WALKING AND HEALTHY? ON THE RELATIONSHIP AMONG UTILITARIAN WALKING, HEALTH, AND RESIDENTIAL CHOICE

By

MIGUEL A LUGO ORTIZ

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2015

ACKNOWLEDGMENTS

I thank my dissertation committee, Dr. Siva, Dr. Steiner, Dr. Washburn, Dr. Lily, and Dr. Yin, for their attentive supervision and mentoring in this endeavor. I extend this appreciation to my work colleagues at Weil Hall, past and present, for their support, and my family and friends in Gainesville, in Puerto Rico, and abroad for their encouragement through this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

WALKING AND HEALTHY? ON THE RELATIONSHIP AMONG UTILITARIAN WALKING, HEALTH, AND RESIDENTIAL CHOICE

By

Miguel A. Lugo Ortiz

August 2015

Chair: Sivaramakrishnan Srinivasan
Major: Civil Engineering

Walking as a mode of transportation, or utilitarian walking, is available to 92.7% of the adults in the United States and as such may be identified as one of the most available modes. While the qualitative impact of walking on health can be inferred from the energy-balance model, empirical studies are necessary to ascertain the magnitude of this impact. Further, such empirical models should capture the effects of land-use patterns. This study addressed this gap in the literature by using data-fusion techniques to merge transportation and health surveys and subsequently develop comprehensive models of land-use, walking, and health. Two recent large cross-sectional surveys from the U.S., the American Time Use Survey and the National Household Travel Survey, are employed to explore these relations by following public health policy recommendations of continuous walking thresholds of at least 10 minutes. The national models showed a negative impact of poor health on utilitarian walking. As an application, a full set of models that links land-use, walking, and health was demonstrated for Florida adults under the age of 70. The applied model associated the presence of pedestrian-oriented facilities up to a half-mile from household with lower BMIs after controlling for socioeconomic characteristics.

CHAPTER 1
INTRODUCTION

Walking as a mode of transportation is available (defined by physical capability) to 92.7% of the adults *(1)* in the United States and as such may be identified as one of the most “available” modes. At the same time, research *(2)* also indicates that destinations are often not accessible by walking, limiting the use of walking as a means of transportation. In this context, walking is estimated to account for 10.4% of all trips in the United States *(3)* and at least 2.8% of all work trips *(4).* As many as 69% of short trips are undertaken by private motorized vehicles *(5),* indicating a potential for increased shifts towards non-motorized modes. Clearly, there are also several reasons to incentivize walking and these include environmental needs *(6, 7)*, alleviation of urban congestion *(8),* and economic impact *(5).* More recently, there has also been an increasing interest in promoting walking from the public health standpoint. Walking can be a catalyst for physical activity *(9)* and thereby discourage sedentary lifestyles and associated ills.

Sedentary lifestyles are characterized as unhealthy as they lead to higher risks of obesity, disease, health conditions and shorter lives *(10, 11).* Walking for transportation, as a moderate-pace activity, can be a strategy used to facilitate daily physical activity, required for a healthy (active) lifestyle in adults. This is supported by the “energy balance” model *(12)* which captures the physiological relationship between walking and health measures such as BMI: walking results in a greater expenditure of energy thereby limiting the accumulation of additional adipose tissue (body fat). Current U.S. Department of Health and Human Services/Centers for Disease Control *(13)* recommendations include at least 10 minutes at a time of moderate-intensity aerobic (e.g., brisk walking) for a total of 150 minutes every week and two days of muscle-strengthening activities. In contrast, the U.S. Surgeon General recommends 30 minutes of moderate physical activities on most days of the week *(14).*

While the qualitative impact of walking on health can be inferred from the energy-balance model, empirical studies are necessary to ascertain the magnitude of the impact of walking on health. Such empirical models should capture the effects of land-use patterns so that these models may be used to determine if and how public health can be improved via changes in land-use patterns. While studies on built environment, walking and physical health exists in the fields of public health, urban planning, and transportation, comprehensive models addressing all their aspects are limited. Arguably one of the major impediments to the development of such models is the lack of data. Transportation surveys while providing details travel partners and spatial information at a fine resolution do not include health data. Health surveys often do no collect travel behavior data and/or do not provide detailed spatial resolution to construct land-use descriptors.

This study addressed this gap in the literature by using data-fusion techniques to first merge transportation and health surveys. The objectives of this research were to develop a methodology for comprehensive disaggregate models of land-use, utilitarian walking, and health of U.S. adults.

This study employed two recent large cross-sectional surveys from the U.S.: the American Time Use Survey (ATUS) and the National Household Travel Survey (NHTS). Once the health data from the ATUS has been suitably linked to NHTS respondents, the final component was to develop a full set of models that linked land-use, walking, and health. Such models can be useful in evaluating the impacts of land-use changes on walking patterns, and ultimately, on public health. The survey samples are not limited to any one geographic area or to a specific population segment. At the same time individual-level time use and health data are modeled.

The rest of this dissertation is organized as follows. Chapter 2 presents a summary of relevant literature. Chapter 3 presents models relating walking to health using the American Time Use Surveys. Chapter 4 presents a data fusion approach to link health and travel data. Chapter 5 provides an application of data fusion techniques in a Florida sample, along with built environment descriptors. Finally, chapter 6 offers a conclusion and recommendations for further study.

CHAPTER 2
LITERATURE REVIEW

Introduction

The purpose of this chapter is to first present a synthesis of empirical findings on the inter-dependencies among physical activity, built environment, and health. Subsequently, literature on data fusion is presented as such methods are required to link health and travel surveys to assemble a comprehensive dataset for modeling the relationships.

The Physical Activity, Built Environment, and Health Framework

Physical activity is defined as “bodily movement produced by the contraction of skeletal muscle that increases energy expenditure above the basal level” *(15).* Physical activity is often characterized by context in which occurs: occupational/mandatory (e.g., work or school), household or maintenance (e.g., cooking and cleaning), leisure time, and transportation. Leisure-time physical activity (LPTA) may include recreational activities (e.g., walking for recreation), sports, and exercise (planned physical activity with the objective of improving or maintaining physical fitness). Active transportation can be defined as physical activity for transportation; across fields, it can also be found as “walking and cycling for transportation”, “non-motorized transport”, “human powered transport”, and “active transport” *(16).* The substantive empirical focus of this study is on utilitarian walking, which is one of the modes of active transport, which in turn is a component of physical activity (Figure 2-1).

Defining built environment is more permissive: e.g., from the field of anthropology, Lawrence *(17)* defines it as “any physical alteration of the natural environment, from heaths to cities, through construction by humans . . . spaces that are defined and bounded." From a planning perspective, the built environment, "comprises urban design, land use, and the transportation system" (*18)* and last, ecological models add that there can be both an objective built environment as well as a perceived built environment *(19).* The empirical focus of this study is on objective measures of land-use and the built environment that may be constructed using geographic databases of land use and transportation systems (Figure 2-1).

Last, health as defined by the President's Council on Physical Fitness and Sports *(20),* "is a state of being associated with freedom from disease and illness that also includes a positive component (wellness) that is associated with a quality of life and positive well-being." A Report of the Surgeon General *(15)* quotes Buchard et al. by defining health on page 22,

> a human condition with physical, social, and psychological dimensions, each characterized on a continuum with positive and negative poles. Positive health is associated with a capacity to enjoy life and to withstand challenges; it is not merely the absence of disease. Negative health is associated with morbidity and in the extreme with premature mortality.

According to the same report, "the ability to relate physical activity to health depends on accurate, precise, and reproducible measures." Often, the use of this term, applies to studies that relate physical activity with the prevention of disease (i.e., obesity, overall mortality, cardiovascular diseases, coronary heart disease, stroke, high blood pressure, cancer, osteoarthritis, osteoporosis, bone fractures, and mental health) or studies that measure physical fitness (e.g., submaximal aerobic capacity) and quality of life (e.g., physiological well-being). In this study (Figure 2-1), we focus on measures of subjective wellbeing as well as objective measures of health (Body Mass Index or BMI).

The next several sections present a summary of literature linking each pair of the three items (physical activity, health, and land use) identified in the framework. To limit the scope of the discussions, it focusses on studies that have addressed the specific measures of interest (such as utilitarian walking in the case of physical activity and BMI and subjective well-being in the case of health). However, as the studies are often not consistent in their choice of measures, additional studies if deemed relevant are included. Finally, it is also useful to note that there are several synthesis papers on the topics of interest. Therefore, the review first draw from these synthesis studies before adding results from recent individual empirical studies.

Active Transportation and Health

As a desirable public health policy, the study of active transportation and health often occurs from the epidemiology/ecological and transportation/planning fields. In both fields, research stems from two general methodologies: the objective measurement of levels of active transportation and the incidental effect of increase in walking or biking rates. Frank et al. *(21)* looked at the Atlanta metropolitan region and provide the latest review of this relationship. The framework employed in this study follows a linear sequence with lifestyles influencing residential location (land use), travel behavior and subsequently health. The study found users that preferred walking and lived on walkable neighborhoods were less likely to drive and more likely to walk. However walking or walking preference was not significant factor for obesity. Further, their study on 2,056 individuals found a significant fraction of the population as living outside of their preferred neighborhood type.

