existing literature. The general consensus among those in public health and urban planning is that proper planning and improved land use and urban design, for example increasing density and mixed land use, may encourage physical activity and help curb obesity (Garden and Jalaludin, 2008). However, this argument is challenged by the finding of recent studies which suggests that obese people may prefer to reside in sprawling places that encourage a sedentary lifestyle and where there is little opportunity to be physically active, while people with ideal weight sort themselves into places where there are opportunities that promote physical activity (Eid et al. 2008; Plantinga and Bernell, 2007).

The contribution of physical activity to health and well-being has been well established. According to the According to the U.S. Department of Health and Human Services (1996),

physical inactivity is associated with increasing the risk of many chronic diseases and conditions, such as obesity, diabetes, hypertension, colon cancer, coronary heart disease and other diseases. However, many adults in the U.S. are inactive. According to the According to the National Center for Chronic Disease Prevention and Health Promotion (2009) 31%, 28% and 25% of the population did not participate in any kind of leisure time physical activity in 1989, 2000 and 2008, respectively.

Recently, the increasing cost of obesity, medical problems and death resulting from diseases caused by obesity necessitated policy makers pay more attention to the problem. At state and local levels, new policies have been initiated and implemented to improve health and diet, and to promote physical activity to maintain healthy body weights (McKinnon et al., 2009). Local government and communities intervene by designing policies and initiatives to support healthy community designs, by promoting facilities such as parks and open space for recreation, through mixed land use development, and by improving access to healthy food.

As obesity and inadequate physical activity become major public health concerns in the U.S., understanding the contributing factors and how those factors interact to contribute to the rise of obesity and the decline of physical activity is important. The problem of obesity is complex and, as indicated earlier, it is the result of the interaction of environmental, socioeconomic, and demographic characteristics, as well as genetics and labor markets. Therefore, understanding the contributing factors to obesity calls for a comprehensive study that examines this wide range of factors. To this end, this study examines multiple factors and how they are interrelated to each other. This study, therefore, will expand the existing literature by investigating the contribution of socioeconomic characteristics along with environmental and other relevant factors in a system of simultaneous equations.

1.2. Objective of the Study

This study is focused on understanding the relationships between obesity, environmental, socioeconomic and labor market factors. It looks at obesity in a comprehensive way, arguing that both environmental and socioeconomic factors contribute to the prevalence of obesity and, thus, need to be studied together. Specific objectives of this study are:

1. To understand the relationship among obesity, physical activity and income
2. To examine the impact of natural amenities, recreational opportunities and aspects of the built environment on obesity and physical activity.
3. To understand how socioeconomic factors relate to obesity and physical activity.
4. To draw relevant empirical conclusions based on the results of the analysis.

1.3. Hypotheses

This study hypothesizes the following:

1. The availability of recreational facilities and natural amenities are negatively related to obesity and positively related to physical activity.
2. Built environment measures such as mixed land use, density, street connectivity, and centeredness are negatively related to obesity and positively related to physical activity.
3. Income as measured by wages and salaries is negatively related to obesity and positively related with physical activity.
4. Physical activity is negatively related to obesity and obesity decreases participation in physical activity.

1.4. Methodology

The primary focus of the study is to investigate the implications of socioeconomic factors, the built environment, natural amenities and recreational opportunities on obesity and physical activity. To understand the relationships, the study uses panel data. A system of equations with endogenous variables for obesity (BMI), physical activity and income are introduced. To account for the simultaneity problems that arise from applying simultaneous equations, the two-stage least squares (2SLS) estimation method is used. In cases where there is a simultaneous relationship, estimating using 2SLS will result in unbiased, consistent and efficient estimates while estimating using ordinary least squares (OLS) will result in biased and inconsistent estimates.

1.5. Overview of the Study Area

This study focuses on 83 U.S. metropolitan areas. These metropolitan areas are part of 44 states and represent nearly half of the U.S. population (Ewing, Pendall and Chen, 2002). However, due to a lack of data for some metropolitan areas, Alaska, Delaware, Idaho, Maine, Mississippi, Montana, New Hampshire, North Dakota, South Dakota, Vermont, Wyoming and West Virginia, are not included in the study. Table 1.1 shows all the included metropolitan areas and their respective states.

Table 1.1. Metropolitan Areas Used in Study

States	Metropolitan Areas
Alabama	Birmingham
Arizona	Tucson, Phoenix
Arkansas	Little Rock, North Little Rock
California	Anaheim, Santa Ana, Oakland, Los Angeles, Long Beach, San Diego, Sacramento, San Jose, Fresno, San Francisco, Riverside, San Bernardino, Oxnard, Ventura, Vallejo, Fairfield, Napa
Colorado	Colorado Springs, Denver

Connecticut	Hartford, New Britain, Middletown, Bristol, New Haven, Waterbury, Meriden, Bridgeport, Stamford, Norwalk, Danbury
Florida	Tampa, St. Petersburg, Clearwater, Jacksonville, Orlando , Fort Lauderdale, Hollywood, Pompano Beach, Miami, Hialeah West Palm Beach, Boca Raton, Delray Beach
Georgia	Atlanta
Hawaii	Honolulu
Illinois	Chicago
Indiana	Indianapolis, Gary, Hammond
Kansas	Wichita, Kansas City
Louisiana	Baton Rouge, New Orleans
Maryland	Baltimore
Massachusetts	Worcester, Fitchburg, Leominster, Springfield, Boston, Lawrence, Salem, Lowell, Brockton
Michigan	Detroit, Grand Rapids
Minnesota	Minneapolis, St. Paul
Missouri	St. Louis
Nebraska	Omaha
Nevada	Las Vegas
New Jersey	Newark, Jersey City
New Mexico	Albuquerque
New York	Syracuse, Albany, Schenectady, Troy, Buffalo, New York, Rochester
North Carolina	Greensboro, Winston, Salem, High Point, Raleigh, Durham
Ohio	Columbus, Cleveland, Cincinnati, Akron, Toledo
Oklahoma	Oklahoma City, Tulsa
Oregon	Portland
Pennsylvania	Pittsburgh, Philadelphia, Allentown, Bethlehem, Easton
Rhode Island	Providence, Pawtucket, Woonsocket
South Carolina	Columbia, Greenville, Spartanburg
Tennessee	Memphis, Knoxville
Texas	Houston, San Antonio, Austin, El Paso, Fort Worth, Arlington, Dallas
Utah	Salt Lake City, Ogden
Virginia	Norfolk, Virginia Beach, Newport News
Wisconsin	Milwaukee
Washington, DC	Washington, DC
Washington	Seattle, Tacoma

1.6. Organization of the Study

This study is organized as follows. Chapter 2 reviews previous studies that dealt with the link between obesity and the physical environment, labor market, and socioeconomic and demographic characteristics of individuals. Chapter 3 outlines the theoretical background of the study. Chapter 4 discusses the methodology and sources of data for the empirical estimation in more detail. Chapter 5 presents and discusses the empirical findings of the study. Lastly, chapter 6 provides a summary, conclusions, limitations, and possible policy recommendations from the research.

Chapter 2

Literature Review

2.1. Definition of the “Built Environment” and “Urban Sprawl”

There is no single accepted definition of the “built environment” and “urban sprawl” that every researcher can agree upon. Different researchers define the built environment differently based on different dimensions and at different geographical scales. One characteristic of the built environment is that it is not stagnant but changes over time in many ways; although the pace of change might not be the same (Handy et al., 2002). For example, expanding new developments or sidewalks or streets may take a few years but the deterioration of these may take a long time. Sprawl can be defined broadly and the definition incorporates multiple aspects of urban areas, such as land use patterns, population density, usage and accessibility of different forms of transportation, etc. Some definitions of the built environment and sprawl are presented in the following section of this chapter.

2.1.1. Definition of the Built Environment

Papas et al. (2007, p. 2) describes the built environment as a place that “encompasses a range of physical and social elements that makes up the structure of a community and may influence obesity.” Handy et al. (2002, p. 65) describes the built environment as including “urban design, land use, and the transportation system, and encompasses patterns of human activity within the physical environment.” According to the authors, urban design refers to “the design of the city and the physical elements within it, including both their arrangement and their appearance, and is concerned with the function and appeal of public spaces;” land use refers to “the distribution of activities across space, including the location and density of different activities, where activities are grouped into relatively coarse categories, such as residential, commercial, office, industrial,

and other activities;” and the transportation system “includes the physical infrastructure of roads, sidewalks, bike paths, railroad tracks, bridges, and so on, as well as the level of service provided as determined by traffic levels, bus frequencies, and the like.” Geurs and van Wee (2004, p. 128) measured the built environment in terms of accessibility, defining accessibility as “the extent to which land-use and transport systems enable (groups of) individuals to reach activities or destinations by means of a (combination of) transport mode(s).” The built environment also can be measured at different scales. For example, Handy et al. (2002) summarize dimensions of the built environment at the neighborhood level as density and intensity, land use mix, street connectivity, street scale, aesthetic qualities and regional structure (Table 2.1).

Table 2.1. Dimensions of the Built Environment

Dimension	Definition	Examples of Measures
Density and Intensity	Amount of activity in a given area	Ratio of commercial floor space to land area, Persons per acre or Jobs per square mile
Land use mix	Proximity of different land uses	Distance from house to nearest store, Share of total land area for different uses, Dissimilarity Index
Street connectivity	Directness and availability of alternative routes through streets	Intersections per square mile of area, Ratio between the straight line distance between two points, Average block length
Street scale	Three-dimensional space along a street as bounded by buildings	Ratio of building heights to street width, Average distance from street to buildings
Aesthetic qualities	Attractiveness and appeal of a place	Percentage of ground in shade at noon, Number of locations with graffiti per square mile
Regional structure	Distribution of activities and transportation facilities across the	Rate of decline in density with distance from downtown,

Source: Handy et al. (2002)

2.1.2. Definition and Measurement of Sprawl

Previous literature provides a wide range of definitions of sprawl that suggest a variety of meanings. The Sierra Club (1999, p. 1) defined sprawl as “low-density development beyond the edge of service and employment, which separates where people live from where they shop, work, recreate and educate-thus requiring cars to move between zones.” They ranked the level of sprawl in metro areas based on population movement to the suburbs from the city, the growth of urban land areas compared to population growth, time spent in traffic and loss of open space. According to their definition of sprawl, Atlanta, Georgia, St. Louis, Missouri, and Washington, D.C. are the most sprawling large metro areas (1 million or more), while Orlando, Florida, Austin, Texas, and Las Vegas, Nevada are the most sprawling medium-sized metro areas (500,000-1,000,000).

Ewing (1997, p. 32) defines sprawl as the combination of three characteristics: “i) leapfrog or scattered development; ii) commercial strip development; and iii) large expanses of low-density or single-use development.” Likewise, the U.S. Department of Housing and Urban Development (1999, p. 33) defined sprawl as “a particular type of suburban development characterized by very low-density settlements, both residential and non-residential; dominance of movement by use of private automobiles, unlimited outward expansion of new subdivisions and leap-frog development of these subdivisions; and segregation of land uses by activity.”

Johnson (2001, p. 721) summarizes sprawl as land use patterns that have the characteristics of “segregated land uses; emphasis on the automobile for transit; a push for growth at the boundary of the metropolitan area; residential and employment densities that are

generally lower than those in further-in suburbs or in the central city.” Likewise, Gillham (2002, p. 8) defined sprawl as “a form of urbanization distinguished by leapfrog patterns of development, commercial strips, low density, separated land uses, automobile dominance, and a minimum of public open space.”

To measure residential sprawl, Burchfield et al. (2006) and Eid et al. (2008) construct a sprawl index by considering the share of underdeveloped land around residential development of the neighborhood. Similarly, Galster et al. (2001) developed a sprawl index using eight variables: density, continuity, concentration, clustering, centrality, mixed use, and proximity. Those variables were identified to capture both the causes and effects of sprawl and used to develop sprawl indices for 13 urban areas. According to the constructed index, Atlanta and Miami are the most sprawling areas in the U.S. while New York City and Philadelphia rank as the least sprawling metro areas. One of the limitations cited by the authors for this index is that it is developed from only the 13 urban areas. Ewing et al. (2003) constructed a sprawl index by considering the average value of a total of four variables: density, land use mix, street accessibility and degree of centering.

Ewing, Pendall and Chen (2002) expanded the definition of sprawl by incorporating both land use and transportation aspects of land use. Ewing, Pendall and Chen (2002, p. 3) defined sprawl as a four dimensional phenomenon and defined it as “a population that is widely dispersed in low density development; rigidly separated homes, shops, and workplaces; a network of roads marked by huge blocks and poor access; and a lack of well-defined, thriving activity centers, such as downtowns and town centers. ” They identified four measures of sprawl: residential density; neighborhood mix of homes, jobs, and services; strength of activity centers

and downtowns; and accessibility of the street network. This sprawl index is the most comprehensive index developed so far as more than 22 variables were included.

2.2. The Impact of the Built Environment and Sprawl on Obesity

According to an urban planning perspective, the current land use pattern observed in the U.S. is characterized by the expansion of sprawl. This expansion is considered one of the primary contributing factors to a sedentary lifestyle and consequent increase in the prevalence of overweight and obesity (Frumkin, 2002; Jackson and Kochtitzky, 2002; Frank, Engelke, and Schmid, 2003). As a result, mixed land use and compact residential density has been perceived as one of the possible alternatives to promote physical activity through walking and cycling for non-work or leisure travel. Places or neighborhoods with mixed land use and high density encourage leisure time walking (Rajamani et al., 2003). Past studies that dealt with urban land use and an individual's choice of mode of transportation advocate for more mixed land use, higher density, and more flexible zoning to promote physical activity such as walking and cycling (Rajamani et al., 2003; Sallis and Glanz, 2006)

A number of studies have linked sprawl and urban planning to the rise of obesity and other diseases (Ewing et al., 2003; Lopez, 2004). Built environments create both opportunities and constraints for physical activities. Pedestrian friendly neighborhoods provide opportunities for residents to be physically active. Areas with compact development and mixed land use patterns encourage walking and bicycling as transportation alternatives while dispersed development patterns increase travel distance, making non-motorized travel more difficult and dangerous (Saelens et al., 2003; Plantinga and Bernell, 2007). The existing designs of commercial and residential areas in many parts of the U.S. seem more favorable for vehicles than

for pedestrians and, hence, create constraints on physical activity and non-motorized transportation alternatives (Jackson and Kochtitzky, 2001).

Frank, Andresen, and Schmid (2004) examined the relationship between land use patterns and obesity. Using survey data of 13 counties around Atlanta, Georgia, the study indicates that more mixed land use led to greater distances walked and lower levels of obesity. The authors also found a positive association between time spent in a car and obesity. In their county level analysis, Lopez-Zetina, Lee and Friis (2006) found a positive correlation between obesity and vehicle mileage traveled and commute time.

Using county-level data, Ewing et al. (2003) found a significant positive relationship between a county's sprawl index and the residents' body mass index (BMI). In the same study, Ewing et al. (2003) found a significant positive relationship between a county-level sprawl index and obesity, and a negative relationship between the sprawl index and minutes of walking. Using metropolitan area data, Ewing, Pendall and Chen (2002) found that living in a sprawling metropolitan area is associated with a higher rate of obesity. Likewise, after accounting for socioeconomic factors such as education, income, race, age and gender, Lopez (2004) found a positive relationship between sprawl and overweight.

Previous studies have found mixed results regarding the relationship between development patterns and obesity. Eid et al. (2008) found no significant relationship between sprawl and BMI. Kelly-Schwartz et al. (2004) using survey data from the National Health and Nutrition Examination Survey of 1988-1994 examined the relationship between obesity and metropolitan sprawl. However, after adjusting for individual level characteristics, the study found no relationship between metropolitan level of sprawl and obesity.

In an attempt to understand whether sprawl causes obesity or obese people choose to live in sprawling neighborhoods, Eid et al. (2008) used National Longitudinal Survey of Youth (NLSY79) data between 1979-1998 to follow an individual's relocation and the subsequent weight change. The study found no evidence that urban sprawl caused obesity. Similarly, by tracking an individual's relocation using NLSY79 data and a county-level sprawl measure Plantinga and Bernell (2007) found individuals who move to a less sprawling county showed significant weight loss for the subsequent two-year period. Both Plantinga and Bernell (2007) and Eid et al. (2008) suggested that, because individuals sort themselves when deciding where to reside, obese people may decide to locate in sprawling locations where the physical environment forces them to depend more on automobiles or limit the opportunities to be physically active. Non-obese individuals, however, may decide to locate in places where there are more opportunities to be physically active. As a result, improving the neighborhood environment in sprawling areas to be more walking and cycling friendly by incorporating sidewalks, trails, parks, etc. may not have the desired impact in reducing obesity (Eid et al., 2008).

2.2.1. Sprawl and Its Impact on Travel Behavior and Obesity

In sprawling regions, cars are driven long distances and people are less likely to take other modes of transportation such as bus, train, bike or walking (Ewing, Pendall and Chen 2002). Frank, Andresen and Schmid (2004) indicate that people who spend more time in the car are more likely to be obese and have a 6 percent greater chance of being obese for every 60 minutes they spend in the car.

In a cross-country comparison, Bassett et al. (2008) studied the relationship of transportation selection and obesity in Europe, North America and Australia. The study found that countries that used active transportation (defined as percentage of trips taken by walking,

bicycling, and public transit) have the lowest obesity rates while countries that use cars for their transportations have the highest rates of obesity.

2.3. The Relationship between Socioeconomic Factors and Obesity

The prevalence of obesity has increased in all ethnic, income, education and age groups. The extent of its prevalence, however, varies across race, gender and income level, and it is more evident in the most disadvantaged groups of society (Paeratakul et al., 2002; Drewnowski, 2004; and Ogden et al., 2010). Higher rates of obesity in the U.S. were observed among Black, Hispanic, lower-income, and less-educated groups (Paeratakul et al., 2002). Drewnowski (2004) found the highest levels of obesity in areas with high poverty levels and less education. The National Center for Health Statistics (2012) reported that adolescent obesity rates were higher in poor households, and lower income people were more likely to be obese than higher income people. The Healthy People 2010 (National Center for Health Statistics, 2000) study also reported African-American women were relatively more obese than their White counterparts.

The 2005-2008 data of the CDC National Health and Nutrition Examination Survey (NHANES) indicates that the prevalence of obesity among men is generally similar at all income levels (Ogden et al., 2010). However, among non-Hispanic black men and Mexican-American men, those with higher incomes are more likely to be obese than those with low incomes (Ogden et al., 2010). The survey also indicated that lower income women are more likely to be obese than higher income women. Among men, there is no significant relationship between obesity and education. However, among women, those with a college degree are less likely to be obese than women with no college degree. The survey also shows that the prevalence of obesity among men and women with college educations is less than among those with some college education.

2.4. The Labor Market and Obesity

The relationship between obesity and the labor market can be examined by looking at the impact of obesity on both the demand and supply sides of the labor market. Greve (2008) explained the effect of body weight on employment in two ways. First, the productivity of obese people may be lower, or, in terms of work motivation and incentives for labor market participation, obese and overweight people may have less motivation and lower incentives than healthy-weighted people or people with ideal weights. Second, from the labor demand side, obese people may face discrimination and prejudice in the employment decision-making process. The challenges obese people face in the labor market may be due to preferences of employers and/or customers for non-obese employees, rather than from the performance of an obese worker (Becker, 1973). A number of experimental studies have indicated that discrimination based on weight exists at every stage of the employment process, starting from the decision of hiring to wage determination and promotion (Puhl and Brownell, 2001; Rooth, 2007). Employers are less inclined to hire obese and overweight people in order to avoid associated health insurance and other health-related expenses (Greve, 2008).