Within the discipline of transportation engineering, most of the literature has focused on walkability and work trips, often treating walking and biking as one category and generally finding limited evidence of lower body weights for adults [see systematic review *(22)*]. The empirical research focused on commute has found significant relationships at aggregate scales. For instance, higher state-level and city-level walk and bike commuting shares have been associated with lower diabetes and obesity rates *(23).* Higher country-level share of biking and walking (all-purpose trips) has also been correlated with lower obesity rates *(23).* In Belgium, a seven-day cross-sectional study found a negative association between BMI and minutes of walking for transportation *(24).* This relationship can also be positive, as one survey in Australia found women that participated in active transportation and lived in socioeconomically disadvantaged neighborhoods as more likely to be obese *(25)*. Similarly, another survey among 30 low-income mothers in England, found those who walked more than their peers had lower Self-Assessed Physical Health Score (SAPHS) values *(26).*

Clinical research, mainly focused on European cities, has found a positive relationship between active transportation (walking or biking) and improved fitness, and reduction in all-cause mortality, type 2 diabetes, and hypertension risks [see *(27)* for systematic review]. However, the definitions of meaningful walking vary across such studies with values starting in 10 min/day up to 30 min/day or 120 min/week. Some studies also employ Metabolic Equivalent Task values (MET) instead of actual walk durations. A smaller number of projects has focused on walking for commute: associating commute times and frequency with reduced health risks and reduced BMI in Denmark and Osaka *(28, 29).* Commuting to work, however, only accounts for approximately 16% of all person trips. The only known prospective cohort study of walking for transportation in general, could not find an association between walking >90 min/wk and reduced all-cause mortality risk *(30).*

In addition to understanding the effect of walking on health, it is also important to understand the effect of health on walking. Research on the effect of health on utilitarian walking has been scarce but greater emphasis has been placed on the impact of health on overall Leisure-Time Physical Activities (LTPA). A synthesis reviews on physical activity (which may include, among other activities, leisure walking but not utilitarian) by Sallis and Owen *(31)* found lack of association between overweight individuals and a reduction of activities. However, an updated review by Trost et al. *(32)* supports a negative association. While the transportation literature is abundant with studies on walking behavior and mode choice for travel, the travelers' health is generally not used as an explanatory factor in these studies.

Built Environment, Physical Activity, and Active Transportation

Though many agree on—or at least have studied—the relationship between built environment and physical activity in adults, a review of current literature shows diverse outcome variables and an apparent disparity on its measurement. The latest study by Harris et al. *(33)*, aimed at understanding the current structure of the physical activity and built environment field, maintains the association between active lifestyles and built environment, both for objective and perceived environment characteristics. In their paper, the authors set the stage by mapping relations and gaps and classifying studies as reviews, delivery, policy, theory/methodology, or discovery. A common issue within the literature and to those studying the effect of transportation systems, the study's classification scheme makes no distinction between recreational or leisure-time physical activities, work-related, and transportation-related, drawing their recommendations from the bulk number of studies focused on the first.

Another review of reviews on built environment, physical activity and obesity by Ding and Gebel *(34)* also notes:

- almost half of the reviews either combined adults with youth or did not specify target age groups
- reviews failing to select conceptually matched associations, or to examine behavior-specific correlates
- reviewers/studies should provide more information on environmental and physical activity measures, and stratify the summary of results based on measurement modes.

Besides the need for more rigorous studies, a review by Ferdinand et al. *(35)* identifies considerable gaps:

- minority populations and with their use of parks, school playgrounds, and active transportation
- rural populations
- facilitating PA for the elderly.

Heath et al. *(36)* raise additional issues on measurements, among these:

- What are the relationships between "objective" (e.g., derived by community and street-scale audits) and "perceived" (e.g., derived by telephone survey) neighborhood characteristics and does this relationship vary by perceived preference?
- In terms of an organizational structure, how should the built environment be conceptualized and what is the best way to measure or quantify components of the built environment (e.g., accessibility, aesthetics, safety, walkability)?
- What is the geographic scale(s) at which the neighborhood environment is most strongly correlated with physical activity?
- Does multivariate adjustment for potential confounding factors (e.g., age, income, gender) change the relationship between the built environment, policies, and physical activity? If so, what potential confounders are most important?
- Is it possible to use existing data to assess the impact of selection bias (e.g., stratifying data sets by income group)?

Furthermore, Ferdinand et al. *(35)* found that studies whose main outcome variable was active transport were less likely to observe a beneficial relationship. Surprisingly, the use of objectively measured PA data (as opposed to self-reported data) was associated with a reduced likelihood in finding a beneficial relationship between the built environment and PA or obesity rates.

A systematic review of physical activity and built-environment as measured by smart-growth principles *(37)* and by the Guide to Community Preventive Services *(36)* also found significant associations for walking or physical activity. The first review focused on 37 cross-sectional and longitudinal studies of adults and children in the U.S. (25), Australia (9), Europe (5), and Canada (3). The authors found a relationship between physical activity and principle seven (open space preservation), and walking (studies could include recreational, transport, or total) and principles one (range of housing choices), six (mixed land use), nine (development toward existing communities) and ten (compact building design). The second review grouped policies in three strategies (community-scale urban design and land use policies and practices to increase physical activity; street-scale urban design and land use policies to increase physical activity; and transportation and travel policies and practices), finding strong evidence in all but transport policies, as the authors could only found one time-series study focused on school mode share.

A recent review specific to green space and physical activity by Lachowycz and Jones *(38)* found mixed evidence for an association between access to green space and physical activity. A majority of studies (66%) found some evidence of a positive association, although only 40% found an association that appeared unambiguous.

A 2010 review by McCormack and Shiell *(39)* found preliminary results towards causality between built environment and physical activity for some of the effects reviewed above.

The discussion thus far has focused on the relationship between physical activities in general and built environment. Studies that narrow focus on the relationship between built environment and active transportation are mainly drawn from the transportation/planning and the public health fields. Pulleyblank Patrick et al. *(40)* provide insight on how both fields have approached this. The report, commissioned by USBGBC and partially funded by U.S. Environmental Protection Agency (EPA) and the Centers for Disease Control and Prevention (CDC), generally concludes a positive association between active transportation and the increase or presence of the following features: regional accessibility, population and employment density, land use mix, access to transit, neighborhood streetscape design that do not inhibit individuals from walking/cycling, higher street connectivity, and on-street parking. However, there is no consensus on what degree or threshold values these characteristics are effective, with some expectation of a collective—rather than isolated—effect. Additional features, such as access to recreational facilities, appear to influence recreational trips only and not utilitarian trips.

In the most recent review on walking and built environment, Saelens and Handy *(41)* find that built environment association found on general physical activity carries to studies focusing on walking, and similarly, “the specifics of the association are less clear.” The studies reviewed, almost all cross-sectional, were conducted in a small number of cities or neighborhoods, where the issue of self-selection remains.

While the issue of self-selection remains, Handy, Cao, and Mokhtarian *(42)* found built environment influence in eight neighborhoods in Northern California even after accounting for self-selection. The study found walking accessibility, physical activity options, safety concerns, socializing opportunities, and neighborhood attractiveness as characteristics that influenced walking trip behavior for 688 neighborhood movers. Another self-selection study, led by Frank et al. *(43)* in 2002 on the Atlanta, Georgia region, found greater walkability associated with greater share of people walking for non-discretionary purposes (work, school and shopping) and discretionary trips (e.g., entertainment, exercise), but that the absolute amount of walking remained extremely low for those who do not prefer a walkable environment. Walkability was defined as an index composed of commercial floor area ratio to the total land area devoted to commercial; land-use mix of residential, commercial, and office use; number of residential units per residential acre; and the number of intersections with more than three legs.

The most recent individual study by Knuiman *(44)* concludes that neighborhood connectivity, land-use mix, number of public transit stops, and the diversity of destination types accessible by walking were significantly related to neighborhood transport walking for the city of Perth, Australia. The 1,703-adults longitudinal study used weekly frequency of 15 min. utilitarian walk.

Built Environment and Health

This section will focus on studies that link built environment characteristics and more direct measures of individual health, mainly physical fitness outcomes. The latest state of research regarding the built environment-health relationship can be found in two epidemiologic syntheses *(45, 46).*

The first epidemiologic review, by Feng et al. *(46)*, looked at a total of 67 studies (61 cross-sectional), the majority coming from the United States (52 papers) and focused on adults (45 studies). In total, 48 studies found statistically significant associations (expected or unexpected direction). Their methodology considered studies related to one of three built environment subdomains: land use/transportation environment, food environment, and physical activity environment. The most common built environment characteristics among studies were density, diversity, design, connectivity, spatial access to facilities, walkability, and sprawl. Among outcome measures, a categorical measure (i.e., normal BMI vs. obese) was predominant, followed by continuous BMI values or Z-score. The authors took a closer look at 22 contextual-based and 15 individually-defined geographic buffer papers and concluded, “existing evidence does not identify a clear and strong role for built environmental risk factors with the possible exception of the county sprawl index and land use mix”. Studies focused on walkability and health outcomes found relations between walkability (be it defined as land use, as an index, or distance to amenities) and obesity in older Americans, by ethnicity, or gender at a neighborhood-scale *(21, 47-50)* and county-scale *(51).* Meanwhile, a comparison of walkability on Black and White neighborhoods by Doyle et al. (*52)* finds that BMI, socioeconomic status, and safety perception were stronger indicators than land-use measures.

The second, independent epidemiology review by Mackenbach et al. *(53),* focuses in adults, and updates the literature to 2013, while assessing the methodological quality of each paper. The authors reviewed the findings of a total of 92 papers, 62 coming from the US, exploring health through use of BMI, overweight or obesity outcomes. Their finding remains consistent with Feng’s discussion: strong associations could only be found for urban sprawl and land use mix in the US. At the same time, Gebel reveals a mismatch between perceived walkability and health benefits *(54).*

A review by Ferdinand et al *(35)* found the use of direct measures of body weight (e.g., BMI) as a measure of physical activity and the main outcome variable in a study was associated with a reduced likelihood in finding a beneficial relationship.

A similar systematic review of built-environment smart-growth principles and health, among other factors, found no strong associations between BMI and smart growth principles *(37),* though not all studies had matching principles. The review focused on 44 cross-sectional studies of adults and children, mostly in North America. The authors suggested the lack of longitudinal studies as a reason; or that smart growth's impact on body mass operates through mediation mechanisms not factored in these studies.

A 2001/2002 study on the Atlanta, Georgia region by Frank et al. *(43)* found greater walkability (by means of a walkability index) associated with lower odds of obesity for 2066 individuals. However, the absolute amount of walking remained low for those who did not prefer a walkable environment.

Recent studies published after the latest reviews are consistent with the impact of measures of BMI and the role of perceptions: Ullman et al. *(55)* found women with better neighborhood perception (safety, social) were associated with lower BMIs and in a longitudinal study of 975 adults in Los Angeles. Similarly, Barry *(56)* found that perceptions and level of physical activity had stronger obesity and diabetes effects on women than men in a cross-sectional study of 4,273 adults in Allegheny County, PA. Lastly, Glazier et al. *(57)* developed multiple land-use indices for a Toronto survey of 9,757 residents age 30 and older, and found that residential density and walkable destinations had strong associations with overweight status and diabetes.