Tunceli, Kemeng, and Williams (2006) found that obesity resulted in lower employment participation, constraints on work activities and, consequently, loss of productivity. Individual problems that resulted from being obese or overweight affect labor market participation either in the form of reduced productivity or increases in the number of missed work days (Greve, 2008). Obese people took more sick days and had a higher rate of absenteeism, and the consequent loss of productivity led to discrimination against obese people in the workplace (Jensen, Greve, and Tranaes, 2005).

In addition to limitations at the workplace, day to day activities of obese individuals may also be impacted, depending on the extent and severity of the obesity (Sturm, Ringel, and Andreyeva, 2004). The authors indicated that the probability of limitations on daily activities increased by 50% and 300% for moderately obese and severely obese people, respectively. However, a study conducted by Cawley (2000), using NLSY79 data on females, suggested that body weight did not cause employment disability. Tunceli, Kemeng, and Williams (2006), using data from the Panel Study of Income Dynamics, examined the impact of obesity on employment and work limitations utilizing a probit model. The study found that obese men had lower levels of employment and that obese and overweight women had higher levels of self-reported limitation in the workplace. Likewise, using survey data from England, Morris (2006) found that as an individual's BMI increased his/her chance of obtaining employment decreased.

2.5. Obesity and Earnings

Studies have shown mixed results regarding the relationship between body weight and wages. For example, Cawley (2004) examines the impact of weight on wages received using Ordinary Least Square (OLS) estimation. The study found that overweight white, black and Hispanic females and Hispanic males earned less than their less heavy counterparts. Gortmaker et al. (1993) found that an overweight male adolescent has 9% less household income, while an overweight female adolescent has 22% less household income. Averett and Korenman (1999) found that the earnings of white obese women were 17% less than those of white women with ideal weight.

Register and Williams (1990) found that obesity had no effect on earnings of male workers but the earnings of obese females was 12% less than those for females with ideal weight. Mitra (2001) found that obesity had no effect on the earnings of obese males while the

wages of obese females in managerial and professional occupations decreased by 2% for every one pound increase in weight. Hence, even though previous studies have examined the implications of obesity on the labor market and wages, it is not clear whether there is a causal relationship between obesity and wages (Cawley, 2004).

2.6. Methodological Approaches of Previous Studies

Previous studies explored relationships among obesity and/or weight, land use patterns and the built environment by applying a number of estimation methods, such as simple linear regression, discrete analysis using probit and/or logit, a system of simultaneous equations, and spatial econometric analysis.

Ewing et al. (2003) employed a single equation linear model to examine the relationship between BMI and a sprawl index. Frank, Andresen, and Schmid (2004) used survey data from the Atlanta region and a discrete analysis. Using a logistic regression, they estimated the odds of being obese based on socio-demographic variables. Plantinga and Bernell (2007) employed both linear regression and a probit model to examine the relationship between the residential choice of new movers and BMI.

Using panel data, Eid et al. (2008) accounted for both unobserved and time invariant factors that contributed to the propensity to be obese. The authors used sex and race to control for time invariant factors to avoid biased and inconsistent results. To capture unobserved time invariant variables they used a first difference equation with respect to time. Chou, Grossman, and Saffer (2004) employed both ordinary least squares estimation and a fixed effects specification using individual level data from the Behavioral Risk Factor Surveillance System (BRFSS) to control unmeasured time invariant variables.

Plantinga and Bernell (2005) used a spatial theoretical framework developed for urban land markets to capture the impacts of consumer preferences, relocation decisions and housing prices on obesity. They argued that previous literature failed to reflect consumers' preferences on weight, residential densities, and commuting costs as simultaneous outcomes. Therefore, studies that did not consider the interaction of obesity and/or overweight and development density as a simultaneous interaction exhibited a simultaneity bias problem. In their follow up study, Plantinga and Bernell (2007) consider sprawl as an endogenous factor in determining obesity and/or overweight. They argued that if residential preferences are affected by weight, then urban sprawl is not an exogenous factor for weight gain and/or obesity. Instead, the association between sprawl and overweight /obesity is a two way relationship in which obesity and weight is a function of land use/urban sprawl (and other factors), and urban sprawl/land use is a function of weight/obesity (and other factors).

Using a three equation system, Rosenberger et al. (2005) examined the interaction among health care expenditures, obesity and physical inactivity. The authors specified healthcare expenditures as a function of physical inactivity and other variables; physical inactivity as a function of obesity; and obesity as a function of physical inactivity. They argued that obesity and physical inactivity are simultaneously determined since obesity creates barriers for people to be physically active, and in turn, all things constant, physical inactivity increases the risk of being obese or overweight. They employed a two-stage least squares (2SLS) estimation method to account for the simultaneity problem. In cases where there is a simultaneous relationship, estimating using 2SLS will result in unbiased, consistent and efficient estimates while estimating using ordinary least squares (OLS) will result in biased and inconsistent estimates.

Chapter 3

Theoretical Framework

3.1. Introduction

To analyze health outcome choices by households, past studies utilized time allocation choices through household utility maximization frameworks in developing a theoretical framework. For example, Becker (1965) introduced time allocation of non-work activities using a basic household utility maximization framework. The study assumed that individuals are both consumers and producers where they produce goods and services by employing both inputs and time. To solve the utility maximizing problem, the theoretical framework incorporated goods and time allocated in producing those goods.

Grossman (1972) developed a theoretical model that reflects demand for health by employing utility maximization theory. Grossman (1972) examined the demand for health as a consumption and/or investment good. He suggested that as a consumption good, consumers consider good health in their utility maximization problem like other goods and services. As an investment factor, consumers demand good health since it increases time availability for other activities, hence, it reduces the monetary value of time lost for those activities.

Extending demand for health by focusing on weight and weight management as an extension of demand for health, Lakdawalla and Philipson (2002) introduced weight as one of the components of a utility maximization framework to understand the dynamic nature of weight management. The utility maximization framework for this study is developed by considering the demand for health developed by Grossman (1972), non-work time allocation by Becker (1962) and a weight management framework developed by Lakdawalla and Philipson (2002).

Individuals maximize their utility by consuming food, other bundles of goods and services, as well as weight and level of physical activity or exercise. In achieving the maximum utility from consumption of those commodities, individuals are constrained by a fixed budget, time and biological factors. The i^{th} individual's utility maximization problem subject to a budget constraint is expressed as;

$$(3.1) \quad \text{Max}V_i = V_i(C, F, W, P) \quad \text{subject to} \quad Y = P_C C + P_F F + P_W W + P_P P$$

where V is utility of individual i , C is consumption of goods and services, W is weight, P is level of physical activity or exercise and F is consumption of food items. Food consumption is classified into healthy food (H) and unhealthy food (U). The income, Y , constraint takes into account the prices for each of these; P_C for the price of consumption goods, P_F for the price of food, P_W for the “price” of weight and P_P for the “price” of physical activity. Price of weight is measured by the expense of the individuals for diagnosis and treatment of health problems associated with overweight and obesity. Similarly, price of physical activity is measured by the expense of the individual associated with engaging in physical activity, for example, gym or health club membership fee, personal trainer fee, coach fee, expense on necessary equipment and accessories. Biological factors or genetics are considered as one of the determining factors of weight. However, due to the difficulty and unavailability of data, biological factors are not explicitly considered.

Utility increases with the consumption of goods and services, food and physical activity. However, utility increases with weight until the individual reaches his/her ideal weight (W). If the weight of the individual is greater than the ideal weight, utility decreases with “consumption” of additional weight. The ideal weight of an individual is subjective and differs from person to person; however, it is assumed to be in the range of normal weight or Body Mass Index (BMI).

The market provides a variety of food choices, both healthy and unhealthy. Consumers, therefore, have a choice to adopt a healthy diet (H) or an unhealthy diet (U). Adoption of a healthy diet consists of consumption of fruits, vegetables, lean proteins and generally more nutrient dense food items. Whereas, unhealthy diets consist of less fruits, vegetables, lean meats and fish, and a high proportion of fat and energy dense foods.

Weight outcome is mainly determined by the balance of calorie intake and calorie expense. Calorie expense can result from normal body functioning and through physical activity or exercise. Weight is, therefore, a result of food consumption (calorie intake) and physical activity (calorie expense). Hence, weight is strictly increasing with food consumption above the maintenance requirement, $\frac{\partial W}{\partial F} > 0$, and decreasing with physical activity, $\frac{\partial W}{\partial P} < 0$.

Individuals have fixed time available and they allocate their time rationally into different activities. Available time is denoted by T , time allocated to work is denoted by K and time for non-work related activities is denoted by N . Therefore, available time is the summation of time spent on work and non-work related activities, $T = K + N$ where $K = T - N$. After rearranging and multiplying both sides by the wage rate (w), wage income is $wK = w(T - N)$; w reflects the opportunity cost of time spent on non-work related activities.

The expanded form of equation 3.1 with income decomposed into wage and non-wage income is:

$$(3.2) \quad V = V(C, F, W, P) + \lambda(I + w(T - N) - P_C C - P_F F - P_W W - P_P P)$$

Total income (Y) is decomposed into wage income (wK) and non-wage income (I). Non-work related activities can be decomposed into time spent on leisure activities (L), travel from work to home and vice versa (R), sleeping (S), and household production (Q). Time spent on leisure

activities is disaggregated into time spent on physical activity or exercise (E) and leisure time spent on non-exercise activities (G). Equation (3.1) can be further expanded by decomposing food consumption into healthy food consumption (H) and unhealthy food consumption (U) as shown in equation (3.3). Similarly, price of food P_F is the summation of price of healthy food P_H and unhealthy food P_U .

Physical activity is a function of weight of the individual and availability and accessibility of recreational facilities and opportunities. Individuals with normal weight may engage in physical activity due to health benefits and/or to maintain their current weight. On the other hand, it can be argued that individuals with ideal weight may be less physically active because they may not need any weight loss, and/or very small weight loss is needed. Overweight and obese individuals may prefer to be more physically active to lose weight to reach their normal or ideal weight level and/or for health-related purposes. It is possible, however, to expect that individuals with high weight may not be physically active due to physical limitations, health problems or other reasons.

The Lagrangian setup for equation (3.1) is:

$$(3.3) \quad V = V[C, F(H, U), W(F(H, U), P), P(W, O)] + \lambda(Y - P_C C - P_H H - P_U U - P_W W - P_P P)$$

The first order conditions for utility maximization can be expressed as:

$$(3.4) \quad \frac{\partial V}{\partial C} = \frac{\partial V(\cdot)}{\partial C} - \lambda P_C = 0$$

$$(3.5) \quad \frac{\partial V}{\partial F} = \frac{\partial V(\cdot)}{\partial F} + \frac{\partial V(\cdot)}{\partial W} \frac{\partial W(\cdot)}{\partial F} - \lambda P_F = 0$$

$$(3.6) \quad \frac{\partial V}{\partial W} = \frac{\partial V(\cdot)}{\partial W} + \frac{\partial V(\cdot)}{\partial P} \frac{\partial P(\cdot)}{\partial W} - \lambda P_W = 0$$

$$(3.7) \quad \frac{\partial V}{\partial P} = \frac{\partial V(\cdot)}{\partial W} \frac{\partial W(\cdot)}{\partial P} + \frac{\partial V(\cdot)}{\partial P} - \lambda P_p = 0$$

$$(3.8) \quad \frac{\partial V}{\partial \lambda} = Y - P_c C - P_f F - P_w W - P_p P = 0$$

$$(3.9) \quad \frac{\partial V}{\partial H} = \frac{\partial V(\cdot)}{\partial F} \frac{\partial F(\cdot)}{\partial H} + \frac{\partial V}{\partial W} \frac{\partial W(\cdot)}{\partial F} \frac{\partial F(\cdot)}{\partial H} - \lambda P_H$$

$$(3.10) \quad \frac{\partial V}{\partial U} = \frac{\partial V(\cdot)}{\partial F} \frac{\partial F(\cdot)}{\partial U} + \frac{\partial V}{\partial W} \frac{\partial W(\cdot)}{\partial F} \frac{\partial F(\cdot)}{\partial U} - \lambda P_U = 0$$

Solving and rearranging from equation (3.4) to equation (3.8) yields:

$$(3.11) \quad \begin{aligned} \frac{\partial V(\cdot)}{P_c \partial C} &= \frac{\partial V(\cdot)}{P_H \partial F} \frac{\partial F(\cdot)}{\partial H} + \frac{\partial V}{P_H \partial W} \frac{\partial W(\cdot)}{\partial F} \frac{\partial F(\cdot)}{\partial H} = \frac{\partial V(\cdot)}{P_U \partial F} \frac{\partial F(\cdot)}{\partial U} + \frac{\partial V}{P_U \partial W} \frac{\partial W(\cdot)}{\partial F} \frac{\partial F(\cdot)}{\partial U} = \\ &= \frac{\partial V(\cdot)}{P_w \partial W} + \frac{\partial V(\cdot)}{P_w \partial P} \frac{\partial P(\cdot)}{\partial W} = \frac{\partial V(\cdot)}{P_p \partial W} \frac{\partial W(\cdot)}{\partial P} + \frac{\partial V(\cdot)}{P_p \partial P} = \lambda \end{aligned}$$

Equation (3.11) implies that the marginal utility derived from the last dollar spent should be the same for all commodities, including weight and physical activity.

3.2. Implicit Demand Functions

Solving the above first order conditions and applying the implicit function theorem, the demands for the choice variables are derived as follows.

$$(3.12) \quad C^* = C(P_c, P_f, P_w, P_p, Y)$$

$$(3.13) \quad H^* = H(P_c, P_H, P_U, Y, P_w)$$

$$(3.14) \quad U^* = U(P_c, P_H, P_U, Y, P_w)$$

$$(3.15) \quad W^* = W(P_c, P_H, P_U, Y, P_w)$$

$$(3.16) \quad P^* = P(P_w, P_p, Y)$$

Equation (3.13) implies that demand for healthy food is a function of the price of other goods and services P_c , its own price P_H , price of unhealthy foods P_U , total income Y and price of

weight P_W . Similarly, demand for unhealthy foods is determined by price of other goods and services P_C , its own price P_U , price of healthy foods P_H , total income Y and price of weight P_W as expressed in equation (3.14). Equation (3.15) indicates that weight is determined by price of goods and services P_C , price of food, both healthy and unhealthy P_H and P_U , and the price of weight P_W . Lastly, equation (3.16) indicates that demand for physical activity is determined by the price of weight P_W , cost of physical activity P_P (such as the expense of a gym membership or equipment purchased to enable physical activity, like a bicycle) and total income Y .

3.3. Comparative Static Analysis

Using comparative statics, the implications of changes in total income and the opportunity cost of time measured by wage income (w) on the weight and physical activity decisions are analyzed. Details of the comparative static calculations are in the appendix.

The Impact of a Change in Total Income on Weight ($\frac{dW}{dY}$) and Physical Activity ($\frac{dP}{dY}$)

The impact of a change in total income on weight differs at different weight and different income levels. In areas with low incomes overall, the initial implication of income growth is to increase weight for all individuals, who were likely underweight. At the individual level, an increase in income could cause an increase in weight or a decrease in weight depending on the initial weight of the individual. For underweight individuals, income growth may lead to weight gain. For overweight and/or obese individuals an increase in income could result in weight loss. The relationship between income and weight suggests that there is a negative income effect for obese and overweight individuals and a positive income effect for underweight individuals. An increase in income may also have no impact on an individual's weight. For example, individuals at an ideal weight may use any increase in income to increase their purchase of consumption goods and maintain levels of food consumption and physical activity that also maintain their

ideal weight. Cawley, Han and Norton (2009) suggested that income could either increase or decrease weight. An increase in income could lead to weight gain if food and a sedentary lifestyle are considered normal goods. Similarly, health and less weight could be considered normal goods leading to investing more time and money on weight loss as income increases.

As income increases, it is expected that the level of physical activity would also increase, up to a certain level. Income growth may mean individuals have enough resources to spend on physical activity; for example, joining a gym, or purchasing necessary equipment and accessories for exercise and other recreational physical activities. Income growth may also lead to an increased preference for leisure time. Assuming that individuals allocate their leisure time to non-sedentary activities, there would be a positive relationship between income and physical activity.

The Impact of the Opportunity Cost of Time on Weight ($\frac{dW}{dw}$) and Physical Activity ($\frac{dP}{dw}$).

A change in the opportunity cost of time (w) spent on non-work related activities may have different implications for low income and high income groups. For low income individuals, an increase in their wage means a higher opportunity cost of non-work time such that less time is made available for leisure, which in turn suggests that less time is available for physical activity. In other words, the higher wage encourages them to work more leading to less leisure time and less physical activity. However, for higher income groups, even though the opportunity cost of non-work related time increases as income increases, they may choose leisure over work because they now need to work fewer hours to maintain the same income level. Assuming that part or most of their leisure time is spent on activities that involve exercise, for higher income groups as the opportunity cost of leisure time increases, the level of physical activity also increases. Hence,

the opportunity cost of leisure may have positive impacts on physical activity for high income groups and negative impacts on low income groups.

As the opportunity cost of leisure time increases, the time spent on physical activity may decrease. This suggests that less leisure time is spent on physical activity and exercise. Given that weight loss is an increasing function of exercise, higher opportunity cost of leisure time may imply lower or no time devoted for physical activity, hence, a negative relationship between weight loss and higher opportunity cost of leisure. Similarly, lower opportunity cost of leisure time may suggest that more time is available for leisure activities. Assuming that part of the leisure time is allocated on physical activity, lower opportunity cost of leisure time may indicate higher level of physical activity and weight loss.

Chapter 4

Empirical Models and Data Description

4.1. Introduction

This chapter focuses on model specification for empirical estimation following the theoretical framework. It begins with a general discussion of methodology in section 4.2. In section 4.3 the general model is introduced. The random effect model is specified in section 4.4. Section 4.5 explains the empirical model specification of the ordered probit estimation method and section 4.6 discusses the data and defines the variables used for estimation.

4.2. Methodology

For empirical estimation panel data of six years is used in this study. Panel data consist of repeated measurements of cross-sections over a period of time. One of the advantages of using panel data is the ability to control for unobserved heterogeneity among the individual subjects. In cross section data, there are numerous unobserved variables that affect the individuals, and not accounting for them may lead to biased estimation (Kennedy, 2003). Another advantage of using panel data for empirical estimation is it increases the variability of the data by increasing the sample size through combining cross-sectional information overtime. This increase in the number of observations may lessen the problem of multicollinearity and may lead to efficient estimation (Kennedy, 2003).

To account for unobserved heterogeneity or time-invariant factors, two estimation approaches are often suggested: fixed effect (FE) and random effect (RE) methods. Section 4.2.1 discusses the FE and RE approaches.

A general specification of panel data with unobserved effects is:

$$(4.1) \quad Y_{it} = \lambda + \beta' X_{it} + \alpha_i + \varepsilon_{it} \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

where X_{it} is a $1 \times K$ vector of explanatory variables. These variables are observable and change across time t and individual i ; α_i denotes the unobserved effect or unobserved heterogeneity of individuals. It is also referred as an individual or group effect. The unobserved effect is assumed to be constant over time. ε_{it} is the error term that is assumed to be independently and identically distributed (i.i.d.).

4.2.1. Fixed Effect and Random Effect Approaches

Panel data can be estimated by using a FE or RE method. The choice of model is based on the assumption of unobserved effects (α_i). Fixed effect models assume the unobserved effect is fixed, in which case this parameter is estimated. The fixed effect model can be specified as:

$$(4.2) \quad Y_{it} = (\lambda + \alpha_i) + \beta' X_{it} + \varepsilon_{it} \quad i = 1, 2, \dots, N; \quad t = 1, 2, \dots, T$$

where Y_{it} denotes the dependent variable for the i^{th} individual's metropolitan county and t time period. β is a $K \times 1$ vector of parameters where K is the number of explanatory variables for the i^{th} observation and time t (denoted by X_{it}). α_i is a scalar of individual fixed effects which capture the unobserved component. The error term for the i^{th} observation and t time is denoted by ε_{it} and is assumed to be i.i.d.