Data Fusion

While there are several datasets available to study each pair of items from the physical activity – built environment – health framework, these do not support the development of a comprehensive model of the entire framework. As an alternative to conducting a new survey to collect additional data, we examine the applicability of data fusion methods to link travel and health surveys to generate the comprehensive dataset needed for the analysis. Specifically, the proposed approach adds health data from the ATUS dataset to the NHTS surveys at the individual level. An overview of the literature on data fusion is presented in this section.

Data linkage of two or more datasets can be conducted through manual inspection, deterministic procedures, or probabilistic matching. Manual inspection methods are not feasible with large datasets. Deterministic procedures require building a statistical model correlating the outcome variable with explanatory factors observed in both datasets and subsequently using the estimated model to make predictions of the outcome variable in the target dataset. In the context of this study, a regression model of BMI as a function of socio-economic-location variables can be built using the ATUS data and in turn applied to the NHTS data. Such methods are also fairly straight-forward.

The rest of this section focuses on probabilistic matching, also known as record linkage within the health/safety literature and several other names outside it. For a list of corresponding terms within various specialties see Gu et al. *(58)*. The process of linking records spans from the assumption that there is a potential pair-match within two sources. As Gomatam et al. *(59)* explain,

> suppose file A has n_a records and file B has n_b records, then the file $A \times B$ contains $n_a \times n_b$ record pairs. Each of the n_b records in file B is a potential match for each of the n_a records in file A. Thus there are $n_a \times n_b$ record pairs whose match=non-match status is to be determined.

Record-pair comparison will lead two three outcomes: (a) match, (b) possible match, and (c) non-match. The degree of separation between datasets can be an indication of the level of difficulty of the linkage and amount of Type I and II errors *(58)*.

The standard algorithm requires a method that minimizes the probability of possible matches, as opposed to true matches and non-matches. This is done by estimating the ratio between conditioned probabilities of observed known identifying fields given that the record pair is a true match and true not-match. In many instances, the expectation-maximization algorithm (EM), as described by Jaro *(59),* is used to estimate the weights or threshold values of this association.

An EM linkage requires initial values for match and unmatched probabilities thresholds, selection of blocking attributes, the threshold of the weights for possible matches, and the type of comparison function. The inclusion of a blocking attribute (e.g., sex) will yield comparisons of pairs within each matching attribute (e.g., for sex as blocking attribute, the comparison of female records exclusively with other female records, and male records with other male records). The decision component, often score-based, relies either on frequency-based distribution *(60)* of attributes (e.g., there are fewer individuals living in rural areas, therefore a match of urban carries less weight than a match of rural) or uniform, average of all matching weights.

Common performance measures are the number of record pairs linked correctly and incorrectly, and the number of pairs unlinked correctly and incorrectly. But, as Baker *(61)* notes, the accuracy of the process is a function of the size of the donor dataset and,—counterintuitively—the use of too many variables may not lead to better linkage.

The transportation literature offers examples of record linkage, such as a study by Kukasabe and Asakura *(62)* linking smart card transit data stops and personal trip surveys, police and hospital road crash records [Amorim et al. *(63),* Rosman *(64)*], work fatigue and crashes [Williamson and Boufous *(65)*], and state and federal motor carrier safety databases *(66).* Pawlak et al. *(67)* used separate datasets on digital lifestyle (ICT use) and physical mobility to match records as an alternative where suitable data cannot be obtained in a single dataset. Similarly, Kressner and Garrow *(68)* also found third-party data from targeted marketing to be interchangeable and useful to incorporating lifestyles variables to transportation survey, in particular with age, gender, income, and presence of children.

Summary

In this chapter, we've shown how the previous relationships have been explored among walking, health, and built environment, with a focus on utilitarian walking. Evidence supports the causal link of walking and better weight or fewer diseases, however, multiple studies rely on aggregate measures of total leisure-time physical activity or on commute-only behavior—the former too broad and the later too limited to employed adults and mandatory activities. The built environment and walking correlates are better known. While there is an expected confounding effect between desirable built environment features, these studies often rely on samples geographically limited to a neighborhood or a particular city. For built environment and health, there is again an agreement, especially when measuring county sprawl index and land use mix, of correlations with obesity (or walking), as some review construct walking behavior as an implicit health benefit without a direct measure of fitness or disease. Similarly, we've shown how there are few studies that attempt to incorporate the three inter-dependencies, much less using a large national sample. Often, the scarcity of rich datasets will involve extensive data fusion techniques. The use of probabilistic record linkage has been proved to be effective with datasets of the same population for various purposes, including crash records and health and Information and Communication Technologies (ICT).

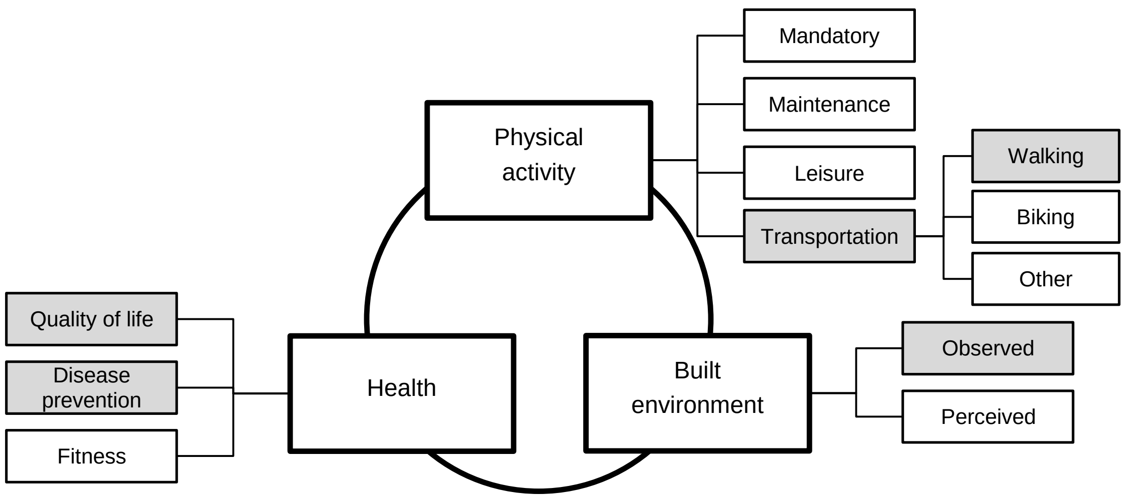

Figure 2-1. Utilitarian walking, health and built environment relationships within the context of this study.

CHAPTER 3
MAKING THE CONNECTION BETWEEN UTILITARIAN WALKING AND HEALTH MEASURES USING THE AMERICAN TIME USE SURVEYS

The object of this chapter is to use a large cross-sectional data from the U.S. to explore the relationship between healthy individuals and their propensity to walk as well as the relation between walking for transportation and a person's health (reported or perceived) through recommended guidelines. The analysis was focused on adults.

Data

This analysis uses data from the 2006, 2007, and 2008 American Time Use Surveys (ATUS) and the corresponding Eating & Health (EH) Modules. ATUS collected detailed socio-economic, demographic, and one-day activity-travel information for a large sample of persons. The survey is limited to one randomly selected household member (individuals age 15 or older living on non-institutionalized facilities) per household and the survey samples were drawn from the Current Population Surveys (CPS) of the same years. The EH Module of the ATUS, administered between 2006-08 collected additional eating, meal preparation, and health information for a random selection of household members participating in the ATUS. To our knowledge, only one other study *(69)* has looked at one year of ATUS/EH data to examine the association between physical activity and BMI, noting lesser BMIs with the presence of active transportation, defined as one minute or more of walk or bike mode.

Overall, both time-use and health data are available for about 36,000 individuals. The socio-economic characteristics of the sample are presented in Table 3-1. Since the ATUS surveys collected weekend time use data for 50% of the samples (and weekday data for the rest), the results are shown separately for the two cases (i.e., weekdays and weekend days). Both the weekday and weekend sample have similar attributes.

The sample is largely Caucasian and there are more females than males (55% versus 45%). The survey respondents had some secondary-degree education and were significantly likely to be employed full-time. The most common household comprises two adults and their children, with a combined income between $35,000-$100,000, and living in an owned household within a Metropolitan Statistical Area according to 2000 Census (have at least one urbanized area of 50,000 or more population, plus adjacent territory that has a high degree of social and economic integration with the core). Another motivation for separating out weekday and weekend samples is that the purpose of the walk trips are significantly different across these two cases (Table 3-2). Work-related trips lead the walk trip purpose on weekdays (18.4%), while social/leisure/recreational activities (27.2%) characterize weekend walk trips.

In order to correlate health measures to walking patterns, it is first required to develop a working definition of a “walker”. Among all individuals surveyed, about 12% walked at some point during the survey day (for any amount of duration). However, many instances of such walking were for rather short durations (such as less than 10 minutes) and these may be (1) incidental trips not reflective of general behavior and/or (2) not be significant enough from the standpoint of health. To be sure, as already indicated CDC recommends at least 10 minutes at a time of moderate-intensity aerobic activity such as brisk walking for a total of 150 minutes every week and the U.S. Surgeon General recommends 30 minutes of moderate physical activities on most days of the week. The ATUS is a one-day survey and therefore longer-term (such as weekly) walking patterns of the respondents are unknown. The number of respondents in the sample who walked continuously for 30 minutes on the survey day were too few.

Therefore, a person is defined to have "walked" on a day if he/she has one or more instances of continuous travel-related walking for 10 or more minutes. With this definition, 5.7% of the persons were observed to have walked on a weekday and 4.7% on a weekend day. As a second definition, the threshold for continuous walking was increased to 15 minutes, the average time it takes a person to walk a mile, yielding 3.4% of the sample to be walkers on weekdays and 3% to be walkers on weekends. As opposed to leisure walking, most utilitarian walking was undertaken on weekdays.

There are two measures of health available from the well-being module of the ATUS. These are the Body Mass Index (BMI) and a Self-Assessed Physical Health Score (SAPHS). BMI is calculated from the self-reported weight ("How much do you weigh without shoes?") and height ("How tall are you without shoes?"). The SAPHS was obtained as the response to the question: "In general, would you say that your physical health was excellent, very good, good, fair, or poor?" Table 3-3 presents summary descriptive statistics on these two heath measures.