In a FE model, it is assumed that the unobserved individual or group characteristics are correlated with observed explanatory variables. That is covariance (X_{it}, α_i) is not equal to 0. λ and α_i are multiple intercepts to be estimated. One way of estimating α_i along with β is to introduce N dummy variables for each cross section observation. The introduction of dummy variables in equation (4.3) controls for the unobserved individual or group effects. The new specification with dummy variables is referred to as a Least Squares Dummy Variables (LSDV) model and can be formulated as:

$$(4.3) \quad y = X\beta + D\alpha + \varepsilon$$

where y is an $NT \times 1$ vector of the dependent variable, D represents the dummy variables of an $NT \times N$ matrix, α is an $N \times 1$ vector of dummy variables (vector of intercepts) to be estimated, β is a $K \times 1$ vector of parameters to be estimated, X is an $NT \times K$ matrix of K regressors and ε is an $NT \times 1$ vector of error terms.

In a random effect (RE) approach α_i is assumed to be a random variable. Random effect implies no correlation between the explanatory variable and the unobserved effects, Covariance $(X_{it}, \alpha_i) = 0, t = 1, 2, \dots, T$. The random effect model can be specified as:

$$(4.4) \quad Y_{it} = \lambda + \beta' X_{it} + \omega_{it} \quad \omega_{it} = \alpha_i + \varepsilon_{it}$$

where ω_{it} is the composite error, α_i is an individual-specific random effect, ε_{it} is combined time series and cross section error component and λ is the intercept. Unlike the fixed effect model, which considered the intercept for each individual or cross-sectional unit, random effect models consider λ as the mean value of all the individual intercepts (Gujarati, 2003).

4.2.2. Ordered Probit Model

A number of references recognize that linear regression is an inappropriate estimation method when the dependent variable is presented in categorical form. The appropriate theoretical model in such a situation is the ordered probit or ordered logit model (Greene, 2000). Over the last three decades these types of models have been widely used as a methodological framework for analyzing ordered data. Ordered data is a form of categorical data in which the information has an intrinsic order.

Suppose the underlying relationship can be characterized as:

$$(4.5) \quad y^* = \alpha + \beta x + \varepsilon$$

where y^* is the exact but unobserved dependent variable, x is the vector of independent variables, and β is the vector of regression coefficients to be estimated. Further, suppose that y^* is not observed but the categories of responses are observed:

$$(4.6) \quad y^* = \begin{cases} y = 0 & \text{if } y^* \leq \mu_1 \\ y = 1 & \text{if } \mu_1 \leq y^* \leq \mu_2 \\ y = 2 & \text{if } \mu_2 \leq y^* \leq \mu_3 \\ \dots\dots \\ = J & \text{if } y^* \geq \mu_J \end{cases}$$

where μ is the unknown threshold parameter to be estimated with β . Threshold parameters determine the estimation for different observed values of y . These threshold parameters can be interpreted as intercepts in equation (4.5). It is assumed that ε is normally distributed with an expected value of zero and variance of one. The marginal effects of the independent variables on the probabilities of Y are not equal to the coefficients. The general probability of ordered probit is specified below.

$$(4.7) \quad \begin{aligned} \text{prob}(Y = 0 | X) &= \Phi(-x' \beta) (\mu_0 = 0) \\ \text{prob}(Y = 1 | X) &= \Phi(\mu_1 - x' \beta) - \Phi(-x' \beta) \\ \text{prob}(Y = 2 | X) &= \Phi(\mu_2 - x' \beta) - \Phi(\mu_1 - x' \beta) \\ \dots\dots \\ \text{prob}(Y = J | X) &= 1 - \Phi(\mu_{J-1} - x' \beta) \end{aligned}$$

For those probabilities the marginal effects of the changes in the independent variables are:

$$\begin{aligned}
& \frac{\partial \text{prob}(y = 0 | X)}{\partial X} = \phi(-x' \beta) \beta \\
& \frac{\partial \text{prob}(y = 1 | X)}{\partial X} = [\phi(-x' \beta) - \phi(\mu_1 - x' \beta)] \beta \\
(4.8) \quad & \frac{\partial \text{prob}(y = 2 | X)}{\partial X} = [\phi(\mu_1 - x' \beta) - \phi(\mu_2 - x' \beta)] \beta \\
& \dots\dots \\
& \frac{\partial \text{prob}(y = J | X)}{\partial X} = \phi(\mu_{J-1} - x' \beta) \beta
\end{aligned}$$

The analytical framework presented above is applied to data described below to determine what factors contribute to an individual's probability of participation in physical activity.

4.3. General Model Specification

In understanding the relationship between obesity and overweight and physical activity one of the concerns is that weight could be endogenous to taking part in physical activity. For example, obesity and overweight may have an adverse impact on participating in physical activities. If this is the case, and obesity and overweight is treated as an exogenous variable in the physical activity equation, the estimated effect of obesity and overweight is likely to be biased. In addition, the endogeneity of obesity and overweight to physical activity may result from both simultaneity and unobserved heterogeneity, suggesting a simultaneous equation model is required to obtain efficient estimates of the impact of obesity and overweight on physical activity. Similarly, obesity and overweight may be endogenous to income. Weight may have negative implications on income through low productivity or discrimination (Cawley, 2004). In this case, considering income as an exogenous variable may result in biased estimation. Using a simultaneous equation model, both the reverse effect of BMI on income and the endogeneity can be estimated and efficient estimates can be obtained.

In the absence of a simultaneity problem, Ordinary Least Squares (OLS) estimation provides consistent and efficient estimates. However, in the presence of simultaneity, OLS does not produce efficient estimates; instead a Two Stage Least Squares (2SLS) method is a more efficient and consistent estimator. A Durbin-Wu-Hausman test is used to determine whether OLS or 2SLS is needed to estimate the BMI and Income equations. The null hypothesis states that there is no endogeneity and both the OLS and 2SLS estimators are consistent, while the alternative hypothesis states that endogeneity exists and only the 2SLS estimator is consistent. The Hausman specification test was conducted and the null hypothesis rejected in both cases: weight and income, and weight and physical activity. The alternative hypothesis, that endogeneity exists is accepted. Thus 2SLS is used in estimating the BMI and Income equations.

To explore the interaction of recreational opportunities based on natural amenities, socioeconomic factors, and the built environment on physical activity and obesity and overweight, the general model used as a basis for the following empirical models is presented. Equations 4.9 through 4.11 specify the general framework.

$$(4.9) \quad BMI_{it} = f(PHYA_{it}, WGSL_{it}, X_{it}, A_{it}, B_{it})$$

$$(4.10) \quad PHYA_{it} = f(BMI_{it}, WGSL_{it}, X_{it}, A_{it}, B_{it})$$

$$(4.11) \quad WGSL_{it} = f(BMI_{it}, PHYA_{it}, X_{it}, O_{it})$$

where BMI_{it} is the body mass index for individual i at time t , $PHYA_{it}$ is the frequency and intensity of physical activity of individual i at time t , $WGSL_{it}$ is total income from wages and salaries of individual i at time t , X_{it} is a vector of the individual's demographic characteristics at time t , A_{it} is a vector of natural amenities and outdoor recreational opportunities that are available

to individual i at time t , B_{it} is a vector of built environment measures related to the residential location of individual i at time t , and O_{it} are economic indicators of the county where individual i resides at time t . Vector X_{it} is composed of socio-demographic variables such as gender, race, level of education, marital status, and number of dependents in the household of individual i . Natural amenity and outdoor recreational opportunities in vector A_{it} consist of land-based, water-based, and winter-based outdoor recreational opportunities, along with general climate of the county. Similarly, the built environment vector represents variables such as mixed land use, street connectivity, density and centeredness. Vector O_{it} includes the unemployment rate of the county at time t and population change of the county at time t .

4.4. Random Effect Model Specification

In section 4.2.1, the theoretical framework of panel estimation methods, Fixed Effect (FE) and Random Effect (RE) are presented for the purpose of comparing the two approaches. However, in this study only the random effect approach is used for empirical estimation. In the presence of time-invariant variables, the fixed effect approach is not generally used because the coefficients of time-invariant variables cannot be estimated. Equations (4.12) through (4.14) are specified to capture the random effect model. The RE model is expressed as:

$$(4.12) \quad BMI_{it} = \lambda + \beta_1 PHYA_{it} + \beta_2 WGSL_{it} + \beta_3 EDUC_{it} + \beta_4 RACE_{it} + \beta_5 GNDR_{it} + \beta_6 MARRD_{it} + \beta_7 DPNDT_{it} + \beta_8 STCN_{it} + \beta_9 CNTD_{it} + \beta_{10} MXLU_{it} + \beta_{11} DNST_{it} + \beta_{12} CLMT_{it} + \beta_{13} LAND_{it} + \beta_{14} WATR_{it} + \beta_{15} WINTR_{it} + (\varepsilon_{it} + \beta_i)$$

$$(4.13) \quad PHYA_{it} = \lambda + \alpha_1 BMI_{it} + \alpha_2 WGSL_{it} + \alpha_3 EDUC_{it} + \alpha_4 RACE_{it} + \alpha_5 GNDR_{it} + \alpha_6 MARRD_{it} + \alpha_7 DPNDT_{it} + \alpha_8 STCN_{it} + \alpha_9 CNTD_{it} + \alpha_{10} MXLU_{it} + \alpha_{11} DNST_{it} + \alpha_{12} CLMT_{it} + \alpha_{13} LAND_{it} + \alpha_{14} WATR_{it} + \alpha_{15} WINTR_{it} + (v_{it} + \alpha_i)$$

$$(4.14) \quad WGSL_{it} = \lambda + \theta_1 BMI_{it} + \theta_2 PHYA_{it} + \theta_3 EDUC_{it} + \theta_4 GNDR_{it} + \theta_5 RACE_{it} + \theta_6 POPCG_{it} + \theta_7 UNEMP_{it} + (\mu_{it} + \theta_i)$$

BMI is specified as a function of physical activity (PHYA), income measured by wages and salary (WGSL), educational attainment (EDUC), and demographic characteristics: gender (GNDR), race (RACE), marital status (MARRD), and number of dependents in the household (DPNDT). In addition, BMI is specified as a function of built environment variables: street connectivity (STCN), centeredness (CNTD), mixed land use (MXLU) and density (DNST); and natural amenities and outdoor recreational opportunities: climate (CLMT), land-based outdoor recreational facilities (LAND), water-based outdoor recreational facilities (WATR) and winter-based outdoor recreational facilities (WINTR).

BMI is expected to be negatively associated with physical activity, suggesting that as the level of physical activity increases, weight decreases and results in lower BMI, and as BMI decreases the level of physical activity goes up. The impact of income (WGSL) on BMI is expected to be negative. As income increases, it is expected that individuals will have more resources to conduct a healthy lifestyle and ultimately achieve normal weight. As income increases, individuals may have the ability to purchase more healthy food items, afford to join a gym or other health clubs and, hence, become more physically active. Education (EDUC) is expected to have a negative impact on weight, which suggests that as education level increases people become more aware of the overall health implications of obesity and, thus, they attempt to maintain normal weight or BMI.

Previous studies, for example, Lopez (2004) and Ewing et al. (2003) found a positive relationship between sprawl and obesity. Sprawling neighborhoods usually exhibit the characteristics of low housing density, separation of schools, work places and shopping centers from residences, and lack of pedestrian friendly sidewalks that lead to dependence on

automobiles for travel and less physical activity. Thus, following previous studies, it is expected that STCN, CNTD, MXLU and DNST are negatively and indirectly related to obesity. Similarly, it is expected that natural amenities and outdoor recreational opportunities (CLMT, LAND, WATR, and WINTR) are negatively related to obesity. The availability of recreational opportunities and favorable weather are likely to encourage people to be physically active.

In equations (4.13), physical activity (PHYA) is specified as a function of BMI and environmental characteristics such as built environment (STCN, CNTD, MXDU and DNST) and natural amenities and recreational opportunities (CLMT, LAND, WATR, and WINTR). It is expected that BMI can be related either positively or negatively with physical activity (PHYA). Obese and overweight people may find it difficult to be physically active because of health and/or physical limitations. On the contrary, an increase in BMI may encourage people to be more physically active to lose weight. Similarly, built environment measures, STCN, CNTD, MXLU and DNST, because they are actually measuring the absence of sprawl, are expected to have positive implications for physical activity. A favorable climate, along with outdoor recreational opportunities, is expected to have a positive impact on physical activity.

An individual's income is specified as a function of BMI and demographic characteristics education (EDUC), gender (GNDR), and race (RACE) along with economic characters of the county in which s/he lives: unemployment rate (UNEMP) and change in population (POPCG). It has been observed that, due to the negative impacts of overweight and obesity on work, for example weight-related health problems and discrimination, obese and overweight people face difficulty penetrating the labor market. Following the results of Tunceli, Kemeng, and Williams (2006) and Greve (2008), BMI is expected to have a negative impact on income (WGSL).

Education, on the other hand, should increase income. It is also expected that there is a wage difference between male and female workers and across different races.

In panel data analysis, a fixed effect method of estimation provides consistent estimates for only time-varying regressors but not time-invariant regressors (Cai, 2010). Even if the BMI equation and income equation can be estimated using either a fixed effect or random effect model, the latent nature of the dependent variable in the data and the simultaneous equation model necessitate the assumption of a random effect for the time-invariant components (Hsiao, 2003). The violation of this assumption may lead to biased estimates. Even if literature supports the use of a random effect model in the presence of time-invariant regressors, a Hausman test should be conducted to choose between fixed effect and random effect estimation. In this study, the Hausman test could not be estimated in SAS due to the presence of time-invariant variables. Therefore, only the random effect method is considered.

4.5. Ordered Choice Models

A set of data is considered as ordinal if the value or observations belonging to it can be ranked or put in order. In the case of this section, the values for physical activity are ranked from 0 to 2. Zero (0) is assigned if the individual does not engage in any physical activity, one (1) is for occasional physical activity, and two (2) is for regular physical activity.

In a situation where the values of dependent variables are expressed in several categories like this, and if there is a natural order in the classification, usually two estimation approaches are used: ordered probit or ordered logit. The difference between the two approaches stems from the underlying assumption of the distribution of the error terms. In an ordered probit model the error term is assumed to be distributed normally while an ordered logit model assumes the cumulative density of the error term is expressed in the form of a logistic function. The basic

ordered choice model is based on the following specification where there is a latent regression expressed as:

$$(4.15) \quad Y_i^* = \beta' X_i + \varepsilon_i, \varepsilon_i \sim F(\varepsilon_i | \theta), E[\varepsilon_i | X_i] = 0 \quad Var[\varepsilon_i | X_i] = 1$$

The latent “preference” variable Y^* is not observed. The observed counterpart of Y^* is Y_i , where it is measured by a set of exogenous variables (X_i) and disturbance term ε_i .

The ordered probit model for physical activity (PHYA) is specified as:

$$(4.16) \quad PHYA_i^* = \gamma_1 BMI_{it} + \gamma_2 WGS L_{it} + \gamma_3 X_{it} + \gamma_4 A_{it} + \gamma_5 S_{it} + \varepsilon_{it}$$

where $PHYA_i^*$ is the unobserved frequency (latent variable) of physical activity for individual i , BMI_{it} is the body mass index for individual i at time t , $WGS L_{it}$ is total income from wages and salaries of individual i at time t , X_{it} is a vector of individual i characteristics at time t , A_{it} is a vector of natural amenities and outdoor recreational opportunities at time t , S_{it} is a vector of built environment measures of the residential location of individual i at time t .

The latent variables need to be linked to their observed discrete counterparts. For physical activity, $PHYA^*$ is linked to three levels of self-reported physical activity.

$$(4.17) \quad PHYA = \begin{cases} PHYA = 0 & \text{if } PHYA^* \leq \mu_1 \\ PHYA = 1 & \text{if } \mu_1 \leq PHYA^* \leq \mu_2 \\ PHYA = 2 & \text{if } \mu_2 \leq PHYA^* \end{cases}$$

where $PHYA=0$ is never engaged in physical activity, $PHYA=1$ is engaged in occasional physical activity, and $PHYA=2$ is engaged in regular physical activity. μ_j are threshold parameters that are estimated along with γ_{it} and ε_{it} in equation 4.16.

A likelihood function can be formed, and once the density function is specified, the model can be estimated. Even though the ordered probit model results from assuming ε is distributed normally, and the ordered logit model results from assuming the cumulative density of ε is the logistic function, in practice the two formulations yield the same results (Kennedy, 2003). Although the threshold parameters, μ_j , can be estimated as either fixed effect or random effect, in a general panel model framework, the latent nature of the dependent variable in the data and simultaneous equations necessitates the assumption of a random effect for the time-invariant components (Hsiao, 2003). The violation of this assumption would lead to biased estimates (Cai, 2010). Therefore, in this study only a random effect probit model is used. The general expression of a random effect ordered probit model is presented as follows:

$$(4.18) \quad \text{Pr } ob(Y_{it} = j) = F[\mu_j - (\beta' X_{it} + \nu_i)] - F[\mu_{j-1} - (\beta' X_{it} + \nu_i)]$$

where ν_i is uncorrelated with X_{it} and the error term ν_{ij} is assumed to be normally distributed.

4.6. Data Sources and Variables Description

To estimate the empirical model discussed in sections 4.3 and 4.4, data is collected from different sources. The study focuses on 83 U.S. metropolitan areas. One of the main sources of data for this study is survey data from the U.S. Bureau of Labor Statistics, National Longitudinal Surveys of Youth (NLSY79). This data is a nationally representative sample of individuals who were born between the years of 1957 to 1964. The survey was conducted annually from 1979 to 1994 and bi-annually after 1994.

For this study, data from 1998-2008 and a total of 1768 individual samples are used. Out of the 1768 individuals 877 are male and 891 are female. These individuals live in 405 counties of the 83 metropolitan areas under consideration. The NLSY79 data is geocoded data where the

county of the respondent is identified. County level data on the natural environment is available from the National Outdoor Recreation Supply Information System (NORSIS) (USDA Forest Service 1997). Using the geocoded data, respondents are directly linked to their county of residence, their county’s amenities, outdoor recreational facilities, and built environment by their Federal Information Processing Standard (FIPS). Table 4.1 presents the variable symbols, descriptions and sources of data.

Table 4.1. Socio-economic & Demographic Variable Descriptions and Data Sources

Variables	Definition	Source of Data
PHYA	Frequency of moderate and Vigorous exercise for > 10 minutes (Computed)	U.S Bureau of Labor Statistics (NLSY79)
BMI	Obesity and overweight measure (computed)	NLSY79
EDUC_1	High school degree or less	NLSY79
EDUC_2	Less than four years of college education	NLSY79
EDUC_3	4 years of college and above	NLSY79
GNDR_F	Female	NLSY79
RACE_B	Black	NLSY79
RACE_H	Hispanic	NLSY79
RACE_O	Non-Black and non-Hispanic	NLSY79
MARRD_N	Not married	NLSY79
DPNDT	Number of dependents in the household	NLSY79
WGSL	Total annual income from wages and salaries (\$000)	NLSY79
UNEMP	Unemployment rate of the county	U.S Bureau of Labor Statistics
POPCG	Total population change for the county	Department of Commerce, Census Bureau

Individuals were asked separate questions of how frequently they engage in moderate and vigorous levels of physical activity for 10 minutes or more. The responses are categorized into five groups of: a) three times or more each week, b) once or twice a week, c) one to three times

a month, d) less than once a month, and e) never. The responses of how often the individuals participate in moderate physical activity and vigorous physical activity are combined to create a physical activity (PHYA) variable. Two steps are taken in creating this physical activity (PHYA) variable (see table 4.2). The first step is assigning values of 0 or 1 for moderate (M) physical activity and vigorous (V) physical activity. A value of 0 (M=0) is assigned if the reply for moderate physical activity is one to three times a month, less than once a month or never participate in moderate physical activity. Likewise, 0 (V=0) is assigned if the reply for vigorous physical activity is one to three times a month, less than once a month, or never. M=1 is assigned if the answer for moderate physical activity is once or twice a week, and V=1 is given if the answer for vigorous activity is once or twice a week. M=2 and V=2 if the individual engaged in moderate or vigorous physical activity three times or more each week, respectively. The second step is combining the above created dummy variables for moderate and vigorous to create a aggregated physical activity (PHYA) variable. As indicated in table 4.2 PHYA= 0 if M=0 and V=0; if M=1 and V=0, and if M=0 and V=1. PHYA=1 if M=0 and V=1; if M=1 and V=1, and if M=2 and V=0. PHYA=2 if M=2 and V=2; if M=1 and V=2; if M=1 and V=1 and if M=0 and V=2. PHYA_0 is defined as never participating in physical activity, PHYA_1 is occasional physical activity and PHYA_2 is considered regular physical activity.