Although BMI is intrinsically a continuous variable (and is treated as such in the models later on), it has been converted into a categorical variable for cross-tabulation purposes. Following the National Institutes of Health classification scheme *(10),* a person is classified into one of eight categories based on their BMI values: Severely Underweight (SU, BMI <14.9), Very Underweight (VU, BMI =15-15.9), Underweight (U, BMI =16-18.4), Normal (N, BMI =18.5-24.9), Overweight (OW, BMI =25-29.9), Obese I (Ob I, BMI =30-34.9), Obese II or Very Obese (Ob II, BMI = 35-39.9), or Obese III or Severely Overweight (Ob III, BMI >40). Due to the very small shares of underweight individuals, categories SU, VU and U are combined into an aggregate category with an overall share of 1.7% of the sample. Note that 37% of the sample have a normal weight and 35% are overweight (see last column of Table 3-3). The rest of the sample is obese with 3.5% of the sample falling under the Obese III category.

Based on the SAPHS, 19% indicated their physical health to be excellent and 34.4% to be very good (see last row of Table 3-3). 4.4% of the respondents indicated poor health and 12.3% indicated fair health. Table 3-3 also presents a cross-tabulation of these two health measures (the top half of the table presents the BMI distribution within each SAPHS level and the bottom half presents the SAPHS distribution within each BMI level). These cross-tabulations show a significant correlation between normal BMI and high ratings on the SAPHS score. Specifically, 54% individuals with excellent SAPHS rating are normal weight (40% for individuals with very good SAPHS rating) while 65% of the normal weight persons reported very good or excellent on the SAPHS.

Next, cross tabulation of health measures (BMI and SAPHS) are presented against walking patterns (defined using 10 and 15 minute walk time thresholds). The intent of this exercise was to present levels of correlations between walking and health at an aggregate level. Table 3-4 presents the distribution of walking patterns by health measures. In general, individuals with a better (healthier) BMI are more likely to be “walkers” (irrespective of the definition of walkers and for both weekdays and weekend days). For example 7.1% of the normal weight persons walked 10 minutes or more on weekdays compared to 2.9% of persons who were obese III. Those with a better SAPHS rating were also found to be more likely to be walkers, although this result is more pronounced in the case of the 10-minute definition of walking. Table 3-5 presents the distribution of health by walking patterns. In general, walkers are more likely (about

45%) to be “normal weight” than non-walkers (about 35%) and this holds irrespective of the definition of walkers and for both weekdays and weekend days. However, the distributions of SAPHS ratings across the walkers and non-walkers do not appear to be pronouncedly different.hese cross tabulations are consistent with the general expectation of a positive correlation between peoples’ walking patterns and their health.

Effect of Walking on Health

This section presents an analysis of the effects of walking on health. As already indicated, two health measures are available. BMI is modeled using a linear regression model and SAPHS is modeled using ordered-probit (0 = “poor”, 1=”fair”, 2 = “good”, 3 = “very good”, and 4 = “excellent”) recognizing the nature of these data. For each health measure, two models are developed capturing the effect of weekdays and weekend walking. Further, given that two definitions of “walking” have been adopted (based on minimum duration thresholds of 10 and 15 minutes), all the models (for the two health measures and the two days of the week) are estimated using each of these definitions of walking leading to a total of eight models. The results for the effect of 10-minute walking on health are presented in Table 3-6 and those for the effect of 15-minute walking are presented in Table 3-7. In each of these tables, the first two major columns present the regression-model results for BMI and the last two major columns present the results of the ordered response model for SAPHS.

The ATUS provides information on whether the respondent walked on the survey day. It would not be entirely appropriate to use this directly as an explanatory variable in the models for health. This is because the short-term (daily) decision of walking on a single day may not be a truly “exogenous” predictor of health outcome measures that are generally descriptors of longer-term conditions. Further, walk episodes are perhaps not undertaken daily. Therefore, an instrumental-variables approach is adopted with a "predicted probability of walking on a day" being used as the instrument. Recognizing that leisure walking/exercise could play a substitute or complementary role, a "first-stage" MNL model was developed to determine the probability of utilitarian walking only, exercising only, and walking and exercising, as a function of the socio-economic characteristics of the individual and his/her household, residential location characteristics, and other characteristics of time-use on that day. These models were developed separately for weekdays and weekend days and using each of the two time definitions of "walking" (four models overall). These models are then used to calculate the predicted probability of walking for each person in the analysis sample. The predicted probabilities of walking are used as the explanatory variables in the "second stage" models for health. Such a two-stage instrumental variable approach is well recognized in the field of statistics as a means to address endogenous explanatory factors and has been employed in transportation research [see *(40)* for an application of such an approach in the context of modeling interdependencies among virtual and real activity participation choices].

The results from the "second stage" models for health outcomes are presented in Tables 3-6 and 3-7. All effects reported are statistically significant at 95% confidence or higher. The predicted probability of weekday/weekend walking (10 minutes or more) has a negative impact on BMI and a positive impact on SAPHS (Table 3-6). This indicates that individuals who are more likely to walk (on either weekday or weekend day) have lower BMI and "feel" healthier (higher SAPHS). On examining the effect of predicted probability of walking 15 minutes or more, the effect on BMI remains statistically significant and of larger magnitude; positive and significant on SAPHS (Table 3-7). Overall, walkers are estimated to “feel” healthier and the effect of walking on BMI is significant even after controlling for exercise.

It is useful to highlight that the models capture the marginal effect of walking after controlling for the effect of a variety of other factors on health. Specifically, health is affected by age, gender, ethnicity, nutrition, exercise, education, and employment status of the person. Further low-income persons are likely to have higher BMI and report lower SAPHS. Similarly, those in metropolitan areas are likely to have lower BMI and report higher SAPHS. Variations in health are also found across the different census regions of the country.

Effect of Health on Walking

This section presents an analysis of the effects of health on walking. Binary logit models are estimating considering the dichotomous outcome variable (walked or not). Separate models are estimated to capture the differential impacts of health on weekday and weekend walking. Further, separate models are also estimated to capture the effects of each of the health outcome variables (BMI and SAPHS). Finally, given that two definitions of “walking” have been adopted (based on minimum duration thresholds of 10 and 15 minutes), all the models (for the effect of each of the two health measures and the two days of the week) are estimated using each of these definitions of walking leading to a total of eight models.

The results for the effects of health on 10-minute walking are presented in Table 3-8 and those for the effect of health on 15-minute walking are presented in Table 3-9. In each of these tables, the first two major columns present the regression-model results for BMI and the last two major columns present the results of the ordered response model for SAPHS. All effects reported are statistically significant at 95% confidence or higher. The Nagelkerke R^2, a goodness of fit value reported on multinomial logistic regression as the normalized likelihood ration, show model performance. The models indicate a negative relationship between BMI and walking for both weekdays and weekend days; i.e., persons with higher BMI are less likely to walk. This result holds irrespective of the definition of walking used. However, the relationship between SAPHS and walking holds only for weekend days. Specifically, those who feel healthier (higher SAPHS) are more likely to walk on weekends but the effect on weekday walking is statistically insignificant. This difference could be because weekday walking is more likely to be for mandatory purposes whereas weekend walking is leisure-oriented.

Once again, the models capture the marginal effect of health on walking after controlling for the effect of several other factors. The likelihood of walking is affected by age, gender, ethnicity, education, and employment status of the person and the magnitude of these impacts vary by day of the week. The daily time-use pattern of the individual also affects (negatively) the likelihood of undertaking utilitarian walking. This is reasonable as the need to invest greater amounts of time in activity participation can motivate people to choose faster modes to get to their destinations given the overall daily time-budget constraint. Variations in walking behavior are also found across the different census regions of the country and across the different seasons of the year.

Findings

The relationship between utilitarian walking and health was investigated using a cross-sectional analysis of a large-sample US time-use and health dataset. In particular, a series of models was developed to analyze (a) the effect of walking on health, and (b) the effect of health on walking. Two thresholds (10-minute bouts and 15-minute bouts) were used in defining walking patterns as part of two health metrics (reported BMI and Self-Assessed Physical Health Score). The impacts of/on weekday and weekend walking were examined separately.

The study indicates that individuals who are more likely to walk (on either weekday or weekend day) have lower BMI and “feel” healthier (higher SAPHS). However, on examining the effect of predicted probability of walking 15 minutes or more, the effect is statistically insignificant for BMI but positive and significant for SAPHS. The study indicates a negative relationship between BMI and walking for both weekdays and weekends; i.e., persons with higher BMI are less likely to walk. This result holds irrespective of the definition of walking used. However, those who feel healthier (higher SAPHS) are more likely to walk on weekends, but the effect on weekday walking is statistically insignificant. This difference could be because weekday walking is more likely to be for mandatory purposes, whereas weekend walking is leisure-oriented.

In general, the results for the positive effects of walking on health are encouraging, particularly considering the large and diverse sample and the inclusion of a large number of control variables. At the same time, the results on the negative effects of BMI on walking highlight the need for encouraging people who are currently not overweight to walk, as the onset of obesity can have a detrimental impact on walking. Overall, this chapter demonstrated the benefits of utilitarian walking on health (decreasing BMI values and fostering better health perceptions) and highlights the need to control obesity, as an increase in this population can reduce expected walking rates.

Table 3-1. Summary characteristics of the analysis sample.