Table 4.2. Physical Activity Dummy Variables

Responses for Moderate Exercise	Moderate Dummy variables	Responses for Vigorous Exercise	Vigorous Dummies variables		PHYA Dummy Variables
A. Never	2	F. Never	2	If M=0 and V=0; If M=1 and V=0	0
B. Less than once a month	1	G. Less than once a month	1	If M=0 and V=1; If M=1 and V=1, and If M=2 and V=0	1

C. One to three times a month	0	H. One to three times a month	0	If M=2 and V=2; If M=1 and V=2; If M=0 and V=2	2
D. Once or twice a week	0	I. Once or twice a week	0		
E. Three times or more each week	0	J. Three times or more each week	0		

Respondents provided information about their height and weight from which their body mass index (BMI) was computed using the following formula: $BMI = \left(\frac{\text{weight in pounds} * 703}{\text{height in inches}^2} \right)$. Education is given in three categories, high school education or less (EDUC_1), less than 4 years of college education (EDUC_1), and college education of four years and beyond (EDUC_2). A dummy variable of 1 is assigned if gender is female and 0 if male (GNDR_F). To understand how obesity and physical activity vary across race and ethnicity, dummy variables are used for Hispanic (RACE_H), Black (RACE_B) and other (RACE_O).

For built environment data, the study relied on the Ewing, Pendall, and Chen (2002) Smart Growth America (SGA) data set. SGA developed four sprawl indices for 83 metropolitan areas to measure four aspects of the built environment. Over 22 variables were identified and a scalar number estimated. Four indices of the built environment are calculated and, based on these, one aggregate sprawl index is constructed. To establish the relationship between the built environment and BMI, studies such as Ewing et al. (2003), Plantinga and Bernell (2007) and Ewing, Brownson, and Berrigan (2006) used the SGA aggregate sprawl index. This study attempts to examine the implications for each of the four sprawl measures on BMI and physical activity. The four sprawl indices used here are: (1) residential density (DNST), (2) neighborhood mix (MXLU), (3) strength of activity centers and downtowns (CNTD), and (4) accessibility of

the street network (STCN). Table 4.3 provides the variable descriptions and data sources of the built environment measures.

Table 4.3. Built Environment Variables Description and Data Sources.

Variables	Definition	Source of Data
STCN	Street Connectivity	Smart Growth America (2002)
CNTD	Centeredness	Smart Growth America (2002)
MXLU	Mixed Land Use	Smart Growth America (2002)
DNST	Density	Smart Growth America (2002)
RSDNT	County of residence of the respondent	NLSY79

Street connectivity (STCN) measures whether the street network is dense or sparse, connected or disconnected, and also the distance, length and the size of the streets. According to SGA, metropolitan areas with the lowest street connectivity are Rochester, NY; Syracuse, NY; Atlanta, GA; Hartford, CT; and Greenville-Spartanburg, SC. On the other hand New York City, Jersey City, NJ; San Francisco, CA and New Orleans, LA scored the highest for street connectivity.

Centeredness (CNTD) refers to the strength of metropolitan centers. It captures how the downtown is active in helping businesses to expand, attract residential development and develop alternative means of transportation. SGA ranks metro areas such as Vallejo, CA; Riverside, CA; Oakland, CA and Gary, IN the lowest in the strength of their metropolitan centers. On the other hand metropolitan areas such as New York City, Honolulu, HI; Columbia, SC; Springfield, MA; and Providence, RI are ranked as the highest in centeredness.

Mixed land use (MXLU) measures the extent of the separation of residential households from other entities such as schools, offices, shopping centers, etc. Based on the SGA ranking, Raleigh, NC has the lowest mixed land use followed by Riverside, CA; Greensboro, NC; Greenville, SC; and West Palm Beach, FL. The top five places which scored highest on the

mixed land use scale are Jersey City, NJ; New Haven, CT; Providence, RI; Oxnard, CA and Bridgeport, CT.

Residential density (DNST) measures the extent to which households are separated from each other. A low density area is where houses are the most spread-out. According to SGA, the metropolitan area with the lowest housing density is Knoxville, TN followed by Greenville, SC; Greensboro, NC; Columbia, SC; Raleigh-Durham, NC; and Birmingham, AL. On the higher end of the residential density scale New York City ranked first followed by the Jersey City, San Francisco, Los Angeles, Chicago and Miami metro areas.

Variables that measure recreational opportunities and natural amenities are collected from NORIS, 1997. Around 26 variables that measure natural amenities and outdoor recreational opportunities are selected. In an attempt to reduce a wide group of recreational opportunities and natural amenity attributes into multiple but similar groups, principal component analysis (PCA) is used. PCA is a method used to condense variables that are related to a single measure (Deller et al., 2001). It reduces a multidimensional data set to lower dimensions in the form of a score or index that can be used in regression analysis (Kwang-koo, Marcouiller and Deller, 2005). PCA creates an index of linear combinations of the original variables where the linear weights are the eigenvectors of the correlation matrix between the set of factor variables. A number of studies relied on PCA to categorize a number of related factors into one measure. For example, Deller et al. (2001) studied the impact of amenities on economic growth by categorizing a large set of amenity variables into five amenity measures such as climate, urban facilities, land, water, and winter amenity attributes using PCA. Similarly, Marcouiller, Kim and Deller (2004) condensed 44 natural amenities into five amenity groups of river, lake, land, summer and winter amenity recreational attributes.

In this study natural amenities and outdoor recreational opportunities are condensed into four indices as shown in table 4.4. These indices include land-based, water-based, and winter-based recreational opportunities along with a climate-based index.

Table 4.4. Definition and Data Sources for Natural and Outdoor Recreational Amenities.

Variable	Definition	Data Source
CLMT	Climate-based amenity index	Computed from NORSIS 1997
LAND	Land-based outdoor recreational amenity index	Computed from NORSIS 1997
WATR	Water-based outdoor recreational amenity index	Computed from NORSIS 1997
WINTR	Winter-based outdoor amenity index	Computed from NORSIS 1997

The land-based outdoor recreational index is constructed from nine components as indicated in table 4.5. These components include parks and recreational departments, playgrounds and recreational centers, private and public swimming pools, organized camps, local, county or regional parks, public and private golf courses, hunting, fishing clubs and lodges and total rail-trail miles. The number of these facilities is based on the American Business Information (ABI) database as listed in NORSIS and the number of rail-trail miles is calculated by the Rails-to-Trails Conservancy (RTC). The retained first component analysis explains 49.85% of the variability in the data. Los Angeles County (CA), Cook County (IL), Maricopa County (AZ) and Harris County (TX) scored highest for the land based outdoor recreational index.

Table 4.5. Results of Principal Component Analysis for Land-based Outdoor Recreational Index

Variable Name	Linear weights (Eigenvector)
ABI number of parks and recreational departments	0.401
ABI number of playgrounds and recreational centers	0.392
ABI number of private and public swimming pools	0.261
ABI number of private and public tennis courts	0.416
ABI number of organized camps	0.364

RTC total rail-trail miles	0.147
ABI number of hunting, fishing preserves, clubs and lodges	0.109
ABI number of local, county or regional parks	0.352
ABI number of private and public golf courses	0.393
<i>Cumulative variance explained</i>	49.85%

The five attributes of water-based outdoor recreation are presented in table 4.6. These attributes are number of marinas, number of canoe outfitters, boat rental firms and raft trip firms, diving instruction and tours and snorkel outfitters, number of guide services and number of fishing camps, private and public fishing lakes, piers, and ponds. The first component analysis explains 28.43% of the data variability. Lake County (FL), Pinellas County (FL), King County (WA) and Duval County (FL) scored the highest in this index.

Table 4.6. Results of Principal Component Analysis for Water-based Outdoor Recreational Index

Variable Name	Linear weights (Eigenvector)
ABI number of marinas	0.517
ABI number of canoe outfitters, boat rental firms and raft trip firms	0.342
ABI number of diving instruction/tours and snorkel outfitters	0.295
ABI number of guide services	0.515
ABI number of fishing camps, private and public fishing lakes, piers, and ponds	0.513
<i>Cumulative variance explained</i>	28.43%

The winter-based outdoor recreational index is created by considering six components as shown in table 4.7. These components are cross country skiing firms, public cross country ski centers, ski firms, ski area trail miles, privately owned ski resorts and publicly owned ski resorts as counted by Inter-Ski Services (ISS) and reported in NORSIS (1997). San Bernardino County (CA), King County (WA), Placer County (CA) and Pima County (AZ) scored the highest for the winter-based outdoor recreational index. The principal component analysis explains 38.39% of the data variation.

Table 4.7. Results of Principal Component Analysis for Winter-based Outdoor Recreational Index

Variable Name	Linear weights (Eigenvector)
Number of cross country skiing firms	0.425
Number of public cross country ski centers	0.444
Number of ski firms	0.456
Ski area groomed trail miles	0.440
ISS number of privately owned ski resorts	0.470
ISS number of publicly owned ski resorts	0.394
<i>Cumulative variance explained</i>	38.39%

The climate index is created using six climatic components of average temperature, average annual precipitation, average January temperature, average annual sunny days, average July temperature and average July humidity. Table 4.8 exhibits the attributes of the climate index. The first component analysis explains 47.41% of the variability of the data. Counties in Florida Metropolitan Statistical Areas (MSA) such as Broward, Palm Beach, Pinellas and Osceola scored the highest for the climate index. These counties also scored the highest average overall temperature, highest average January temperature and highest average January sunshine.

Table 4.8. Results of Principal Component Analysis for Climate Index

Variable Name	Linear weights (Eigenvector)
Average temperature	0.563
Average annual precipitation	0.209
Average January temperature	0.536
Average January sunny days	0.324
Average July temperature	0.485
Average July humidity	0.112
<i>Cumulative variance explained</i>	47.41%

The prevalence of obesity varies across states and regions. For example, according to the BRFSS report (CDC, 2012), in 2011 the prevalence of obesity ranged from 20.7% in Colorado to

34.9% in Mississippi. During this year no state had a prevalence of obesity less than 20%. Thirty-nine states had a prevalence of 25% or more; 12 of these states had a prevalence of 30% or more. Out of the twelve states with obesity prevalence of 30% or more, ten states (Alabama, Arkansas, Indiana, Kentucky, Louisiana, Mississippi, Oklahoma, South Carolina, Texas, and West Virginia) are located in the southern region and two states (Michigan and Missouri) are located in the Midwest. Historically, states that are located in the South have higher prevalence of obesity followed by Midwest states. Recently, the BRFSS 2011 report (CDC, 2012) indicates that states located in the South had the highest prevalence of obesity (29.5%), followed by states in the Midwest (29.0%), in the Northeast (25.3%) and in the West (24.3%).

To account for the regional variation of obesity prevalence, dummy variables are created to represent each of the four regions (South, Midwest, West and Northeast). The regional categories are created following the Bureau of Labor Statistics Consumer Price Index (CPI) categories. As presented in table 4.9, the Southern region is comprised of 12 states and Washington DC, the West has 9 states, the Midwest region also includes 9 states, and the Northeast consists of 6 states.

Table 4.9. Regional Categorization of States

Regions	States
South	Alabama, Arkansas, Florida, Georgia, Louisiana, Maryland, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia
West	Arizona, California, Colorado, Hawaii, Nevada, New Mexico, Oregon, Utah and Washington
Midwest	Illinois, Indiana, Kansas, Michigan, Minnesota, Missouri, Nebraska, Ohio and Wisconsin
Northeast	Connecticut, Massachusetts, New Jersey, New York, Pennsylvania and Rhode Island

Descriptive statistics of the variables used for estimations are reported in table 4.10. For estimating the BMI equation, three dummy variables for physical activity were used, PHYA_0,

PHYA_1 and PHYA_2, as defined above. For estimation of the factors associated with physical activity (as the dependent variable), an ordered physical activity variable (PHYA) was used with values equal to zero, one and two, following the steps provided in table 4.2. The minimum BMI in the sample, 13.99, suggests that this individual is very severely underweight (less than 15), and the maximum, 68.65, represents a person that morbidly obese (greater than 40). No errors were found for these individuals so, even though they appear to be outliers they were included in the sample. The average BMI of 28.56 falls into what is generally considered the overweight category that ranges from 25 to 30.

Average income in the sample is \$41,333 which is similar to the national average annual wage in 2008 (\$41,335) according to the U.S. Social Security Administration. The U.S. Census Bureau (2010) reported that adults who are married declined in the last decade from 57.3 percent in 2000 to 54.1 percent in 2010. Married individuals in this study are thus over-represented since 60 percent of the sample is married (40 percent not married shown in table 4.10). The sample used for this study is 6 percent Hispanic, 32 percent Black, and 62 percent other than Black or Hispanic. This can be compared to the U.S. racial and ethnic distribution in 1980 since this study uses data from individuals who were interviewed through time beginning in 1979. In the U.S. in 1980, Hispanics were 6 percent of the population, and Blacks, and non-Blacks and non-Hispanics were 12 and 80 percent of the population, respectively. Thus, the percentage of Hispanics in this study is in line with the 1980 U.S. Census information, however, the percentage of Blacks is higher and the percentage of non-Blacks and non-Hispanics in the sample is lower. This could bias this study somewhat since Blacks tend to have relatively high rates of obesity (Paeratakul et al., 2002).

Table 4.10. Summary of Descriptive Statistics

Variables	Mean	Std. Dev.	Minimum	Maximum
BMI	28.56	6.10	13.99	68.65
PHYA_0	0.24	0.43	0	1
PHYA_1	0.38	0.49	0	1
PHYA_2	0.38	0.48	0	1
PHYA	0.86	0.40	0	2
WGSL (dollars)	41,333	42,986	0.00	307,823
MARRD_N	0.40	0.49	0	1
RACE_O	0.62	0.49	0	1
RACE_B	0.32	0.47	0	1
RACE_H	0.06	0.24	0	1
GNDR_F	0.50	0.50	0	1
EDUC_1	0.48	0.50	0	1
EDUC_2	0.27	0.45	0	1
EDUC_3	0.25	0.43	0	1
DPNDT	1.59	1.34	0	8
CLMT	0.22	1.79	-12.34	9.25
LAND	3.08	5.40	-27.51	22.02
WATR	0.79	1.74	-3.30	15.68
WINTR	0.70	2.47	-0.50	13.50
STCN	102.15	26.03	37.20	166.80
CNTD	102.20	26.03	40.90	167.30

MXLU	99.12	24.12	39.50	172.90
DNST	107.88	35.09	71.20	242.50
UNEMP	4.87	3.67	1.10	14.30
POPCG	13,673	27,073	-246,640	128,987
RGN_W	0.25	0.43	0	1
RGN_M	0.25	0.43	0	1
RGN_N	0.18	0.38	0	1
RGN_S	0.32	0.47	0	1

The descriptive statistics indicate that this sample is representative of the U.S. educational attainment distribution. In the sample, 48 percent of the individuals have a high school diploma or less education, 27 percent have some years of college education and 28 percent have four years of college education or beyond. Comparing this to national metro area educational attainment, according to the U.S. Department of Commerce report of 1998-2008, in 2006, 45 percent of individuals had a high school diploma or less, 27.7 percent had some years of college education, and 27.3 percent had four years of college education and beyond. In the sample, on average, the metropolitan areas gained 13,673 individuals. This coincides with growth of 10.8 percent in metro areas from 2000 to 2010 in the U.S. (U.S. Census, 2011).

Looking at the regional distribution of the sample data, one fourth of the sample lives in the West and one fourth in the Midwest, which is comparable to 2010 population distribution that was 22 percent in the Midwest and 23 percent in the West. In the sample, one third lives in the South compared to 37 percent in the South in 2010, and 18 percent lives in the Northeast, the same percentage as in 2010 (U.S. Census, 2011).

Chapter 5

Results and Analysis

5.1. Empirical Estimation

In this chapter, models presented in chapter 4 are empirically estimated and results are presented. The model examines the relationship among endogenous variables BMI, physical activity, and income, and the exogenous variables of socio-demographic characteristics, built environment measures, natural amenities and outdoor recreational opportunities. The model has three equations: BMI, physical activity (PHYA) and income (WGSL).

5.2. Findings and Analysis

In a system of simultaneous equations, the endogenous variables or dependent variables are determined by simultaneous interaction of several relationships. In this system of equations all of the endogenous variables are random variables; therefore, any change in the error term or in the disturbance will affect all of the endogenous variables.

One of the limitations of a system of simultaneous equations is that it violates the basic ordinary least squares assumption. This assumption states that in repeated samples, exogenous variables can be considered as fixed. Simultaneous equation models (SEMs) violate this assumption because, in SEMs, at least one endogenous variable is included as an exogenous variable. As a result, the endogenous variable cannot be considered fixed in a repeated sample (Kennedy, 2003). Secondly, in SEMs the endogenous variables used as explanatory variables can be contemporaneously correlated with the error term. The endogenous variables used as explanatory variables might not be distributed independently, indicating the change in the disturbance or error term will directly affect the endogenous variable it determines. These in turn

affect the other endogenous variables since they are determined simultaneously (Kennedy, 2003). Under such circumstances, ordinary least squares (OLS) is a biased estimator. Among other techniques, two-stage least squares (2SLS) is the most popular for estimating a system of simultaneous equations. Two-stage least squares estimation is generally consistent, unbiased and robust, i.e., the results are not sensitive to other estimation problems such as specification error and multicollinearity (Kennedy, 2003). Thus, this study uses a 2SLS estimator to deal with the issue of consistency and biased estimates from using OLS. The panel nature of the data necessitates using Fixed Effect and Random Effect estimation methods. However, one of the limitations of using a Fixed Effect method is estimation of time-invariant variables will not be possible (Kennedy, 2003). The time-invariant variables included in this study are gender, race, built environment measures, and amenities and natural recreational opportunities. Even though, in reality, built environment, natural amenities and recreational opportunities change over time, due to lack of data, only one time period is used. Therefore, the Random Effect estimation method is employed for the analysis. In addition, because of the categorical nature of the physical activity equation, an Ordered Probit model is used for estimation.

5.2.1 BMI Equation: Findings and Analysis

The BMI equation is estimated as a function of socioeconomic, physical activity, built environment and natural amenities and recreational opportunity measures as presented in table 5.1. The overall fit of the estimation is measured by an R^2 of 0.0134. This R^2 value is very low, but it is consistent with similar studies (for example, Eid, et al., 2008, Chou, Grossman, and Saffer, 2004 and Schmeiser, 2009) that used panel data for their estimations.