Classification	Description	Weekday	Weekend
Sample Size	Randomly-selected member per household	17,607	17,992
Sex	Male	45.3%	44.6%
	Female	54.7%	55.4%
Age	Person is under 20	7.4%	7.6%
	Person is between 20 and 39	30.9%	31.7%
	Person is between 40 and 59	37.6%	37.4%
	Person 60 or older	24.2%	23.3%
Race	Identified as Caucasian	69.8%	68.9%
	Identified as Black/African America	13.0%	12.8%
	Identified as Hispanic	12.3%	13.4%
	Identified as Asian	3.0%	2.8%
	Other ethnicities	1.9%	2.1%
Exercise	Some exercise/active leisure (non-transportation)	14.3%	13.6%
Nutrition	Food stamps recipient household	6.8%	7.0%
Education	Did not finish high-school	16.6%	16.9%
	High-school/GED degree	26.5%	26.6%
	Completed or some advanced-degree education	56.9%	56.5%
Employment	Person enrolled in school (part- or full-time)	5.5%	6.0%
	Person identified as unemployed	17.4%	17.5%
	Person identified as part-time worker	11.8%	11.1%
	Person identified as full-time worker	53.8%	54.8%
	Person identified as retired from work	17.1%	16.8%
Household	Lives alone	25.6%	25.3%
	Couple without children or relatives household	20.9%	20.0%
	Household with minor(s) and two parents	35.4%	36.0%
	Household with minor(s) and one adult	5.5%	5.6%
	Other household types: roommates, siblings, etc.	12.5%	13.2%
Income	Combined household income less than $35k	31.0%	31.8%
	Combined household income between $35k-$100k	40.3%	39.9%
	Combined household income greater than $100k	14.7%	15.5%
HH Tenure	Residence owned by a household member	75.2%	74.6%
Region	Household located in Northeast census region	16.7%	17.2%
	Household located in Midwest census region	24.9%	24.7%
	Household located in South census region	36.5%	36.2%
	Household located in West census region	21.8%	21.9%
Metro	Household located within a Metropolitan Area	81.5%	82.0%
	Household located outside a Metropolitan Area	18.5%	18.0%
	Missing household location	0.7%	0.8%
Season	Survey day between Nov 01 - Feb 28	33.1%	33.8%
	Survey day between Mar 01 - April 30	17.3%	17.9%
	Survey day between May 01 - Jun 14	12.7%	11.5%
	Survey day between Jun 15 - Aug 14	16.5%	16.4%
	Survey day between Aug 15 - Sep 14	7.7%	7.1%
	Survey day between Sep 15 - Oct 31	12.7%	13.4%

Table 3-2. Purpose of weekday- and weekend- walk trips.

Walk trip purpose	Weekday	Walk trip purpose	Weekend
Work	18.4%	Socializing, relaxing and leisure	27.2%
Socializing, relaxing and leisure	18.2%	Consumer purchases	22.5%
Consumer purchases	14.5%	Eating and drinking	10.9%
Caring for household members	11.7%	Work	6.1%
Eating and drinking	10.2%	Sports, exercise & recreation	5.9%
School	5.7%	Household activities	5.7%
Household activities	4.1%	Caring for household members	4.6%
Sports, exercise & recreation	4.1%	Caring for non-household members	3.9%
Caring for non-household members	3.3%	Religious and spiritual activities	3.4%
Professional & personal care services	1.9%	Volunteering activities	1.5%
Volunteering activities	1.4%	Professional & personal care services	1.4%
Other discretionary activities	1.0%	Other discretionary activities	1.2%
Religious and spiritual activities	0.9%	Personal care	1.0%
Personal care	0.8%	Household services	0.6%
Household services	0.3%	School	0.5%

Table 3-3. Summary Statistics on health measures.

	Category	Poor	Fair	Good	Very good	Excellent	Total
Within SAPHS (%)	SU-U	2.2	1.6	1.5	1.6	2.2	1.7
	N	25.8	24.0	29.3	39.8	54.7	37.0
	OW	29.8	31.8	35.4	37.7	32.6	35.0
	Ob I	20.3	22.3	21.9	15.1	7.9	16.9
	Ob II	11.0	11.2	7.7	4.2	1.8	5.9
	Ob III	10.9	9.0	4.2	1.6	.7	3.5
Within BMI (%)	SU-U	5.7	11.7	26.0	31.8	24.8	
	N	3.1	8.0	23.5	37.0	28.3	
	OW	3.8	11.2	30.0	37.1	17.8	
	Ob I	5.4	16.3	38.6	30.8	9.0	
	Ob II	8.2	23.3	38.3	24.5	5.7	
	Ob III	13.7	31.4	35.3	15.9	3.7	
	Total	4.4	12.3	29.7	34.4	19.1	

Table 3-4. Distribution of walking behavior by health measures.

		Weekday				Weekend			
Type	Category	Nonwalkers (%)		Walkers (%)		Nonwalkers (%)		Walkers (%)	
	Threshold	10 min	15 min	10 min	15 min	10 min	15 min	10 min	15 min
BMI	SU-U	91.2	95.6	8.8	4.4	95.1	95.9	4.9	4.1
	N	92.9	95.8	7.1	4.2	94.2	96.4	5.8	3.6
	OW	94.8	96.9	5.2	3.1	95.5	97.0	4.5	3.0
	Ob I	95.9	97.3	4.1	2.7	96.2	97.8	3.8	2.2
	Ob II	95.6	97.4	4.4	2.6	97.0	98.4	3.0	1.6
	Ob III	97.1	98.5	2.9	1.5	96.5	97.7	3.5	2.3
SAPHS	Poor	95.5	96.9	4.5	3.1	97.3	98.2	2.7	1.8
	Fair	94.7	96.5	5.3	3.5	95.3	96.6	4.7	3.4
	Good	94.0	96.3	6.0	3.7	95.2	97.0	4.8	3.0
	Very good	94.6	96.9	5.4	3.2	95.4	97.1	4.6	2.9
	Excellent	93.8	96.7	6.2	3.3	94.7	96.7	5.3	3.3

Table 3-5. Distribution of health measures by walking behavior.

		Weekday				Weekend			
Type	Category	Nonwalkers (%)		Walkers (%)		Nonwalkers (%)		Walkers (%)	
	Threshold	10 min	15 min	10 min	15 min	10 min	15 min	10 min	15 min
BMI	SU-U	1.5	1.5	2.4	2.0	1.9	1.9	2.	2.6
	N	36.3	36.6	46.4	45.6	36.6	36.8	45.1	44.6
	OW	35.3	35.2	32.4	32.7	34.9	34.8	33.1	34.5
	Ob I	17.3	17.1	12.4	13.6	16.9	16.8	13.4	12.4
	Ob II	6.1	6.0	4.6	4.5	6.0	6.0	3.8	3.1
	Ob III	3.5	3.5	1.7	1.5	3.8	3.7	2.7	2.8
SAPHS	Poor	4.4	4.4	3.5	4.0	4.6	4.6	2.6	2.8
	Fair	12.4	12.3	11.6	13.0	12.3	12.2	12.3	13.8
	Good	29.6	29.6	31.3	32.3	29.6	29.7	30.3	29.3
	Very good	34.9	34.8	33.0	32.5	34.2	34.2	33.1	32.8
	Excellent	18.7	18.9	20.5	18.2	19.3	19.3	21.7	21.2

Table 3-6. Effect of walking (10-minute threshold) on health.

	BMI [Weekday]		BMI [Weekend]		SAPHS [Weekday]		SAPHS [Weekend]	
	B	S.E.	B	S.E.	B	S.E.	B	S.E.
Walk probability	-7.42	1.63	-8.16	2.13	1.56	0.302	1.00	0.365
Walk + Exercise probability			-8.97	2.90			2.05	0.877
Exercise probability					0.895	0.130	0.559	0.146
Female	-0.883	0.084	-1.09	0.088	0.125	0.018	0.077	0.018
Age 20-39	3.10	0.220	3.06	0.210	-0.431	0.047	-0.271	0.047
Age 40-59	3.84	0.224	4.20	0.214	-0.641	0.046	-0.506	0.048
Age 60+	3.76	0.254	3.86	0.254	-0.743	0.053	-0.660	0.056
Black	2.01	0.150	2.01	0.153	-0.214	0.027	-0.210	0.027
Asian	-2.12	0.217	-2.54	0.198	-0.286	0.049	-0.106	0.049
Hispanic	0.768	0.152	0.922	0.152	-0.237	0.029	-0.253	0.027
Income <$35k	0.561	0.109	0.621	0.109	-0.187	0.021	-0.210	0.021
HS/GED education					0.218	0.030		
College	-0.542	0.098	-0.528	0.099	0.479	0.030	0.346	0.019
Student	-0.513	0.224	-0.611	0.210			0.149	0.044
Retired	-0.699	0.169	-0.788	0.180			-0.167	0.040
Part-time job	-0.746	0.131	-0.502	0.142	0.303	0.038	0.252	0.039
Full-time job					0.360	0.035	0.314	0.035
Unemployed					-0.158	0.041	0.190	0.028
Own house			-0.461	0.152	0.138	0.027	0.052	0.021
Nuclear household					0.107	0.019	-0.087	0.027
Other multi-adult household	0.606	0.140	0.385	0.134				
Food stamps recipient	1.58	0.229	1.35	0.219	-0.380	0.037	-0.327	0.036
Constant	24.6	0.227	25.0	0.281	1.72	0.072	1.82	0.073
tau1	n/a		n/a		0.878	0.070	0.863	0.072
tau2	n/a		n/a		1.88	0.070	1.86	0.072
tau3	n/a		n/a		2.93	0.071	2.90	0.072
Goodness of fit								
McFadden's R^2	0.0743		0.0837		0.0655		0.0671	

Table 3-7. Effect of walking (15-minute threshold) on health.

	BMI [weekday]		BMI [weekend]		SAPHS [weekday]		SAPHS [weekend]	
Parameter	B	S.E.	B	S.E.	B	S.E.	B	S.E.
Walk probability	-11.979	2.724	-7.33	3.29	2.44	0.494	2.10	0.555
Walk + exercise			-13.27	5.55			2.89	1.031
Exercise					0.840	0.127	0.558	0.125
Female	-0.891	0.084	-1.10	0.09	0.127	0.018	0.078	0.018
Age 20-39	3.543	0.173	3.31	0.20	-0.444	0.047	-0.398	0.053
Age 40-59	4.333	0.168	4.43	0.20	-0.656	0.046	-0.628	0.053
Age 60+	4.301	0.204	4.10	0.24	-0.765	0.053	-0.771	0.060
Black	1.961	0.148	1.99	0.15	-0.216	0.027	-0.204	0.027
Asian	-2.246	0.208	-2.52	0.20	-0.267	0.048	-0.112	0.049
Hispanic	0.743	0.149	0.916	0.16	-0.239	0.029	-0.242	0.028
Income <$35k	0.476	0.112	0.698	0.11	-0.196	0.021	-0.204	0.021
HS/GED education					0.231	0.030	0.235	0.030
College	-0.618	0.097	-0.567	0.10	0.495	0.031	0.524	0.030
Student			-0.473	0.21	0.163	0.046	0.193	0.026
Retired	-0.734	0.170	-0.811	0.18				
Part-time job	-0.712	0.130	-0.479	0.14	0.283	0.039	0.235	0.039
Full-time job					0.351	0.035	0.306	0.035
Unemployed					-0.178	0.042	-0.169	0.040
Own house	-0.442	0.149			0.143	0.028		
Nuclear household					0.109	0.019	0.060	0.021
Other multi-adult household	0.594	0.139	0.264	0.13			-0.093	0.027
Food stamps recipient	1.524	0.228	1.33	0.22	-0.388	0.037	-0.317	0.036
Constant	24.575	0.235	24.30	0.20	1.737	0.070	1.762	0.072
tau1	n/a		n/a		0.878	0.069	0.869	0.071
tau2	n/a		n/a		1.876	0.069	1.872	0.070
tau3	n/a		n/a		2.93	0.070	2.91	0.071
Goodness of fit								
McFadden's R^2	0.0742		0.083		0.0656		0.0685	

Table 3-8. Effect of health on walking (10-minute threshold).

	Weekday		Weekend		Weekday		Weekend	
	B	S.E.	B	S.E.	B	S.E.	B	S.E.
BMI	-0.039	0.007	-0.040	0.007	n/a		n/a	
SAPHS	n/a		n/a		n.s.		0.137	0.037
Female			-0.248	0.078			-0.212	0.078
Age 20-39	-0.654	0.132			-0.770	0.130		
Age 40-59	-0.590	0.137			-0.743	0.134		
Age >=60	-0.576	0.173			-0.723	0.171		
Asian	0.608	0.159			0.697	0.158		
Black	0.669	0.098	0.610	0.105	0.603	0.097	0.561	0.104
Hispanic	0.672	0.095	0.565	0.099	0.639	0.094	0.549	0.099
HS/GED			-0.541	0.110			-0.583	0.110
College education			-0.466	0.100			-0.523	0.100
Part-time work	-0.569	0.122			-0.549	0.122		
Full-time work	-0.627	0.105			-0.640	0.105	-0.241	0.083
Couple household	-0.409	0.108			-0.425	0.108		
Single-parent household	-0.391	0.154	-0.347	0.161	-0.409	0.154	-0.363	0.161
Nuclear household	-0.554	0.105	-0.282	0.085	-0.570	0.105	-0.293	0.085
Other multi-adult hh	-0.427	0.121			-0.460	0.120		
Retired	-0.455	0.156			-0.452	0.155		
At-home maintenance	-0.170	0.022	-0.136	0.021	-0.167	0.022	-0.134	0.021
Non-home maintenance	-0.117	0.013	-0.083	0.017	-0.118	0.013	-0.082	0.017
At-home discretionary			0.143	0.058			0.141	0.058
At-home discretionary^2	-0.005	<0.001	-0.009	0.002	-0.005	<0.001	-0.009	0.002
Non-home discretionary^2	-0.007	0.002	-0.005	0.002	-0.007	0.002	-0.005	0.002
Northeast US Census	0.527	0.091			0.527	0.091	0.611	0.098
Midwest US Census	-0.299	0.104			-0.311	0.104	-0.276	0.113
Southern US Census	-0.681	0.098			-0.695	0.098	-0.626	0.106
HH tenure	-0.850	0.078	-1.035	0.079	-0.835	0.077	-1.051	0.079
Metropolitan area	0.640	0.123	0.658	0.133	0.667	0.123	0.666	0.133
Jun-Aug Season			0.294	0.091			0.279	0.091
Constant	1.29	0.294	-0.848	0.541	0.408	0.251	-2.206	0.516
Goodness of Fit								
Nagelkerke R^2	0.143		0.137		0.137		0.0134	

Note: n/a (not applicable)

Table 3-9. Effect of health on walking (15-minute threshold).

	Weekday		Weekend		Weekday		Weekend	
	B	S.E.	B	S.E.	B	S.E.	B	S.E.
BMI	-0.04	0.008	-0.045	0.009	n/a		n/a	
SAPHS	n/a		n/a		n.s.		0.099	0.044
Female			-0.275	0.093			-0.245	0.093
Age 15-19	0.518	0.15			0.672	0.146		
Age >=60					0.503	0.215		
Asian					0.631	0.12		
Black	0.674	0.12	0.552	0.131	0.661	0.117	0.469	0.132
Hispanic	0.661	0.116	0.727	0.115			0.655	0.117
Dropped school			0.399	0.114			0.468	0.114
Student					0.622	0.111		
Unemployed	0.631	0.111	0.361	0.112			0.423	0.113
Part-time work					0.412	0.103		
Single-person household	0.392	0.103						
Nuclear household							-0.206	0.105
At-home mandatory	-0.209	0.032	-0.172	0.036			-0.156	0.036
Non-home mandatory	-0.218	0.019	-0.172	0.022	-0.203	0.032	-0.168	0.022
At-home maintenance	-0.227	0.029	-0.189	0.026	-0.217	0.019	-0.187	0.026
Non-home discretionary	-0.148	0.025			-0.221	0.029		
At-home discretionary^2	-0.008	0.001	-0.006	0.001	-0.146	0.025	-0.006	0.001
Non-home discretionary^2			-0.012	0.002	-0.008	0.001	-0.012	0.002
Northeast US Census	0.608	0.116	0.745	0.103			0.615	0.118
Midwest US Census	-0.29	0.137			0.617	0.116	-0.284	0.142
Southern US Census	-0.549	0.125	-0.51	0.118	-0.287	0.137	-0.64	0.131
HH tenure	-0.989	0.093	-0.982	0.095	-0.549	0.126	-0.934	0.098
Metropolitan area	0.598	0.157	0.733	0.171	-0.975	0.093	0.726	0.172
Nov-Mar Season	-0.239	0.095			0.61	0.157		
Mar-May Season					-0.239	0.095		
Constant	0.64	0.401	-0.235	0.363	-0.505	0.329	-1.576	0.316
Goodness of Fit								
Nagelkerke R^2	0.149		0.134		0.144		0.129	

Note: n/a (not applicable)

CHAPTER 4
IMPUTING HEALTH MEASURES BY DATA FUSION

Introduction

The previous chapter presented models that capture the relationships between utilitarian walking and health measures. However, these models did not include any controls for land-use variables (i.e., characteristics of the respondents' residential location). This was primarily because the ATUS data used in the analysis do not provide detailed spatial information about the residential location of the survey respondents. Other surveys of daily travel such as the National Household Travel Surveys (NHTS) do provide detailed (latitude and longitude) residential location information. However, these surveys do not collect information on the health of the survey respondents. Therefore, developing models that examine walking, health, and land-use either require new surveys that collect all these data or the use of data-fusion approaches that systematically link individual records from ATUS and NHTS surveys to generate a comprehensive dataset. In this study, we explore the latter approach of data fusion. Insights from this study can help provide insights into the need for additional surveys in the future to comprehensively study the land-use – transportation – health relationships.

In this chapter, two methods for performing disaggregate data fusion were examined. The first method (regression-based matching) involves first developing a regression model of BMI using socio-economic predictor variables using a "Donor" dataset and, subsequently, using this model to predict the BMI for the respondents in a "Receiver" dataset. The second method (probabilistic matching) involves directly comparing the records of the "Receiver" dataset with records of the "Donor" dataset to identify the record from the "Donor" that best matches the "Receiver" on a set of predefined socio-economic attributes. The BMI of the matched “Donor” is then set as the BMI of the “Receiver”.

In order to evaluate these methods, it is necessary to know the BMI of the respondents in the “Receiver” dataset so that the matched BMI may be compared to the true BMI. Therefore, this chapter splits the ATUS data into a “Receiver” sample (10,000 cases) and a “Donor” sample (26,000 cases) for comparing the regression-based- and the probabilistic- matching methods. The probabilistic-matching method is shown to be better and is subsequently used (next chapter) to link the ATUS and NHTS records.

Application of the Two Methods

A comparison of Donor and Receiver datasets and socioeconomic indicators is shown on Table 4-1. Both ATUS samples follow similar shares, including a female majority (55%), as well as some higher education (57%), and full-time employed (55%). The datasets provide a fair representation of non-civilian U.S. adults, including participation of all family types and income levels.

The regression-based matching method is conceptually straightforward and requires a regression model developed on the Donor dataset to be applied to the Receiver dataset. Table 4-2 presents the results of such OLS regression as a function of matching attributes significant at 95% confidence or higher.

This probabilistic data fusion employed Link Plus, a probabilistic record linkage program developed at the U.S. CDC, based on the theoretical framework of Fellegi and Sunter and Demptser et al. *(57)*. A review of key elements for common record linkage and previous applications can be found in Chapter 2, under Data Fusion.

Disaggregate (individual level) linking is performed on the basis of several variables. The software allows these variables of interest to be specified as either blocking (Table 4-3) or matching (Table 4-4) variables. In this study, gender and the census region of the household (Northeast, Midwest, South, and West) were used as simultaneous “Blocking” variables. This means that a matching record for a male from the Northeast region (in the receiver dataset) will necessarily be obtained from males from the Northeast in the donor dataset. As such the matching on these attributes is deterministic (the software will report that a matching record could not be found is such a deterministic match is not possible).

Variables that are not blocked are matched probabilistically. Three matching methods are employed: (a) exact matching (binary all-or-nothing comparison), (b) generic string (comparison of characters and the needed number of operations needed to transform one string into the other), and value-specific (sets weights matching values based on the frequencies of values in the file being compared).