Physical Activity

The results indicate that there is a positive and significant association between BMI and physical activity (PHYA_1 and PHYA_2). Individuals who participate in occasional physical activity (PHYA_1) have higher BMI than those who do not participate (PHYA_0) by 0.35 BMI units. Similarly, individuals who regularly participate in physical activity (PHYA_2) have higher BMI of 0.21 units compared to those who do not participate in physical activities (PHYA_0). This positive association suggests that even as physical activity increases, it might not be sufficient to actually decrease weight. Further, if an increase in physical activity is not accompanied by calorie intake control, exercise alone may not be able to have a significant impact on weight. If increasing exercise levels leads to increased calorie intake, this could more than offset the calories burned through exercise and weight gain could occur. Another possible explanation for the positive relationship is that an increase in physical activity may result in an increase of body muscle and an increase in muscle may translate into higher BMI. Wen and Jones (2012) related higher BMI with work-related activities indicating higher BMI may suggest healthy weight due to participating in physically labor-intensive occupations. Similarly, Eid et al. (2008) and Lakdawalla and Philipson (2002) reported a positive link between work related exercise and BMI, indicating that individuals who participate in jobs that require more strength tend to have higher BMI¹. One of the limitations of using the BMI index is that it does not distinguish between body muscle and body fat (Mei et al., 2002). According to BMI calculations, fit individuals with high muscle mass may score a higher BMI and be considered obese or overweight. Therefore, the positive implication of PHYA_1 and PHYA_2 on BMI could be the result of greater muscle mass.

¹ The physical activity data does not differentiate between leisure time physical activity and job related physical activity. Hence it is possible that individuals who work in a more labor intensive industry report higher physical activity participation.

Table 5.1. BMI Equation Results

Variable	Estimate	P Value
PHYA_1	0.35*** (5.51)	<0.0001
PHYA_2	0.21*** (3.27)	0.0011
WGSL	0.0001*** (3.96)	<0.0001
EDUC_2	-1.12*** (-2.63)	0.0085
EDUC_3	-4.75*** (-4.66)	<0.0001
RACE_B	3.25*** (7.66)	<0.0001
RACE_H	1.35*** (2.26)	0.0236
GNDR_F	1.99*** (2.65)	0.0081
DPNDT	-0.14 (-1.29)	0.1957
MARRD_N	-0.24** (-2.49)	0.0126
CLMT	0.10** (2.04)	0.0415
LAND	-0.02 (-0.90)	0.3676
WATR	-0.04 (-0.86)	0.3913
WINTR	0.06** (2.13)	0.0335
DNST	-0.002** (-1.98)	0.0439
CNTD	-0.0005 (-0.13)	0.8956
MXLU	0.02*** (3.50)	0.0005
STCN	-0.01 (-1.16)	0.2441
RGN_W	-0.09 (-0.86)	0.3904
RGN_M	-0.30*** (-2.71)	0.0067
RGN_N	-0.16 (-1.19)	0.2351
R Squared	0.0134	

Note: t-statistics given in parenthesis

*significant at 0.1 level, **significant at 0.05 level and ***significant at 0.01 level.

Socioeconomic factors

The results show a positive and significant association between income (WGSL) and BMI, although the impact is very small. An increase in annual income of \$10,000 increases BMI by 0.0001 point. One possible explanation could be that an increase in income (measured by wages and salaries) may lead to working longer hours (Finkelstein et al. 2009), which could, in turn, lead to a higher BMI due to limited time available to prepare food at home, resulting in increased consumption in restaurants or of convenience foods or prepackaged foods that are more calorie-dense (Chou et al., 2004).

The result indicates that there is a negative association between less than four years of college education (EDUC_2) and of 4 years of college and beyond (EDUC_3) and BMI. Individuals with some college (EDUC_2) have lower BMI than individuals with only a high school diploma by 1.12 points. Likewise, individuals with 4 years of college or beyond (EDUC_3) have lower BMI than individuals with a high school education or less by 4.75 points. With more education, awareness and understanding of overall health and the negative consequences of weight on health may also increase. More educated people may be better able to utilize health-related information and maintain normal weight. This finding is consistent with those of Michimi and Wimberly (2012) who found a negative relationship between BMI and four years of college education and beyond, and Eid et al. (2008) who found a negative relationship between more years of schooling and BMI.

The results show a positive association between BMI and Blacks (RACE_B); compared to non-Blacks and non-Hispanics, the BMI of Blacks is higher by 3.25 points. Similarly, Hispanics have higher BMI than non-Blacks and non-Hispanics; holding other factors constant, the BMI of Hispanics is higher by 1.35 points. Females have a positive significant association

with BMI such that their BMI is 1.99 points higher than that of males. There is no statistically significant association between BMI and number of dependents (DPNDT) in the household.

Compared to married individuals, unmarried individuals have lower BMI by 0.24 points.

Climate and Outdoor Recreational Opportunities

A positive and significant association between BMI and climate (CLMT) is found. This result is unexpected. One possible explanation may be related to the variation of climate across regions.

In this study, the highest climate index values are observed in counties that are located in the southern region of the U.S. (Florida, Texas, Louisiana, Georgia, and South Carolina). Thus the observed positive association may be explained by the fact that individuals with high BMI are located in those southern counties that have a higher prevalence of obesity.

The results indicate that there is no significant association between BMI and land-based recreational opportunities (LAND) and water-based recreational opportunities (WATR).

Although this was not expected, it may be that the measures of water-based recreation used here, for example fishing, sailing and cruising in a power boat, may not require intense physical activity that burns enough calories to actually have a negative impact on weight. The study found a positive association between WINTR and BMI, suggesting that individuals who live in counties with plenty of winter-based recreational opportunities have 0.06 higher BMI. Winter-based recreational opportunities are generally located in areas where there is plenty of snow and a long winter. The observed higher level of BMI in counties with a high level of winter-based recreational opportunities, therefore, may be the result of lack of or inadequate physical activity due to the fact that wintertime activities, such as snow skiing, are not convenient and accessible for many people (Merrill et al., 2005). As a result, physical inactivity or insufficient physical activity may lead to a sedentary lifestyle and subsequent weight gain.

Built Environment

A negative and significant association between density (DNST) and BMI is observed. The result suggests that individuals who live in counties with high residential density have 0.002 lower BMI. This result is consistent with the findings of Wen and Jones (2012). No significant relationship between BMI and centeredness (CNTD) is found. One possible explanation for this lack of association could be the broad definition of centeredness at the county level in this study, which may not be a good proxy for a neighborhood-level measure. Also, no significant relationship is found between street connectivity (STCN) and BMI. This result is consistent with the findings of Ball et al. (2012), Burgoine, Alvanides, and Lake (2011) and McDonald, Oakes, and Forsyth (2012). The general belief that well-connected and well-integrated streets encourage physical activity and may lead to lower BMI is not supported by the evidence.

A positive and significant relationship is observed between mixed land use (MXLU) and BMI, such that a one point increase in the mixed land use index increases BMI by 0.02 point. Although mixed land use is expected to encourage walking or biking, and thus lead to lower BMIs, higher weight could result from individuals having easier access to calories if the mixed use includes restaurants or food stores. Or, mixed land use alone may not be sufficient to lower BMI. Another possible explanation could be location preference or neighborhood selection of individuals. The location of individuals in a specific neighborhood or area is not a random occurrence; rather socioeconomic factors drive particular individuals to locate in a specific location (McDonald, Oakes, and Forsyth, 2012). For example, African-Americans and Hispanics are disproportionately located in downtown areas that are characterized as low residential sprawl and high mixed land use (Eid et al. 2008). The availability and accessibility of public transportation and affordable housing in cities may be a major factor for economically

disadvantaged groups to locate in such areas. Hence, the positive association between mixed-land use and BMI could be explained by the fact that a disproportionate section of the population (low-income groups, African-Americans and Hispanics) that are more likely to be obese and overweight may have located in neighborhoods with high mixed land use. The literature shows mixed results. For example, Eid et al. (2008) found no significant relationship between mixed use and BMI, whereas Ewing et al. (2003) found a negative relationship.

Regional Variation

Regional dummies were introduced to account for regional variation in the prevalence of obesity. The result shows a negative and significant association between BMI and living in the Midwest (RGN_M) suggesting individuals who live in the Midwest have lower BMI compared to individuals who live in the South by 0.3 point. Results for the Northeast (RGN_N) and West (RGN_W) indicate no significant variation in BMI compared to the South (RGN_S).

5.2.2. Physical Activity Equation: Findings and Analysis

The Ordered Probit estimation method is used to estimate the physical activity (PHYA) equation. The findings are presented in table 5.2.. The marginal effects of not participating in physical activity (PHYA_0), occasional participation in physical activity (PHYA_1) and regular participation in physical activity (PHYA_2) are reported in columns 2, 3 and 4, respectively.

BMI

There are no significant associations between BMI and not-participating in physical activity (PHYA_0) or occasionally (PHYA_1) or regularly participating (PHYA_2) in physical activity. Dunbar-Jacob et al. (1998) argue that knowledge of important risk factors and health-related behaviors do not necessarily motivate changes in practices. Motivational factors, for example, perceived vulnerability to sickness, perceived consequences of the sickness and belief that

recommended action is appropriate or effective to reduce risk may be viewed as important but not sufficient enough to motivate change (Merrill, Friedrichs, and Larsen, 2002).

Socioeconomic factors

The results indicate that there is a positive and significant association between income (WGSL) and participating in regular physical activity (PHYA_2) such that an increase in income of \$1,000 is associated with a 13 percent increase in the likelihood of participating in regular physical activity. This suggests that as income increases individuals may better afford goods and services that promote physical activity such as a health club membership, joining a gym, paying a fee to visit a community pool or recreation center and buying the necessary gear and home exercise equipment. In addition, with an increase in income individuals are able to afford to buy services so that they will have free time to allocate to physical activity. The result may also suggest that working hours increase as wages increase, and as a result physical activity increases if the job is labor intensive. The negative and significant relationship between income (WGSL) and not-participating in physical activity (PHYA_0) shows that an increase in income of \$1,000 is associated with an 11 percent decrease in the likelihood of not-participating in physical activity. In contrast to the above results, an increase in income of \$1,000 is linked with a 20 percent decrease in the probability of engaging in only occasional physical activity (PHYA_1).

Table 5.2. Physical Activity Equation Results

Variables	PHYA_0	PHYA_1	PHYA_2
BMI	0.895 (1.588)	0.162 (1.586)	-0.106 (-1.562)
WGSL	-0.112*** (-3.061)	-0.203*** (3.040)	0.132*** (3.245)
EDUC_2	0.156*** (44.726)	0.220*** (13.098)	-0.177*** (-9.992)
EDUC_3	0.470*** (35.876)	0.271*** (35.437)	-0.497*** (-31.892)
RACE_B	-0.119*** (-15.122)	-0.242*** (-11.649)	0.143*** (7.247)

RACE_H	-0.524*** (-8.077)	-0.109*** (-6.926)	0.633*** (3.339)
GNDR_F	-0.229*** (-23.550)	-0.391*** (-17.246)	0.268*** (13.079)
MARRD_N	-0.253*** (-4.105)	-0.465*** (-3.824)	0.299 (1.564)
DPNDT	0.229** (2.481)	0.415** (2.469)	-0.271*** (-47.831)
CLMT	-0.419 (-1.542)	-0.759 (-1.541)	0.495 (1.568)
WATR	0.519 (0.211)	0.940 (.211)	-0.613 (-0.212)
LAND	-0.124* (-1.664)	-0.224* (-1.662)	0.146* (1.676)
WINTR	-0.590*** (-3.347)	-0.107*** (-3.327)	0.697*** (3.487)
STCN	0.968*** (3.501)	0.175*** (3.472)	-0.114*** (-3.698)
DNST	-0.0153 (-1.067)	-0.277 (-1.067)	0.181 (1.059)
CNTD	-0.161 (-1.204)	-0.292 (-1.202)	0.191 (.2275)
MXLU	-0.895** (-2.437)	-0.162** (-2.428)	0.106*** (2.541)
RGN_M	0.131** (2.380)	0.232** (2.482)	-0.154 (-0.817)
RGN_N	0.220*** (4.103)	0.381*** (4.407)	-0.258 (-1.375)
RGN_W	0.562 (1.003)	0.101 (1.021)	-0.153 (-0.817)

Note: t-statistics given in parenthesis

*significant at 0.1 level, **significant at 0.05 level and ***significant at 0.01 level.

A negative association between some college education and the likelihood of participating in regular physical activity (PHYA_2) is observed. The probability of participating in regular physical activity decreases by 18 percent for individuals with some college education (EDUC_2) compared to individuals with a high school education or less. Similarly, for individuals with an undergraduate degree and beyond (EDUC_3), the likelihood of participating in regular physical activity (PHYA_2) is 50 percent lower as compared to individuals with a high school diploma or less. There are several possible explanations for this result. First, work activities for people with higher education could be in an office and involve long hours of sitting

with little movement. Moreover, the nature of the job may require them to spend more time at the workplace with little time left over to spend on other activities, such as exercising. The sedentary work environment coupled with lack of time may contribute to declining participation in regular physical activity. Second, although it was expected that educated people would engage in physical activity regularly because they have knowledge of the health benefits of participating in physical activity, knowledge of the importance of physical activity may not necessarily translate to being physically active (Merrill, Friedrichs, and Larsen, 2002). Similarly, the likelihood of not-participating in any physical activity (PHYA_0) is positively associated with education for individuals with less than 4 years of college education (EDUC_2) and for those with a college degree or more (EDUC_3). The result indicates that, compared to individuals with a high school diploma or less, individuals with some college (EDUC_2) and individuals with a college degree or beyond (EDUC_3) are 16 percent and 47 percent more likely to not-participate in physical activity (PHYA_0), respectively. However, results for engaging in occasional physical activity (PHYA_1) show a positive association with education, where, compared to individuals with a high school diploma or less, individuals with some college (EDUC_2) and individuals with a college degree or beyond (EDUC_3) are 22 percent and 27 percent more likely to participate in occasional physical activity, respectively.

The likelihood of participating in physical activity varies across racial groups. The likelihood of engaging in regular physical activity (PHYA_2) is higher by 14 percent for Blacks compared to non-Blacks and non-Hispanics, and the likelihood of not-participating (PHYA_0) is 12 percent lower for Blacks compared to non-Blacks and non-Hispanics. The probability of Blacks only occasionally participating in physical activity (PHYA_1) is 24 percent lower than for non-Blacks and non-Hispanics. For Hispanics, the probability of engaging in regular physical

activity (PHYA_2) is higher than for non-Blacks and non-Hispanics by 63 percent. Hispanics are less likely not to participate in physical activity (PHYA=0) by 52 percent and less likely to participate in occasional physical (PHYA_1) activity by 11 percent as compared to non-Blacks and non-Hispanics. As compared to males, females are 27 percent more likely to engage in regular physical activity (PHYA_2) and 23 percent less likely to engage in no physical activity (PHYA_0). Whereas, females are 39 percent less likely than males to engage in only occasional physical activity (PHYA_1).

There is no significant association between marital status and the likelihood of participating in regular physical activity (PHYA_2). Individuals who are not married (MARRD_N) are 25 percent less likely to never participate in physical activity (PHYA_0) than those who are married, but unmarried individuals also are 47 percent less likely to occasionally participate in physical activity (PHYA_1). The likelihood of engaging in regular physical activity (PHYA_2) decreases as the number of dependents (DPNDT) increases, and the probability of not participating in physical activity increases as the number of dependents increases, by 27 percent and 23 percent, respectively. As more time is allocated for childrearing and other household activities, little or no time is available for other activities such as exercising. However, the likelihood of only occasional participation in physical (PHYA_1) increases by 42 percent as the household gains one more dependent.

Climate and Outdoor Recreational Opportunities

There is no significant association between climate (CLMT) and the probability of participating in regular, occasional or no physical activity. In addition, no significant association is found between water-based recreational opportunities (WATR) and the likelihood of engaging in physical activity or not. One possible explanation could be related to the construction of the

WATR index, which is computed using number of marinas, number of boat rental firms, number of guide services, number of fish camps and private and public fishing lakes, as indicated in section 4.6. These fishing- and boating-related activities may require little to no physical activity. Hence, WATR may not be a good proxy active water-based recreational opportunities.

A positive association between land-based recreational opportunities (LAND) and the probability of regular physical activity (PHYA_2) is found, such that a one unit increase in this index is associated with a 15 percent increase in the probability of engaging in regular physical activity. In addition, those living in counties with a higher land-based recreation index are 12 percent less likely to never engage in any physical activity (PHYA_0). These findings are consistent with prior expectations. The results also indicate a negative and significant association between land-based recreational opportunities and only occasional participation in physical activity (PHYA_1); a one unit increase in the index is associated with a 22 percent decrease in the likelihood of occasional participation in physical activity.

The results reveal a positive and significant association between winter-based recreational opportunities (WINTR) and the probability of engaging in regular physical activity (PHYA_2). A one unit increase in the index is associated with a 70 percent increase in the probability of regular physical activity participation. A one unit increase in this index also is associated with a 59 percent decrease in the probability of never engaging in physical activity. However, a one unit increase in the WINTR index is associated with an 11 percent decrease in the probability of occasional participation in physical activity (PHYA_1).

Built Environment Measures

Living in counties with higher street connectivity is, surprisingly, associated with an 11 percent lower probability of participating in regular physical activity (PHYA_2) and a 97 percent higher probability of engaging in no physical activity (PHYA_0). Living in counties with higher street connectivity, however, is associated with an 18 percent increase in the likelihood of occasionally participating in physical activity (PHYA_1). No significant association is found between density (DNST) or centeredness (CNTD) and any of the physical activity variables.

The association between the likelihood of engaging in regular physical activity (PHYA_2) and the built environment measure of mixed land use (MXLU) is positive and significant. A one unit increase in the mixed land use index increases the probability of engaging in regular physical activity by 11 percent. Mixed land use (MXLU) is negatively associated with the likelihood of not participating in physical activity (PHYA_0) and only occasional physical activity (PHYA_1). A one unit increase in the mixed land use index decreases the likelihood of never engaging in physical (PHYA_0) and only occasional physical activity (PHYA_1) by 90 percent, and 16 percent, respectively.

Regional Variation

The results examining regional variation indicate no significant association with the probability of engaging in regular physical activity (PHYA_2) and living in the Midwest (RGN_M), Northeast (RGN_N) or West (RGN_W) as compared to the South (RGN_S). A positive and significant association is found with the Midwest and Northeast and not engaging in any physical activity (PHYA_0), indicating that the Midwest and Northeast are more likely to never exercise, by 13 percent and 22 percent, respectively, when compared to the South. However, a positive and significant association between these regions and engaging in only occasional physical activity (PHYA_1) indicates that the Midwest and Northeast are 23 percent and 38 percent more

likely to occasionally participate in physical activity when compared to the South. The West is not significantly associated with the likelihood of not participating and occasional participation in physical activity.

5.2.3 Income Equation: Findings and Analysis

The income (WGSL) equation is estimated using a random effect estimation method. The results are presented in table 5.3. The overall fit of the estimation as measured by the R^2 is 0.0718.

Similar to the BMI equation, the value of the R^2 is very low. However, the low R^2 observed here is consistent with similar studies that used panel data (for example, Eid, et al., 2008, Chou, Grossman, and Saffer, 2004 and Schmeiser, 2009).

The findings indicate a positive and significant association between BMI and income (WGSL) suggesting individuals with higher BMI tend to earn more, such that a one point increase in BMI is associated with an increase in annual income of \$4,917. Literature indicates that the relationship between income and BMI is complex, mixed and varies across gender, age, race and employment sector (private vs. public). For example, Cawley (2004) pointed out that White females, Black females, Hispanic females, and Hispanic males with higher BMI tend to earn less, whereas Black males with higher BMI tend to earn more. In an examination of the relationship between BMI and wages in the private and public sectors, Greve (2008) found a positive association between BMI and wages for public sector employees between ages 18-30. Because the current study did not make comparisons between sectors, and did not account for gender-race interactions or the type of employment (labor intensive or not) this finding of a positive relationship should be interpreted cautiously.