Matching variables and methods are included in Tables 4-3 and 4-4. In preparation for matching, certain variables were recoded from their usual binary or continuous values. The use of generic string in age, race, employment and income accounts for the ordinal associations or expected associations by coding these as a series of similar strings (i.e., match young adults with other young adults (20-39), similarly for middle adults (40-59) and late adults (60+); increase likelihood of matching minorities with other minorities (Hispanic, Asian, Black, other), and Hispanics with Caucasian and Black; decrease likelihood of full-time or part-time employment to be matched with unemployed individuals). The age string was constructed as a letter describing age range (A, <20; 20 < B < 40; 40 < C < 60; D, >60). The ethnicity/race string uses and XYZ format, where X assigns minority status for Hispanics, Asians, and

Blacks (yes/no), Y differentiates between Caucasian, Hispanic, Black, Asian, and Other, and Z assigns a similar value to Caucasian and Blacks due to the possible misrepresentation of Hispanic populations as Caucasian or Black. The race match hierarchy follows: perfect match (e.g., Caucasian with Caucasian), Hispanic-related minorities (e.g., Hispanic with Black), minorities (e.g., Asian with Black) or Hispanic-related (e.g., Hispanic with Caucasian), Caucasian and non-Hispanic-related (e.g., Caucasian with Asian), and last any catch-all “Other” ethnicity (e.g., Other with Hispanic). Employment match follows the next hierarchy: perfect match (e.g., unemployed with unemployed), any employed (e.g., Part-time with Full-time), unemployed/part-time, and last unemployed/full-time. Income hierarchy follows: perfect match (e.g., low-income with low-income), one-level difference (e.g., high-income with medium-income), two-level difference (e.g., high-income with low-income), presence of missing (e.g., missing with high-income).The employment of value-specific type matching for household size lies in the expected resolution to prioritize match of unusual cases, usually higher household sizes, with each other (e.g., household of six with another household of six).

The M-probability is the probability that a matching variable agrees given that the comparison pair being examined is a match; the same M-probability applies for all records. Likewise one-to-many matches are possible and allowed. The u-probabilities are calculated based on the donor dataset, and represents the probability that a matching variable agrees given that comparison pair being examined as a non-match. The u-probabilities are estimated based on the frequencies of reported values in the donor dataset.

For each record pair, the weights across all fields are calculated from M-probabilities and u-probabilities to obtain the total weight. The highest weights corresponds to pairs which agree across all fields prior probability that a randomly selected record from file A matches a randomly selected record from file B. Threshold score values assign true matches, matches for review, and non-matches. For this exercise, matches for review are assumed to be true matches and the cut-off value for non-matches is a weight of zero.

Multiple donor dataset size iterations revealed the largest available dataset (full donor sample) as the best-performing. However, the duplication of cases within the sample that share matching attributes but differ on reported BMI (matches for review) led to the adoption of the multi-run mean as a way to recognize these differences. The mean BMI was determined by resorting and rematching the donor dataset five times and finding the average of the five matched BMI values.

Validation

A comparison of matched BMI, predicted BMI by regression and reported BMI shows the record linkage being able to provide a more varied range of health measures than the regression values (BMI range of 31.6 versus 13.3). Likewise, the loose linkage criteria allowed the matching of Obese II category individuals, values that were not predicted by regression models. Figures 4-1 thru 4-4 show a comparison of predicted BMI, the matched ATUS BMI, frequency and their error distribution. Inspection of errors (Table 4-5) shows similar performance of the mean matched model in comparison to the deterministic model's Mean Absolute Error (MAE), variance and average error. A comparison of matched and reported BMI, as shown on Figure 4-1 (each color represents a different matching run) allowed for some assignment of extreme values interchangeably, undesirable but arguably an example of similarities across unhealthy BMIs in both spectrums (under and overweight). Conversely, the deterministic model offered a narrower match, bounded by normal and overweight predictions. This contrast between matched and deterministic approach is also evident when considering the frequency distributions shown in Figure 4-3 (B) and Figure 4-3 (H) against the known reported values as shown on Figure 4-3 (A).

This probabilistic linkage method allowed for matching of ATUS BMI and SAPHS information to NHTS pairs. The validation of record linkage within the ATUS cases allowed for all but 5 cases to receive a record pair. While a deterministic approach is dominated by the sample mean, the probabilistic linkage may reflect BMI ranges (i.e., normal, overweight, obese) more closely to expected values.

Table 4-1. Descriptive statistics of ATUS donor and receiver databases.

Classification	Description	Donor	Receiver
Sample Size	No. randomly-selected member per household	25,599	10,000
Sex	Male	45.0%	44.9%
	Female	55.0%	55.0%
Age	Average age of respondent	46.1	45.9
Race	Identifies as Caucasian	69.0%	69.0%
	Identifies as Black/African American	13.0%	13.0%
	Identifies as Hispanic	13.0%	13.0%
	Identifies as Asian	3.0%	3.0%
Education	Did not finish high-school	16.7%	16.9%
	High-school/GED degree	28.5%	26.2%
	Completed or some advanced-degree education	57.0%	57.0%
Employment	Person enrolled in school (any level, part- or full-time)	6.0%	5.9%
	Person identified as unemployed	18.0%	17.0%
	Person identified as part-time worker	12.0%	11.0%
	Person identified as full-time worker	54.0%	55.0%
Household	Household members	2.75	2.80
Income	Combined household income up to $35k	31.4%	31.5%
	Combined household income between $35k-$75k	29.6%	28.7%
	Combined household income greater than $75k	25.6%	26.3%
HH Tenure	Residence owned by a household member	75.0%	75.0%
Region	Household located in Northeast census region	17.0%	16.0%
	Household located in Midwest census region	25.0%	24.0%
	Household located in South census region	36.0%	37.0%
	Household located in West census region	22.0%	23.0%
Metro	Household located within a Metropolitan Area	82.0%	82.0%

Table 4-2. Deterministic model Dataset B employed in validation.

Parameter	B	S.E.	p-value
Constant	24.8	0.30	<0.01
Age	0.04	0.00	<0.01
Female	-0.84	0.09	<0.01
Black	2.02	0.13	<0.01
Hispanic	0.70	0.14	<0.01
Asian	-2.54	0.25	<0.01
Student	-2.19	0.21	<0.01
Unemployed	0.59	0.16	<0.01
Full-time work	0.92	0.14	<0.01
Retired (work)	-0.71	0.19	<0.01
College education	-0.21	0.09	0.02
Low income HH	0.70	0.10	<0.01
Other multi-adult HH	0.38	0.13	<0.01
Household size	0.10	0.03	<0.01
HH tenure	-0.30	0.11	0.01
Metropolitan area	-0.47	0.11	<0.01
Midwest US	0.64	0.13	<0.01
Southern US	0.40	0.12	<0.01
Northeast US	0.35	0.14	0.01

Table 4-3. Blocking variables and method.

Data item	Type	Matching method
Gender	Binary	Exact*
Midwest census region	Binary	Exact*
Northeast census region	Binary	Exact*
Southern census region	Binary	Exact*
West census region	Binary	Exact*

*Note: Blocking all variables as AND and not OR

Table 4-4. Matching variables and methods.

Data item	Type	Matching method	M-prob	U-prob
College education	binary	exact	0.95	0.5186
Metropolitan area	binary	exact	0.95	0.6593
Household tenure	binary	exact	0.95	0.6953
Student	binary	exact	0.95	0.9087
Age	string	generic string	0.95	0.0168
Race	string	generic string	0.95	0.6360
Employment	string	generic string	0.95	0.3788
Income	string	generic string	0.95	0.3309
Household size	continuous	value-specific	0.95	0.2191

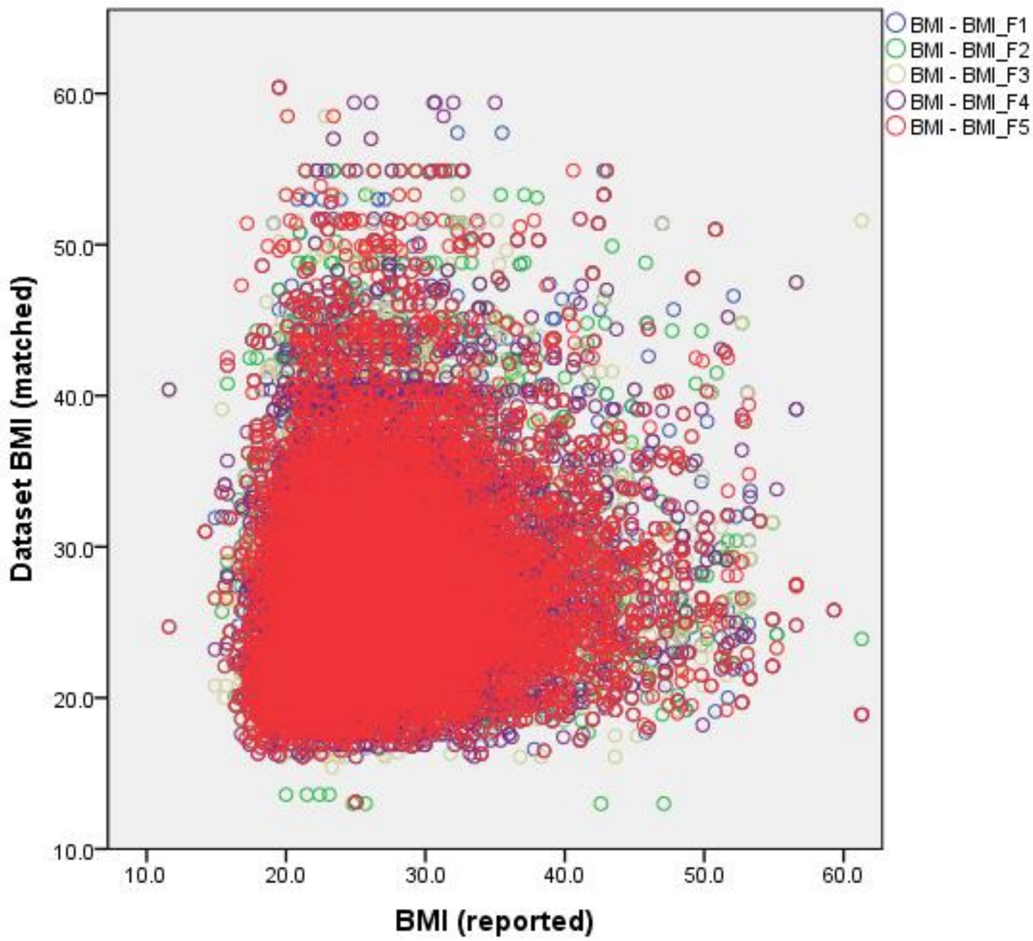

Figure 4-1. Matched BMI vs. reported BMI for five runs.