The results indicate that occasional participation in physical activity (PHYA_1) is associated with income (WGSL) that is \$5,108 higher than when individuals engage in no

physical activity. Similarly, participation in regular physical activity (PHYA_2) is associated with \$2,031 more in annual income compared to no participation in physical activity. Even though physical activity in this case does not differentiate between work related physical activities, physical activity as a means of work transportation (from and to work) or leisure time physical activity, the association between some frequency of physical activity and higher income could be explained by participation in labor intensive jobs and other job-related exercise. Or it could mean that individuals who are more physically active are healthier, such that they can work more hours and earn greater income over the course of a year.

Table 5.3. Income Equation Results

Variable	Estimate	P Value
BMI	4,916.55*** (3.32)	0.0009
PHYA_1	5,108.35*** (7.90)	<.0001
PHYA_2	2,031.13*** (3.04)	0.0023
EDUC_2	11,112.16*** (6.48)	<.0001
EDUC_3	38,890.75*** (18.72)	<.0001
GNDR_F	-21,431.10*** (-11.60)	<.0001
RACE_B	-20,788.70*** (-6.28)	<.0001
RACE_H	-10,047.10*** (-2.88)	0.0040
POPCG	-0.01 (-0.88)	0.3784
UNEMP	86.53 (1.24)	0.2135
RGN_W	390.84 (0.39)	0.6997
RGN_M	1,868.91** (1.91)	0.0559
RGN_N	451.13 (0.37)	0.7131
R ²	0.0718	

Note: t-statistics given in parenthesis

*significant at 0.1 level, **significant at 0.05 level and ***significant at 0.01 level.

As expected, the results indicate a positive association between education and income. Individuals with less than 4 years of college education (EDUC_2) have a higher annual income, by \$11,112, compared to an individual with a high school diploma or less (EDUC_1). Similarly, having 4 years of college or more (EDUC_3) increases annual income by \$38,891 compared to that of an individual with a high school diploma or less. Not surprisingly, women (GNDR_F) earn \$21,431 less than men. Blacks earn \$20,789 less in annual income compared to non-Blacks and non-Hispanics. Hispanics earn \$10,047 less in annual income, again, as compared to non-Blacks and non-Hispanics. In this study county population change (POPCG) and county unemployment rate (UNEMP) are not significantly associated with individual level income.

Regional Variation

The results indicate that personal income is higher in the Midwest (RGN_M) compared to the South (RGN_S) by an average of \$1,869 per year. Income in the Northeast (RGN_N) and West (RGN_W) is found to not be statistically different from the South.

5.3. Revisiting the Hypotheses

Hypothesis No. 1

The availability of outdoor recreational opportunities and natural amenities are negatively related to obesity (BMI) and positively related to physical activity.

Land-based, water-based, winter-based recreational facilities and climate are introduced to understand the impact of outdoor recreational opportunities and natural amenities on obesity measured by BMI and on physical activity.

Implications for BMI

It is expected that abundant outdoor recreational opportunities and favorable climate promote physical activity, which in turn, decreases the BMI of the individual. Hence, individuals living in counties with plenty of outdoor recreational opportunities and favorable climate would have lower BMI. Contrary to this hypothesis, the climate index used here and the measure of winter-based recreational opportunities are positively related to BMI. Land-based and water-based recreational opportunities have no significant association with BMI. These results mean we cannot accept the first part of hypothesis number one.

Implications for Physical Activity

The availability of favorable climate and outdoor recreational opportunities were expected to increase the probability of participating in physical activity. The results find a positive and significant association between land-based recreational opportunities and the probability of engaging in frequent physical activity and a negative association with the probability of not being physical active and occasional physical activity. In addition, winter-based outdoor recreational opportunities are positively associated with the probability of participating in regular physical activity and negatively associated with the likelihood of no physical activity and occasional physical activity. Contrary to the stated hypothesis, climate and water-based outdoor recreational opportunities are found not to be significantly linked to the probability of not being physically active, and engaging in occasional and regular physical activity. Because all natural amenities and outdoor recreational opportunities are not positively associated with participating in physical activity, hypothesis one can only be partially accepted.

Hypothesis No. 2

Built environment measures are negatively related to obesity and positively related to physical activity.

Built environment measured by mixed land use, density, street connectivity, and centeredness were introduced to examine their influence on BMI and the likelihood of participating in occasional and frequent physical activity. Thus, counties that have high built environment measures would have less obesity and more individuals engaging in physical activities.

Implications for BMI

Examining the influence of built environment measures on BMI, a negative and significant association is found between density and BMI. Street connectivity and centeredness are not significantly related with BMI. Contrary to expectations, there is a positive association between mixed land use and BMI. Density is the only built environment measure consistent with hypothesis two; hence, the hypothesis of a negative association of BMI with built environment measures is not fully accepted.

Implications for Physical Activity

This hypothesis tests whether built environment measures of mixed land use, density, street connectivity, and centeredness promote physical activity. A negative and significant association between street connectivity and regular physical activity is found. Street connectivity, however, is positively associated with not engaging in any physical activity and participating in physical activity occasionally. Density and centeredness are not statistically significant. Consistent with expectations, the probability of engaging in regular physical activity is positively associated with mixed land use. Moreover, mixed land use is negatively linked with not engaging in any physical activity and participating in only occasional physical activity. Not all built environment measures are found to be as hypothesized. The results for mixed land use support the hypothesis.

The hypothesis is rejected in the case of street connectivity and not supported in the case of density and centeredness.

Hypothesis No. 3.

Income measured by wages and salaries is negatively related to obesity and positively related with physical activity.

Implications for BMI

Because a positive and significant association between income and BMI is found hypothesis number three is not supported by the evidence.

Implications for Physical Activity

It was hypothesized that an increase in income would increase the probability of engaging in physical activities by allowing individuals to afford goods and services necessary for facilitating exercise. The results show a positive association of income with regular physical activity and a negative association with no physical activity, both supporting hypothesis three. But a negative association of income with occasional physical activity means the hypothesis is not fully supported by the findings.

Hypothesis No. 4.

Physical activity negatively affects obesity and obesity decreases participation in physical activity.

Implications for BMI

Weight gain is the result of calorie imbalance where calorie intake is greater than calorie expenditure. Calorie expenditure can be accomplished by engaging in different kinds of physical activities. Regular and occasional participation in physical activity were introduced to capture calorie expenditures. The estimates indicate a positive association between occasional and

regular physical activity and BMI. This finding is contrary to the hypothesis of physical activity decreasing BMI. Therefore, the hypothesis is not accepted.

Implications for Physical Activity

Examining whether BMI is associated with the probability of occasional and/or regular physical activity participation, the results suggest that BMI is not significantly associated with physical inactivity, occasional, or regular physical activity. Therefore, the hypothesis that BMI is negatively associated with physical activity is not accepted.

Chapter 6

Summary and Conclusions

6.1. Summary and Conclusions

This study set out to determine associations between the physical environment, obesity, physical activity, and income. To address the potential endogeneity problem of weight and income, and weight and physical activity, a system of simultaneous equations is introduced. The system of equations is estimated using a two stage least squares with random effect approach. The ordinal nature of the physical activity variable necessitated the use of an ordered probit estimation method. The study focused on metropolitan areas of the United States. Panel data of six years, from 1998-2008 (collected every two years), for 1,768 individuals is used.

It is increasingly recognized that biology, socioeconomic, and physical environment factors contribute to rising obesity. One of the goals of this study was to explore the impact of the physical environment on weight. The physical environment in this case consists of the built environment, natural amenities and outdoor recreational opportunities. The built environment is broadly defined, however, in this study the built environment is measured by density, centeredness, mixed land use and street connectivity. Likewise, natural amenities and outdoor recreational opportunities are measured by climate, and land-based, water-based and winter-based recreational opportunities.

With regard to obesity, this study has a number of interesting outcomes. Surprisingly, participating in occasional and regular physical activity is positively associated with BMI. Compared to the BMI of individuals who do not engage in physical activity, the BMI of individuals who are occasionally and regularly active is higher. This may have something to do

with the way the physical activity data is measured, where the question asks if an individual was active for more than 10 minutes. An individual's activity could have lasted for 15 minutes or for more than an hour, however, no distinction can be made for this analysis. Moreover, whether the activity is job or leisure-related, which could affect its intensity and impact, is also not clearly specified. Despite measurement limitations, the result may signal that physical activity alone may not be sufficient to reduce weight, and there are other factors that will also have to be looked at simultaneously to determine overall effects on weight.

Another surprising result, though the literature is not conclusive, is that of the income-BMI relationship. As income increases, the results show that BMI is expected to increase as well. One argument to support this is that as income increases calorie intake increases and leads to weight gain. Similarly, this study finds that BMI is positively associated with income in the reverse relationship, such that as BMI increases income also increases. This shows the relationship is more complex than previously thought, where higher incomes are assumed to be associated with lower BMI and higher BMI is assumed to be associated with lower income. This could mean that in future studies the way income is included in the analysis may need to be refined. For example, a distinction may need to be made regarding whether income is increasing from a low level or increasing from a relatively high level. The fact that the average income in this study is \$41,333, which is similar to median US income, shows that as income increases from such levels, on average, calorie consumption may increase leading to weight gain.

There has been a growing understanding of the role environmental factors play in promoting physical activity and reducing obesity. The built environment can affect energy balance by presenting opportunities or barriers for physical activity. Implications regarding the built environment on BMI are inconclusive and mixed. The study does not find evidence that

street connectivity and centeredness are associated with BMI. Mixed land use is found to be positively linked with BMI. However, density is negatively associated with BMI. Until recently, there was a strong belief that areas with a favorable built environment promote physical activity such as walking, jogging, and biking, and thus reduce BMI, but this is now changing with new research that establishes no significant association between built environments and weight. The findings here will contribute to the growing body of literature that found an inconclusive association between the built environment and BMI, despite the partial conclusion drawn from density.

The implication of natural amenities and outdoor recreational opportunities on BMI is also found to be mixed and inconclusive. No compelling evidence is found that associates living in counties with abundant land-based and water-based recreational opportunities with lower BMI. Availability of these resources does not necessarily guarantee their usage. Climate and winter-based recreational opportunities, contrary to expectations, were found to be positively associated with BMI. The climate variable represents many southern metropolitan areas, where obesity rates are higher. Also, areas that are favorable to winter-based outdoor recreation may be associated with a more sedentary lifestyle, at least during a long winter, leading to higher BMI.

The fact that social factors, such as race, gender and marital status are also related with weight suggests that a more complex set of factors shapes obesity. This is particularly so for minorities, women and married individuals who have a higher tendency for higher weight. These results underlie the importance of taking into account social characteristics when addressing the broader obesity challenge. This result, combined with the finding that education is associated with obesity reduction, means that awareness and education in targeted social settings may have significant obesity control effects. Regional variations in obesity signal limited importance of

geography. Examining the distribution of obesity across regions, the evidence shows that individuals living in the Midwest have lower BMI than those in the South, but the evidence does not support any difference between the West or Northeast and the South.

Physical activity, as one among many feasible solutions for obesity and overweight, is often viewed as an important indicator of weight outcomes. An assessment in this study of the determinants of physical activity itself reveals that income plays an important role, as it is positively associated with regular physical activity, as an increase in income increases the affordability of goods and services that promote physical activity. Likewise, an increase in income decreases the probability of not being physically active. Generally, natural amenities and recreational opportunities are hypothesized to play a positive role in promoting physical activity. Living in counties with more land-based and winter-based recreational opportunities is associated with increasing the likelihood of engaging in regular physical activity, and decreasing the likelihood of physical inactivity, and only occasional physical activity. Climate and water-based recreational opportunities, however, do not influence physical activity or inactivity. Hence, the implication of natural amenities and outdoor recreational opportunities for promoting physical activity is limited and mixed.

The influence of the built environment on physical activity participation is also mixed. Mixed land use is associated with increasing the likelihood of regular physical activity and discouraging the probability of physical inactivity, but also discourages occasional physical activity. A surprising negative association between street connectivity and regular physical activity, and positive association between street connectivity and physical inactivity, is found. However, street connectivity increases the likelihood of participating in occasional physical activity. On the other hand, density and centeredness are not significant factors in facilitating or

hindering participation in physical activity or inactivity. While there is some evidence of the role of the built environment on physical activity, the findings are not consistent across a broad measure of the built environment.

The likelihood of participating in regular physical activity is also related to social factors. Interestingly, while there is no statistical difference in regular physical activity between married and unmarried individuals (though unmarried individuals are less likely to be physically inactive and occasionally active), the tendency of Hispanics and Blacks and females to be more active is an interesting finding. This may be due to work-related physical activity and/or the activity involved in caring for young children, although the data used for this study did not allow for this detailed of an examination. An increase in the number of dependents in the household decreases the likelihood of engaging in regular physical activity but increases the probability of occasional physical activity. As the number of dependents increases, the probability of not being physically active also increases. The overall finding is that physical activity is endogenous, and that income, demographic factors, natural amenities and the built environment partially or fully play a role, indicating the importance of both personal and surrounding exogenous factors in shaping physical activity choices. Policies related to public health for managing obesity will need to consider these aspects in designing effective interventions to encourage more physical activity.

With regard to income, individuals who engage in occasional and regular physical activity are more likely to have higher incomes than those who are not physically active. As mentioned earlier, while the measurement of physical activity may limit the broad nature of this conclusion, it can be said that physical activity has an income dividend. Results observed in prior studies are also found in this study: there are gender, educational attainment and race disparities in income. Blacks and Hispanics earn less compared to non-Blacks and non-

Hispanics, females earn less than males, and individuals with at least some college education earn more than individuals with a high school diploma or less.

Overall, this study contributes to the literature in a number of ways. First, by testing factors in the context of 83 metropolitan areas of the US, over a course of 10 years (with 6 years of observation) in longitudinal data helps verify the validity of prior conclusions in the context of wider metropolitan areas. Second, simultaneity among BMI, physical activity and income is a contribution, one with significant implications to health policy as discussed above. Third, by examining the impact of natural amenities on obesity in metropolitan areas, this study sheds some light on the relevance of the impact of natural amenities on health.

6.2. Recommendations

Based on findings in this study, the following recommendations are suggested.

1. The study indicates that not all built environment measures are associated with lower BMI and promoting physical activity. Thus, from a public health perspective, investments to change or alter the existing built environment structure without identifying their effectiveness at reducing obesity or promoting physical activity may not accomplish the intended purpose. Therefore, policy makers should identify and invest in changes to the built environment that are most effective in reducing obesity and promoting physical activity, particularly those that enhance density.
2. The availability of outdoor recreational opportunities does not necessarily translate into usage of these resources by residents of the county in which they are located for the purpose of engaging in physical activity and weight reduction. Therefore, understanding why local residents may or may not be taking advantage of these local resources is

important in encouraging their usage through outreach, promotion and awareness programs.

3. If the goal of physical activity is to reduce weight, it is important to educate people about the intensity and frequency of physical activity required for successful weight control over time. The evidence from this study that a short duration of activity is not related with lower BMI may suggest that physical activity alone is not sufficient in tackling BMI, and physical activity intensity and duration is an important factor. Therefore, the level and length of time that is necessary to be spent being physically active to reduce BMI will need to be widely known through proper awareness programs.
4. The fact that social factors, such as race, gender and marital status, are related with weight suggests a more complex set of factors shape obesity. This is particularly so for minorities, women and married individuals who have a higher tendency for higher weight. Obesity management policies and initiatives should therefore pay attention to these groups in society when putting in place effective interventions. .
5. The findings that physical activity is endogenous, and that income, social factors, family size, natural amenities and the built environment partially or fully matter signals the importance of both personal and surrounding exogenous factors in shaping physical activity choices. Policies that target physical activity to managing obesity will need to consider these underlying drivers of physical activity in designing effective programs.

6.3. Limitations of the Study

One limitation of this study is the sample used for analysis. Although 12,686 individuals were initially included, eliminating those without location information and limiting the study to metro areas reduced the sample to only 2,594 individuals. This was necessary to allow for analysis of

measures of sprawl and natural amenities and recreational opportunities. Once individuals with missing data for variables of interest were also eliminated, only 1,768 individuals remained on which to base the estimations. This could mean there is bias in the results due to unknown characteristics of those who were eliminated, however, the sample is generally representative of national characteristics considered here.

According to the U.S. Department of Health and Human Services 2008 physical activity guidelines, the recommended weekly minutes of physical activity for adults is 150 minutes of moderate physical activity or 75 minutes of vigorous physical activity. However, NLSY79 data only asked questions about the frequency of moderate and vigorous exercise of ten minutes or more. The questions do not specify whether the individuals follow the recommended level of exercise, and do not specify how many total minutes per week they exercise. In both the moderate and vigorous exercise questions, equal weight is given for participating in a physical activity for more than ten minutes regardless of the duration. In addition, the physical activity variable does not distinguish between leisure time physical activities or job-related physical activity.

Another limitation of the study is related to the study area in consideration. The prevalence of obesity is a nationwide issue, and may be even more of a problem in rural areas, however, due to lack of available data the study is only focused on U.S. metro areas. The prevalence of obesity varies across states and it is usually higher in states with high poverty rates. The potential influence of state or county level poverty is not included in this study.

Even though BMI is widely used in research to measure obesity and overweight there are some limitations associated with it. The BMI calculation does not take into account factors such as age, gender or muscle mass. In addition, it does not distinguish between lean body mass and

fat mass. As a result, some people, such as heavily muscled athletes, may be considered overweight or obese even though they don't have a high percentage of body fat. Similarly in elderly people, BMI may appear normal even though they have less muscle mass due to aging. Despite the limitations of BMI it is an accepted measure of obesity and overweight.

The features of the physical environment can change over time due to natural or human-made causes. This study used the National Outdoor Recreation Supply Information System (NORSIS) for constructing natural amenity and recreational facilities indices. Even though it is a comprehensive data set, it is not the most recent data and as a result it might not reflect the current status of the amenities and natural environment. The same is true with regard to sprawl measures. Since recent natural environment and sprawl data are not available, the study assumes there is no change in those variables in recent years. In addition, due to lack of data availability of indoor recreational opportunities were not included in the study.

6.4. Future Research

This study could be expanded further in a number of ways. Addressing the limitations of the study could be a starting point to advance the existing body of knowledge regarding the characteristics of neighborhoods and obesity. Focusing on the following will improve the existing literature.

1. Future studies to understand the effect of sprawl and natural amenities on obesity should focus on using smaller scale and more recent sprawl and natural amenity data for better understanding of the relationships. Moreover, estimating more specific natural amenities not aggregated indices may provide better understanding of their contributions.
2. Social support and social capital of neighborhoods where many people are seen exercising may promote physical activity. As such it is important to investigate the

effects of social capital on physical activity and obesity. Moreover, safety is another factor that may promote or hinder outdoor physical activity, thus, it should be examined for further understanding.

3. Future research should be focused at a local level in order to have a precise measure of obesity, physical activity and environmental factors. Doing this research can provide a strong foundation for understanding the interaction between the behavior of individuals and their neighborhoods.
4. Future studies should examine the impact of obesity on long term income to better understand its overall economic implications.

References

- Averett, S. and S. Korenman. 1999. "Black White Differences in Social and Economic Consequences of Obesity." *International Journal of Obesity* 23: 166–173.
- Ball, K., K. Lamb, N. Travaglini, and A. Ellaway. 2012. "Street Connectivity and Obesity in Glasgow, Scotland: Impact of Age, Sex and Socioeconomic Position." *Health & Place* 18(6): 1307–1313.
- Bassett D.R., J. Pucher, R. Buehler, D.L. Thompson, and S.E. Crouter. 2008. "Walking, Cycling, and Obesity Rates in Europe, North America, and Australia." *Journal of Physical Activity and Health* 5: 795-814.
- Becker, G.S. 1965. "A Theory of the Allocation of Time." *The Economic Journal*, 75(299): 493-517.
- Becker, G.S. 1973. *The Economics of Discrimination*. The University of Chicago Press.
- Bhattacharya, J. and N. Sood. 2004. "Health Insurance, Obesity, and Its Economic Costs." The Economics of Obesity, Economic Research Service, USDA.
- Burchfield, M., H.G. Overman, D. Puga, and M.A. Turner. 2006. "Causes of Sprawl: A Portrait from Space." *Quarterly Journal of Economics* 121(2): 587–633.
- Burgoine, T., S. Alvanides, and A.A. Lake. 2011. "Assessing the Obesogenic Environment of North East England." *Health & Place* 17(3): 738–747.
- Catlin, A., C. Cowan, M. Hartman, and S. Heffler. 2008. "National Healthcare Spending in 2006." *Health Affairs* 27: 14 -29.
- Cai, L. 2010. "The Relationship between Health and Labor Force Participation: Evidence from a Panel Data Simultaneous Equation Model." *Labor Economics* 17:77-90.
- Cawley, J. 2000. "An Instrumental Variables Approach to Measuring the Effect of Body Weight on Employment Disability." *Health Services Research* 35: 1159 –1179.
- Cawley, J. 2004. "The Impact of Obesity on Wages." *Journal of Human Resource* 39(2): 451–474.
- Cawley, J, E. Han and E. Norton, 2009. "Obesity and Labor Market Outcomes among Legal Immigrants to the United States from Developing Countries." *Economics and Human Biology* 7: 153–164

- Center for Disease Control and Prevention (CDC). 2010. *Behavioral Risk Factor Surveillance System (BRFSS)*. Available at http://www.cdc.gov/obesity/downloads/obesity_trends_2010.ppt. Accessed April 20, 2011.
- Centers for Disease Control and Prevention (CDC). 2012. *Overweight and Obesity*. Available at <http://www.cdc.gov/obesity/adult/causes/index.html>. Accessed September, 2012.
- Chou, S., M. Grossman and H. Saffer. 2004. “An Economic Analysis of Adult Obesity: Results from the Behavioral Risk Factor Surveillance System.” *Journal of Health Economics* 23(3): 565-587.
- Davison, K. and C.T. Lawson. 2006. “Do Attributes in the Physical Environment Influence Children’s Physical Activity? A Review of the Literature.” *International Journal of Behavioral Nutrition and Physical Activity* 3(19).
- Deller, S.C., T-H Tsai, D.W. Marcouiller, and D.B.K. English, 2001. “The Role of Amenities and Quality of Life in Rural Economic Growth.” *American Journal of Agricultural Economics* 83(2):352-365.
- Drewnowski, A. 2004. “Obesity and the Food Environment Dietary Energy Density and Diet Costs.” *American Journal of Preventive Medicine* 27(3S): 154–162.
- Dunbar-Jacob, J.M, Schlenk E.A, Burke L.E, Matthews J.T. 1998. “Predictors of Patient Adherence: Patient Characteristics.” In Shumaker S.A, Schron E.B, Ockene J.K and McBee W.L (Eds.) *The Handbook of Health Behavior Change*, 2nd Edition. New York: Springer Publishing Company, 491-511.
- Eid, J., H.G. Overman, D. Pugad, and M.A. Turner. 2008. “Fat City: Questioning the Relationship between Urban Sprawl and Obesity.” *Journal of Urban Economics* 63: 385–404.
- Ewing, R. 1997. “Is Los Angeles-style Sprawl Desirable?” *Journal of the American Planning Association* 63(1): 107- 126.
- Ewing, R., R. Pendall, and D. Chen. 2002. “Measuring Sprawl and its Impact.” *Smart Growth America*, Washington, D.C.
- Ewing, R., T. Schmid, R. Killingsworth, A. Zlot, and S. Raudenbush. 2003. “Relationship between Urban Sprawl and Physical Activity, Obesity, and Morbidity.” *American Journal of Health Promotion* 47–57.
- Ewing, R., R.C. Brownson and D. Berrigan. 2006. “Relationship between Urban Sprawl and Weight of United States Youth.” *American Journal of Preventive Medicine* 31 (6): 464–474.

- Frank, L.D., M. A. Andresen, and T.L. Schmid. 2004. "Obesity Relationships with Community Design, Physical Activity, and Time Spent in Cars." *American Journal of Preventive Medicine* 27: 87–96.
- Frank, L.D., P.O. Engelke, and T.L. Schmid. 2003. *Health and Community Design: The Impact of the Built Environment on Physical Activity*. Washington, DC: Island Press.
- Finkelstein, E.A, J.G. Trogon, J.W. Cohen, and W. Dietz. 2009. "Annual medical spending attributable to obesity: Payer- and service-specific estimates." *Health Affairs* 28(5): w822-w831.
- Frumkin, H. 2002. "Urban Sprawl and Public Health." *Public Health Report* Vol. 117.
- Galster, G., R. Hanson, M. Ratcliffe, H. Wolman, S. Coleman, and J. Freihage. 2001. "Wrestling Sprawl to the Ground: Defining and Measuring an Elusive Concept," *Housing Policy Debate* 12(4): 685.
- Garden, F.L. and B.B. Jalaludin. 2008. "Impact of Urban Sprawl on Overweight, Obesity and Physical Activity in Sydney, Australia." *Journal of Urban Health: Bulletin of the New York Academy of Medicine* 86(1): 19-30.
- Geurs, K.T. and B. van Wee. 2004. "Accessibility Evaluation of Land-use and Transport Strategies: Review and Research Directions." *Journal of Transport Geography* 2: 127–140.
- Giles-Corti B., and R.J. Donovan. 2002. "The Relative Influence of Individual, Social and Physical Environment Determinants of Physical Activity." *Social Science and Medicine*. 54(12): 1793-1812.
- Gillham, G. 2002. *The Limitless City: A Primer on the urban sprawl debate*. Washington, DC: Island Press.
- Gortmaker, S.L., A. Must, J.M. Perrin, A.M. Sobol, and W. Dietz. 1993. "Social and Economic Consequences of Overweight in Adolescence and Young Adulthood." *New England Journal of Medicine* 329 (14): 1008–1012.
- Greene, W.H. 2000. *Econometric Analysis*, 4th ed., Upper Saddle River, NJ: Prentice Hall.
- Greve, J. 2008. "Obesity and Labor Market Outcomes in Denmark." *Economics and Human Biology* 6: 350–362.
- Grossman, M. 1972. "On the Concept of Health Capital and the Demand for Health." *The Journal of Political Economy* 80(2): 223-255.
- Gujarati, D.N. 2003. *Basic Econometrics*, 4th ed. Singapore: McGraw Hill.

- Hammond, R.A., and R. Levine. 2010. "The Economic Impact of Obesity in the United States." *Diabetes, Metabolic Syndrome and Obesity* 3: 285–295.
- Handy, S.L., M.G. Boarnet, R. Ewing, and R.E. Killingsworth. 2002. "How the Built Environment Affects Physical Activity: Views from Urban Planning." *American Journal of Preventive Medicine* 23(2S): 64 –73.
- Hill, J.O. and F.L.Trowbridge. 1998. "Childhood Obesity: Future Directions and Research Priorities." *Pediatrics* 101: 570-574.
- Hsiao, C. 2003. *Analysis of Panel Data*. 2nd ed. Cambridge: Cambridge University Press.
- Huang, T.T. and T.A. Glass. 2008. "Transforming Research Strategies for Understanding and Preventing Obesity." *Journal of the American Medical Association* 300: 1811–1813.
- Jackson, R.J., and C. Kochtitzky. 2002. *Creating a Healthy Environment: The Impact of the Built Environment on Public Health*." Centers for Disease Control and Prevention.
- Jensen, P., J. Greve and T. Tranaes. 2005. "Statistical Discrimination of Disabled Workers?" Working paper presented at SOLE/EALE World Conference, June.
- Johnson, M.P. 2001. "Environmental Impacts of Urban Sprawl: A Survey of the Literature and Proposed Research Agenda." *Environment and Planning A* 3: 717–735.
- Keehan, S., A. Sisko, and C. Truffer. 2008. "Health Spending Projections through 2017: The Baby-boom Generation is Coming to Medicare." *Health Affairs* 27(2):w145–w155.
- Kelly-Schwartz, A.C., J. Stockard, S. Doyle, and M. Schlossberg. 2004. "Is Sprawl Unhealthy? A Multilevel Analysis of the Relationship of Metropolitan Sprawl to the Health of Individuals." *Journal of Planning Education and Research* 24:184-196.
- Kwang-Koo, K. D.W. Marcouiller and S. C. Deller. 2005. "Natural Amenities and Rural Development: Understanding Spatial and Distributional Attributes." *Growth and Change* 36(2):273-297.
- Kennedy, P. 2003. *Introduction to Econometrics. A guide to econometrics*. Cambridge: MIT Press.
- Lakdawalla, D. and T. Philipson. 2002. "The Growth of Obesity and Technological Change: A Theoretical and Empirical Examination." NBER Working Paper, No. 8946.
- Lee, I-M., R. Ewing, and H.D. Sesso. 2009. "The Built Environment and Physical Activity Levels." *American Journal of Preventive Medicine* 37(4): 293-298.
- Lopez, R. 2004. "Urban Sprawl and Risk for Being Overweight or Obese." *American Journal of Public Health* 94: 1574–1579.

- Lopez-Zetina, J., H. Lee, and R. Friis. 2006. "The Link between Obesity and the Built Environment. Evidence from an Ecological Analysis of Obesity and Vehicle Miles of Travel in California." *Health Place* 12: 656–664.
- Marcouiller, D.W., K.K. Kim and S.C. Deller. 2004. "Natural Amenities, Tourism and Income Distribution" *Annals of Tourism Research* 31: 1031-1050.
- Mei, Z., L.M. Grummer-Strawn, A. Pietrobelli, A. Goulding, A. M.I. Goran and W.H. Dietz. 2002. "Validity of Body Mass Index Compared with Other Body-Composition Screening Indexes for the Assessment of Body Fatness in Children and Adolescents." *American Journal of Clinical Nutrition* 75(6): 978-985.
- McDonald, K.N., J.M. Oakes, and A. Forsyth. 2012. "Effect of Street Connectivity and Density on Adult BMI: Results from the Twin Cities Walking Study." *Journal of Epidemiology Community Health* 66(7): 636-640.
- McKinnon, R.A., J. Reedy, S.L. Handy, and A.B. Rodgers. 2009. "Measuring the Food and Physical Activity Environments Shaping the Research Agenda." *American Journal of Preventive Medicine* 36(4S): S81-S85.
- Merrill, R.M., M. Friedrichs, and L. Larsen. 2002. "Perceptions of Healthy Behaviors versus Health Practices." *Health Promotion Practice* 3(4): 497-500.
- Merrill, R.M., E. Shields, G.L. White, and D. Druce. 2005. "Climate Conditions and Physical Activity in the United States." *American Journal of Health Behavior* 29(4): 371-381.
- Michimi, A. and M.C. Wimberly. 2012. "Natural Environments, Obesity, and Physical Activity in Nonmetropolitan Areas of the United States." *The Journal of Rural Health* 28(4) 398–407.
- Mitra, A. 2001. "Effects of Physical Attributes on the Wages of Males and Females." *Applied Economics Letter* 8: 731–735.
- Morris, S. (2006) *The Impact of Obesity on Employment*. Imperial College London, United Kingdom.
- National Center for Chronic Disease Prevention and Health Promotion, Center for Disease Control and Promotion. 2009.. *The Power of Prevention*. Available at <http://www.cdc.gov/chronicdisease/pdf/2009-power-of-prevention.pdf>
- National Center for Health Statistics, CDC. 2004. *Trends in Intake of Energy and Macronutrients*, United States, 1971-2000. February.

- National Center for Health Statistics. 2012. *Healthy People 2010 Final Review*. Hyattsville, MD, December.
- National Institute of Health, National Heart Lung and Blood Institute. Obesity Education Initiative. 1998. *Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults*. SEPTEMBER. Available online: http://www.nhlbi.nih.gov/guidelines/obesity/ob_gdlns.pdf
- Northridge, M.E, E. Sclar, P. Biswas. 2003. "Sorting out the Connections between the Built Environment and Health: A Conceptual Framework for Navigating Pathways and Planning Healthy Cities." *Journal of Urban Health* 80:556-568.
- Ogden, C.L, M.M. Lamb, M.D. Carroll and K.M. Flegal. 2010. *Obesity and Socioeconomic Status in Adults: United States, 2005–2008*. U.S. Department of Health and Human Services Centers for Disease Control and Prevention, National Center for Health Statistics, NCHS Data Brief No. 50, December.
- Ogden, C.L., M.D. Carroll, K.K. Brian, and K.M. Flegal. 2012. *Prevalence of Obesity in the United States, 2009–2010*. Centers for Disease Control and Prevention National Center for Health Statistics, NCHS Data Brief, No. 82, January.
- Paeratakul, S., J.C. Lovejoy, D.H. Ryan, and G.A. Bray. 2002. "The Relation of Gender, Race and Socioeconomic Status to Obesity and Obesity Morbidities in a Sample of US Adults." *International Journal of Obesity and Related Metabolic Disorder* 26: 1205–1210.
- Papas, M.A., A.J. Alberg, R. Ewing, K.J. Helzlsouer, T.L. Gary, and A.C. Klassen. 2007. "The Built Environment and Obesity." *Epidemiologic Reviews* 29: 129–143.
- Plantinga, A.J. and S. Bernell. 2005. "A Spatial Economic Analysis of Urban Land Use and Obesity." *Journal of Regional Science*, 45(3): 473–492.
- Plantinga, A. J., and S. Bernell. 2007. "The Association between Urban Sprawl and Obesity: Is It a Two-Way Street?" *Journal of Regional Science* 47(5): 857–879.
- Puhl, R., and K.D. Brownell. 2001. "Bias, Discrimination, and Obesity." *Obesity Research* 9: 788–805.
- Rajamani, J., C.R. Bhat, S. Handy, G. Knaap, and Y. Song. 2003. "Assessing the Impacts of Urban form Measures in Non-work Trip Mode Choices after Controlling for Demographic and Level of Service Effects." *Transportation Research Record* 1831: 158-165.
- Register, C. A. and D. R. Williams. 1990. "Wage Effect of Obesity among Young Workers." *Social Science Quarterly* 71: 130-41.

- Rooth, D.O. 2007. "Evidence of Unequal Treatment in Hiring Against Obese Applicants: A Field Experiment. Discussion Paper NO. 2775.
- Rosenberger, R.S., Y. Sneh, T. Phipps and R. Gurvitch. 2005. "A Spatial Analysis of Linkages between Health Care Expenditure, Physical Inactivity, Obesity and Recreation Supply." *Journal of Leisure Research* 37(2): 216-235.
- Rubenstein, A. H. 2005. "Obesity: A Modern Epidemic." *Transaction of the American Clinical and Climatological Association* 116:103-11.
- Saelens, B.E., J.F. Sallis, J.B. Black and D. Chen. 2003. "Neighborhood- based Differences in Physical Activity: An Environment Scale Evaluation." *American Journal of Public Health* 93: 1552-1558.
- Sallis, J.F. and K. Glanz. 2006. "The Role of Built Environments in Physical Activity, Eating, and Obesity in Childhood." *The Future of Children* 16(1): 89-108.
- Schmeiser, M.D. 2009. "Expanding Wallets and Waistlines: The Impact of Family Income on the BMI of Women and Men Eligible for the Earned Income Tax Credit." *Health Economics* 18:1277-1294.
- Sturm, R., J.S. Ringel, and T. Andreyeva. 2004. "Increasing Obesity Rates and Disability Trends." *Health Affairs* 23: 199-205.
- The Sierra Club. 1999. "The dark side of the American Dream: the costs and consequences of suburban sprawl." The Sierra Club, San Francisco, CA. Available at: <http://www.sierraclub.org>.
- Tunceli, K., L. Kemeng and L. K. Williams. 2006. "Long-Term Effects of Obesity on Employment and Work Limitations among U.S. Adults, 1986 to 1999." *Obesity* 14(9): 1637-1646.
- U.S. Department of Commerce, U.S. Census Bureau. 2010. Available at: http://www.census.gov/newsroom/releases/archives/families_households/cb10-174.html. Accessed May 2013.
- U.S. Department of Commerce, U.S. Census Bureau. 2011. Population Distribution and Change: 200 to 2010. Available at: <http://www.census.gov/prod/cen2010/briefs/c2010br-01.pdf>. Accessed May 2013.
- U.S. Department of Commerce, U.S. Census Bureau. Current Population survey on Educational Attainment. 1998-2008. Available at: <https://www.census.gov/hhes/socdemo/education/data/cps/>. Accessed May 2013.

- U.S. Department of Commerce, U.S. Census Bureau. Population Estimates. 2009. Available at: <http://www.census.gov/popest/data/counties/totals/2009/CO-EST2009-01.html>. Accessed May 2012.
- U.S. Bureau of Labor Statistics. Local Area Unemployment Rate Statistics. 1998-2008. Data Available at: <http://www.bls.gov/lau/#cntyaa>. Accessed August 2012.
- U.S. Bureau of Labor Statistics, National Longitudinal Surveys of Youth (NLSY79). 2010. Available at: <https://www.nlsinfo.org/investigator/pages/login.jsp>. Accessed January 15, 2012.
- USDA Forest Service, National Outdoor Recreation Supply Information System (NORSIS). 1997. Available upon request from USDA Forest Service Southern Station.
- U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion. 1996. *Physical Activity and Health: A Report of the Surgeon General*.
- U.S. Department of Health and Human Services, National Institutes of Health. 2010. *Overweight and Obesity Statistics*.
- U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Division of Nutrition, Physical Activity, and Obesity Physical Activity Guidelines for Adults . 2008. 2008 *Physical Activity Guidelines for Americans*. Available at <http://www.health.gov/paguidelines/pdf/paguide.pdf>. Accessed October 2012.
- U.S. Department of Health and Human Services, National Institute of Health , Overweight and Obesity Statistics, Weight Control Information Network. 2012. Available at: Available at <http://win.niddk.nih.gov/publications/PDFs/stat904z.pdf>. Accessed December 2012
- U.S. Department of Housing and Urban Development (USHUD). 1999. *The State of the Cities*. Washington, DC, Third Annual Report.
- U.S. Social Security Administration. *Social Security Index*. 2012. Available at <http://www.ssa.gov/oact/cola/AWI.html>. Accessed May 2013.
- Wang, Y., and M. A. Beydoun. 2007. “The Obesity Epidemic in the United States-Gender, Age, Socioeconomic, Racial/Ethnic, and Geographic Characteristics: A Systematic Review and Meta-Regression Analysis.” *Epidemiologic Reviews* 29: 6–28.
- Wells, H.F. and J.C. Buzby. 2008. “Dietary Assessment of Major Trends in U.S. Food Consumption, 1970-2005.” *United States Department of Agriculture, Economic Research Service, Economic Information Bulletin*, No. 33.

Wen, M. and L.K. Jones. 2012. "The built environment and risk of obesity in the United States: Racial-ethnic disparities." *Health & Place* 18: 1314-1322.

Appendix

To maximize utility, consumers choose C, F, W, P subject to the income constraint. The Lagrangian for this problem is:

$$(A.1) \quad V = V(C, F, W, P) + \lambda(Y - P_C C - P_F F - P_W W - P_P P)$$

The first order condition for optimal choice of C, F, W and P are found by partially differentiating V with respect to the choice variables and the lagrangian multiplier.

$$(A.2) \quad \frac{\partial V}{\partial C} = V_C(C, F, W, P) - \lambda P_C = 0$$

$$(A.3) \quad \frac{\partial V}{\partial F} = V_F(C, F, W, P) - \lambda P_F = 0$$

$$(A.4) \quad \frac{\partial V}{\partial W} = V_W(C, F, W, P) - \lambda P_W = 0$$

$$(A.5) \quad \frac{\partial V}{\partial P} = V_P(C, F, W, P) - \lambda P_P = 0$$

$$(A.6) \quad \frac{\partial V}{\partial \lambda} = Y - P_C C - P_F F - P_W W - P_P P = 0$$

Totally differentiating each first order conditions and setting the total differential equal to zero, the system of equations can be expressed as:.

$$(A.7) \quad \begin{array}{l} V_{CC}dC + V_{CF}dF + V_{CW}dW + V_{CP}dP - P_C d\lambda = \lambda dP_C \\ V_{FC}dC + V_{FF}dF + V_{FW}dW + V_{FP}dP - P_F d\lambda = \lambda dP_F \\ V_{WC}dC + V_{WF}dF + V_{WW}dW + V_{WP}dP - P_W d\lambda = \lambda dP_W \\ V_{PC}dC + V_{PF}dF + V_{PW}dW + V_{PP}dP - P_P d\lambda = \lambda dP_P \\ -P_C dC - P_F dF - P_W dW - P_P dP \quad 0 = CdP_C + FdP_F + WdP_W + PdP_P - dY \end{array}$$

The matrix form of the system of total differential equations can be expressed as:

$$(A.8) \quad \begin{bmatrix} V_{CC} & V_{FC} & V_{CW} & V_{CP} & -P_C \\ V_{CF} & V_{FF} & V_{FW} & V_{FP} & -P_F \\ V_{WC} & V_{WF} & V_{WW} & V_{WP} & -P_W \\ V_{PC} & V_{PF} & V_{PW} & V_{PP} & -P_P \\ -P_C & -P_F & -P_W & -P_P & 0 \end{bmatrix} \begin{bmatrix} dC \\ dF \\ dW \\ dP \\ d\lambda \end{bmatrix} = \begin{bmatrix} \lambda dP_C \\ \lambda dP_F \\ \lambda dP_W \\ \lambda dP_P \\ CdP_C + FdP_F + WdP_W + PdP_P - dY \end{bmatrix}$$

The Impact of Income on Weight ($\frac{\partial W}{\partial Y}$) and Physical Activity ($\frac{\partial P}{\partial Y}$)

To find the effect of income (Y) on W and P, totally differentiate the system of total differential equations with respect to dY , by holding dP_C , dP_F , dP_W , and dP_P constant .

$$(A.9) \quad \begin{bmatrix} V_{CC} & V_{CF} & 0 & V_{CP} & -P_C \\ V_{FC} & V_{FF} & 0 & V_{FP} & -P_F \\ V_{WC} & V_{WF} & 0 & V_{WP} & -P_W \\ V_{PC} & V_{PF} & 0 & V_{PP} & -P_P \\ -P_C & -P_F & -1 & -P_P & 0 \end{bmatrix} \begin{bmatrix} \frac{dC}{dY} \\ \frac{dF}{dY} \\ \frac{dW}{dY} \\ \frac{dP}{dY} \\ \frac{d\lambda}{dY} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

The comparative statics derivatives in matrix form are expressed as:

$$(A.10) \quad \frac{\partial W}{\partial Y} = \frac{|A_1|}{|A|} = \frac{\begin{bmatrix} V_{CC} & V_{CF} & 0 & V_{CP} & -P_C \\ V_{FC} & V_{FF} & 0 & V_{FP} & -P_F \\ V_{WC} & V_{WF} & 0 & V_{WP} & -P_W \\ V_{PC} & V_{PF} & 0 & V_{PP} & -P_P \\ -P_C & -P_F & -1 & -P_P & 0 \end{bmatrix}}{\begin{bmatrix} V_{CC} & V_{CF} & V_{CW} & V_{CP} & -P_C \\ V_{FC} & V_{FF} & V_{FW} & V_{FP} & -P_F \\ V_{WC} & V_{WF} & V_{WW} & V_{WP} & -P_W \\ V_{PC} & V_{PF} & V_{PW} & V_{PP} & -P_P \\ -P_C & -P_F & -P_W & -P_P & 0 \end{bmatrix}}$$

$$|A_1| =$$

$$V_{CC}V_{FF}[V_{WP}P_P - V_{PP}P_W] + V_{CC}V_{FP}[-V_{WF}P_P + V_{PF}P_W] + V_{CC}P_F[V_{WF}V_{PP} - V_{WP}V_{PF}] - V_{CF}V_{FC}[V_{WP}P_P - V_{PP}P_W] - V_{CF}V_{FP}[-V_{WC}P_P + V_{PC}P_W] + V_{CF}P_F[V_{WC}V_{PP} - V_{WP}V_{PC}] + V_{CF}V_{FC}[-V_{WF}P_P + V_{PF}P_W] + V_{CF}V_{FF}[V_{WC}P_P + V_{PC}P_W] - V_{CP}P_F[-V_{WC}V_{PF} + V_{WF}V_{PC}] - V_{FC}P_C[V_{WF}V_{PP} - V_{WP}V_{PF}] + V_{FF}P_C[V_{WC}V_{PP} - V_{WP}V_{PC}] + V_{FP}P_C[-V_{WC}V_{PF} - V_{WF}V_{PC}]$$

$$|A| =$$

$$V_{CC}[-V_{WW}P_P^2 + V_{WP}P_P P_W - P_W(-V_{PW}P_P + V_{PP}P_W)] - V_{CC}V_{FW}[-V_{WF}P_P^2 + V_{WP}P_P P_F - P_W(-V_{PF}P_P + V_{PP}P_F)] + V_{CC}V_{FP}[V_{WF}(-P_P P_W) + V_{WW}P_P P_F - P_W(-V_{PF}P_W + V_{PW}P_F)] + V_{CC}P_F[V_{WF}(-V_{PW}P_P + V_{PP}P_W) - V_{WW}(-V_{PF}P_P + V_{PP}P_F) + V_{WP}(V_{PF}P_W + V_{PW}P_F)] - V_{CF}V_{FC}[-V_{WW}P_P^2 - V_{WP}P_P P_W - P_W(-V_{PW}P_P + V_{PP}P_W)] + V_{CF}V_{FW}[-V_{WC}P_P^2 - V_{WP}P_P P_C - P_W(-V_{PC}P_C + V_{PP}P_C)] - V_{CF}V_{FP}[-V_{WC}P_P P_W + V_{WW}P_P P_W - P_W(-V_{PC}P_W + V_{PW}P_C)] - V_{CF}P_F[V_{WC}(-V_{PW}P_P + V_{PP}P_W) - V_{WW}(-V_{PC}P_P + V_{PP}P_C) + V_{WP}(-V_{PC}P_W + V_{PW}P_C)] + V_{CW}V_{FC}[-V_{WF}P_P^2 + V_{WP}(P_P P_F) - P_W(-V_{PF}P_P + V_{PP}P_F)] - V_{CW}V_{FF}[-V_{WC}P_P^2 + V_{WP}P_P P_C - P_W(-V_{PC}P_P + V_{PP}P_C)] + V_{CW}V_{FP}[-V_{WC}P_P P_F - V_{WF}(P_P P_C) - P_W(-V_{PC}P_F + V_{PF}P_C)] + V_{CW}P_F[V_{WC}(-V_{PF}P_P + V_{PP}P_F) - V_{WF}(-V_{PC}P_P + V_{PP}P_C) + V_{WP}(-V_{PC}P_F + V_{PF}P_C)] - V_{PC}V_{FC}[-V_{WF}P_P P_F + V_{WW}P_P P_F - P_W(-V_{PF}P_W + V_{PW}P_F)] + V_{CP}V_{FF}[-V_{WC}P_P P_P + V_{WW}P_P P_C - P_W(-V_{PC}P_W + V_{PW}P_C)] - V_{CP}V_{FW}[-V_{WC}P_P P_F + V_{WF}P_P P_C - P_W(-V_{PC}P_F + V_{PF}P_C)] - V_{CP}P_F[V_{WC}(-V_{PF}P_W + V_{PW}P_F) - V_{WF}(-V_{PC}P_W + V_{PW}P_C) + V_{WW}(-V_{PC}P_F + V_{PF}P_C)] - V_{FC}P_C[V_{WF}(-V_{PW}P_P + V_{PP}P_W) - V_{WW}(-V_{PF}P_P + V_{PP}P_F) + V_{WP}(-V_{PF}P_W + V_{PW}P_F)] + V_{FF}P_C[V_{WC}(-V_{PW}P_P + V_{PP}P_W) - V_{WW}(-V_{PC}P_P + V_{PP}P_C) + V_{WP}(-V_{PC}P_W + V_{PW}P_C)] - V_{FW}P_C[V_{WC}(-V_{PF}P_P + V_{PP}P_F) - V_{WF}(-V_{PC}P_P + V_{PP}P_C) + V_{WP}(-V_{PC}P_F + V_{PF}P_C)] + V_{FP}P_C[V_{WC}(-V_{WC}P_W + V_{PW}P_F) + V_{WF}(-V_{PC}P_W + V_{PW}P_C) + V_{WW}(-V_{PC}P_F + V_{PF}P_C)]$$

After solving the determinant, substituting and rearranging the mathematical

representation of the impact of income on weigh is:

$$\frac{\partial W}{\partial Y} = <, > 0$$

The Impact of Income on Physical Activity ($\frac{\partial P}{\partial Y}$)

Totally differentiate physical activity with respect to total income (dY) can be determined by

holding prices dP_C , dP_F , dP_W , and dP_P constant. $\frac{dP}{dY}$ can be determined following Cramer's rule.

The comparative statics in matrix form is given as:

$$(A.11) \quad \frac{\partial P}{\partial Y} = \frac{|A_2|}{|A|} = \frac{\begin{bmatrix} V_{CC} & V_{CF} & V_{CW} & 0 & -P_C \\ V_{FC} & V_{FF} & V_{FW} & 0 & -P_F \\ V_{WC} & V_{WF} & V_{WW} & 0 & -P_W \\ V_{PC} & V_{PF} & V_{PW} & 0 & -P_P \\ -P_C & -P_F & -P_W & -1 & 0 \end{bmatrix}}{\begin{bmatrix} V_{CC} & V_{CF} & V_{CW} & V_{CP} & -P_C \\ V_{FC} & V_{FF} & V_{FW} & V_{FP} & -P_F \\ V_{WC} & V_{WF} & V_{WW} & V_{WP} & -P_W \\ V_{PC} & V_{PF} & V_{PW} & V_{PP} & -P_P \\ -P_C & -P_F & -P_W & -P_P & 0 \end{bmatrix}}$$

Solving the above matrix provides;

$$|A_2| = V_{CC}V_{FF}[-V_{WW}P_P - V_{PW}P_W] - V_{CC}V_{FW}[-V_{WF}P_P + V_{PF}P_W] + V_{CC}P_F[-V_{WF}V_{PW} + V_{WW}V_{PF}] - V_{CF}V_{FC}[-V_{WW}P_P + V_{PW}P_W] + V_{CF}V_{FW}[-V_{WC}P_P + V_{PC}P_W] - V_{CF}P_F[-V_{WC}V_{PW} + V_{WW}V_{PC}] + V_{CW}V_{FC}[-V_{WF}P_P - V_{PF}P_W] - V_{CW}V_{FF}[-V_{WC}P_P + V_{PC}P_W] + V_{CW}P_F[-V_{WC}V_{PF} + V_{WF}V_{PC}] - V_{FC}P_C[-V_{WF}V_{PW} + V_{WW}V_{PF}] + V_{FF}P_C[-V_{WC}V_{PW} + V_{WW}V_{PC}] - V_{FW}P_C[-V_{WC}V_{PF} + V_{WF}V_{PC}]$$

$$\frac{\partial P}{\partial Y} = > 0$$

The Impact of the Opportunity Cost of Time on Weight $(\frac{\partial W}{\partial w})$ and *Physical Activity* $(\frac{\partial P}{\partial w})$.

To examine the implication of the opportunity cost of time, the basic utility maximization function is utilized, with total income (Y) decomposed into wage $w(T-N)$ and non-wage income I components. The opportunity cost of time is measured by wage rate (w).

The expanded Lagrangian with wage and non-wage income is;

$$(A.12) \quad V = V[C, H, U, W, P] + \lambda(I + w(T - N) - P_C C - P_H H - P_U U - P_W W - P_P P)$$

The system of total differential equations are expressed in matrix form as:

Equation shows the first order condition in matrix form.

$$(A.13) \quad \begin{bmatrix} V_{CC} & V_{FC} & V_{CW} & V_{CP} & -P_C \\ V_{CF} & V_{FF} & V_{FW} & V_{FP} & -P_F \\ V_{WC} & V_{WF} & V_{WW} & V_{WP} & -P_W \\ V_{PC} & V_{PF} & V_{PW} & V_{PP} & -P_P \\ -P_C & -P_F & -P_W & -P_P & 0 \end{bmatrix} \begin{bmatrix} dC \\ dF \\ dW \\ dP \\ d\lambda \end{bmatrix} = \begin{bmatrix} \lambda dP_C \\ \lambda dP_F \\ \lambda dP_W \\ \lambda dP_P \\ CdP_C + FdP_F + WdP_W + PdP_P - dI - wd(T-N) - (T-N)dw \end{bmatrix}$$

Totally differentiating physical activity with respect to the opportunity cost of time (w), and by keeping dP_C , dP_F , dP_W , dP_P , dI and $d(T-N)$ constant, the system of differential equations in matrix form is given as:

$$(A.14) \quad \begin{bmatrix} V_{CC} & V_{CF} & V_{CW} & V_{CP} & -P_C \\ V_{FC} & V_{FF} & V_{FW} & V_{FP} & -P_F \\ V_{WC} & V_{WF} & V_{WW} & V_{WP} & -P_W \\ V_{PC} & V_{PF} & V_{PW} & V_{PP} & -P_P \\ -P_C & -P_F & -P_W & -P_P & 0 \end{bmatrix} \begin{bmatrix} \frac{dC}{dw} \\ \frac{dF}{dw} \\ \frac{dW}{dw} \\ \frac{dP}{dw} \\ \frac{d\lambda}{dw} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ (T-N) \end{bmatrix}$$

The comparative statics derivative is given as;

$$(A.15) \quad \frac{\partial P}{\partial w} = \frac{|A_3|}{|A|} = \frac{\begin{bmatrix} V_{CC} & V_{CF} & V_{CW} & 0 & -P_C \\ V_{FC} & V_{FF} & V_{FW} & 0 & -P_F \\ V_{WC} & V_{WF} & V_{WW} & 0 & -P_W \\ V_{PC} & V_{PF} & V_{PW} & 0 & -P_P \\ -P_C & -P_F & -P_W & (T-N) & 0 \end{bmatrix}}{\begin{bmatrix} V_{CC} & V_{CF} & V_{CW} & V_{CP} & -P_C \\ V_{FC} & V_{FF} & V_{FW} & V_{FP} & -P_F \\ V_{WC} & V_{WF} & V_{WW} & V_{WP} & -P_W \\ V_{PC} & V_{PF} & V_{PW} & V_{PP} & -P_P \\ -P_C & -P_F & -P_W & -P_P & 0 \end{bmatrix}}$$

$$|A_3| =$$

$$\begin{aligned} & V_{CC}V_{FF}[-V_{WW}P_P(T-N) - V_{PW}P_W(T-N)] + V_{CC}V_{FW}[-V_{WF}P_P(T-N) - V_{PF}(T-N)P_W] + V_{CC}P_F[-V_{WF}V_{PW}(T-N) + V_{WW}V_{PF}(T-N)] - V_{CF}V_{FC}[-V_{WW}P_P(T-N) + V_{PW}P_W(T-N)] + \\ & V_{CF}V_{FW}[-V_{WC}P_P(T-N) + V_{PC}P_W(T-N)] - V_{CF}P_F[-V_{WC}V_{PW}(T-N) + V_{WW}V_{PC}(T-N)] + V_{CW}V_{FC}[-V_{WF}P_P(T-N) + V_{PF}P_W(T-N)] - V_{CW}V_{FF}[-V_{WC}P_P(T-N) + V_{PC}P_W(T-N)] + \\ & V_{CW}P_F[-V_{WC}V_{PF}(T-N) + V_{WF}V_{PC}(T-N)] - V_{FC}P_C[-V_{WF}V_{PW}(T-N) + V_{WW}V_{PF}(T-N)] + V_{FF}P_C[-V_{WC}V_{PW}(T-N) + V_{WW}V_{PC}(T-N)] - V_{FW}P_C[-V_{WC}V_{PF}(T-N) + V_{WF}V_{PC}(T-N)] \end{aligned}$$

Solving the determinants, substituting and rearranging gives:

$$\frac{\partial P}{\partial w} = <, > 0.$$

To analyze the effect of w on weight, (W), totally differentiate w with respect to W by holding dP_C , dP_F , dP_W , dP_P , dI and $d(T-N)$ constant. The comparative statics derivative is:

$$(A.16) \quad \frac{\partial W}{\partial w} = \frac{|A_4|}{|A|} = \frac{\begin{bmatrix} V_{CC} & V_{CF} & 0 & V_{CP} & -P_C \\ V_{FC} & V_{FF} & 0 & V_{FP} & -P_F \\ V_{WC} & V_{WF} & 0 & V_{WP} & -P_W \\ V_{PC} & V_{PF} & 0 & V_{PP} & -P_P \\ -P_C & -P_F & (T-N) & -P_P & 0 \end{bmatrix}}{\begin{bmatrix} V_{CC} & V_{CF} & V_{CW} & V_{CP} & -P_C \\ V_{FC} & V_{FF} & V_{FW} & V_{FP} & -P_F \\ V_{WC} & V_{WF} & V_{WW} & V_{WP} & -P_W \\ V_{PC} & V_{PF} & V_{PW} & V_{PP} & -P_P \\ -P_C & -P_F & -P_W & -P_P & 0 \end{bmatrix}}$$

$$|A_4| =$$

$$\begin{aligned} & V_{CC}V_{FF}[V_{WP}(T-N)P_P - V_{PP}P_W(T-N)] + V_{CC}V_{FP}[-V_{WF}(T-N)P_P + V_{PF}(T-N)P_W] + V_{CC}P_F[V_{WF}V_{PP}(T-N) - V_{WP}V_{PF}(T-N)] - V_{CF}V_{FC}[V_{WP}(T-N)P_P - V_{PP}P_W(T-N)] \\ & - V_{CF}V_{FP}[-V_{WC}(T-N)P_P + V_{PC}P_W(T-N)] - V_{CF}P_F[V_{WC}V_{PP}(T-N) - V_{WP}V_{PC}(T-N)] - V_{CP}V_{FC}[-V_{WF}P_P(T-N) + V_{PF}P_W(T-N)] + V_{CP}V_{FF}[-V_{WC}P_P(T-N) - V_{PC}P_W(T-N)] \\ & - V_{CP}P_F[-V_{WC}V_{PF}(T-N) + V_{WF}V_{PC}(T-N)] - V_{FC}P_C[V_{WF}V_{PP}(T-N) - V_{WP}V_{PF}(T-N)] + V_{FF}P_C[V_{WC}V_{PP}(T-N) - V_{WP}V_{PC}(T-N)] + V_{FP}P_C[-V_{WC}V_{PF}(T-N) + V_{WF}V_{PC}(T-N)] \end{aligned}$$

Solving equation A.16 provides:

$$\frac{\partial W}{\partial w} = <0.$$