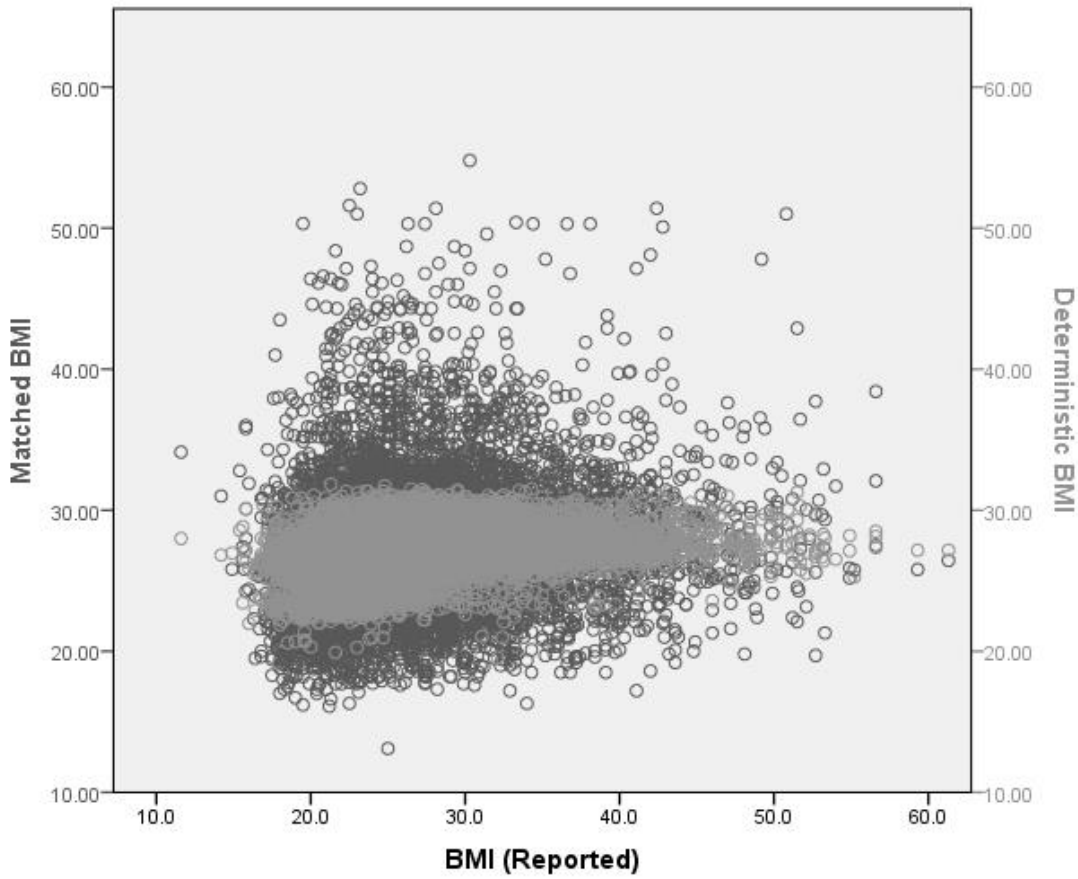

Figure 4-2. Mean matched BMI (dark) and predicted BMI (light) vs. reported values.

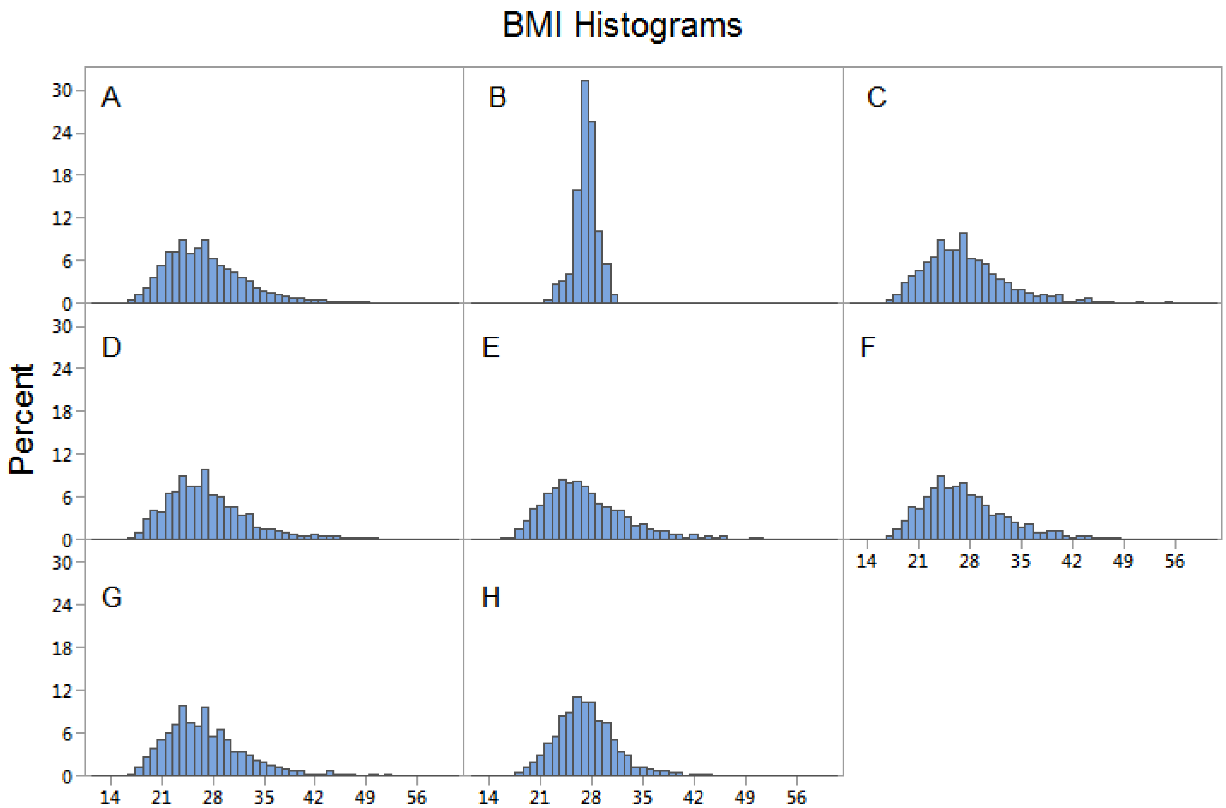

Figure 4-3. Frequency distribution of BMI: (A) reported, (B) deterministic model, (C-G) matched runs 1-5, and (H) mean matched.

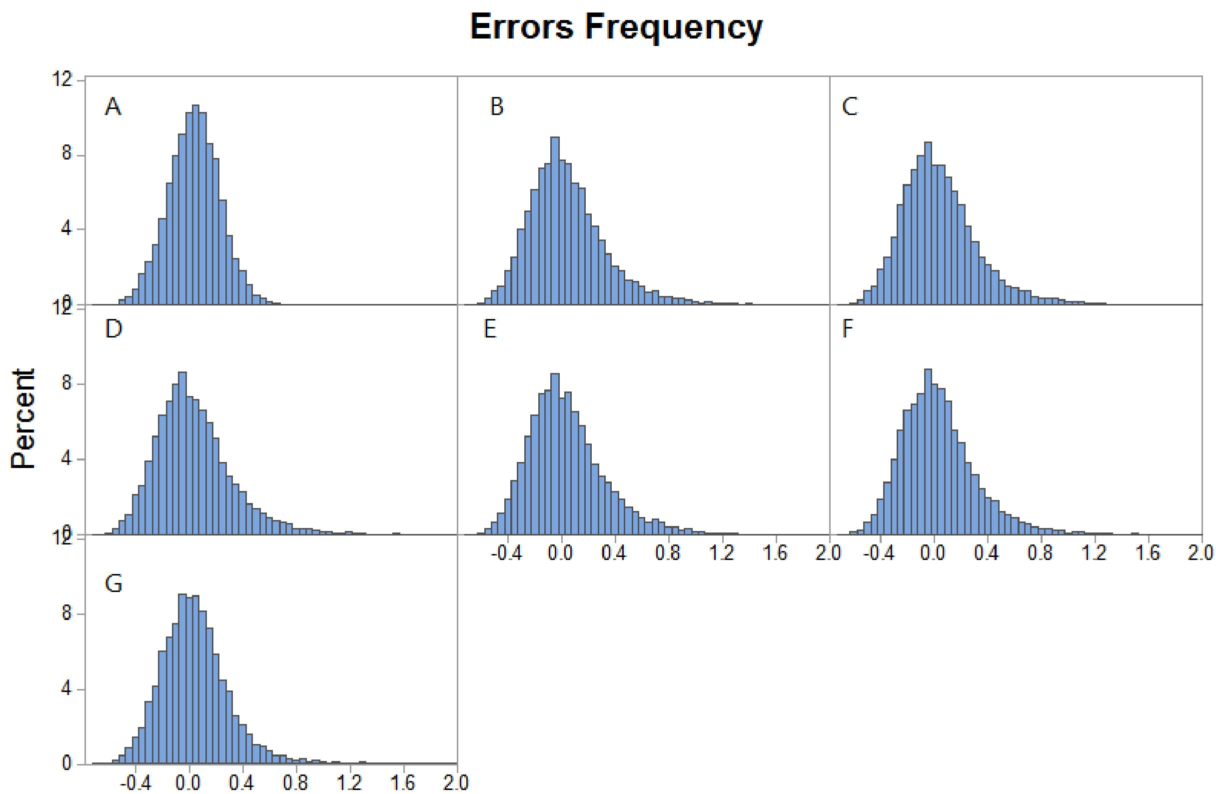

Figure 4-4. Error distributions: (A) deterministic model, (B-F) matched runs 1-5, (G) mean matched.

Table 4-5. Errors across models.

Model	MAE	Average error	Variance
Deterministic	0.16	0.038	0.037
Match 1	0.22	0.043	0.083
Match 2	0.22	0.041	0.082
Match 3	0.22	0.039	0.085
Match 4	0.22	0.041	0.084
Match 5	0.21	0.033	0.082
Mean match	0.19	0.039	0.060

CHAPTER 5
USE OF THE FLORIDA NHTS TO LINK LAND USE AND HEALTH

The object of this chapter is to explore the relationship between land use, healthy individuals and their propensity to walk. This chapter employs matched health information from ATUS with of Florida households that participated in the NHTS.

Data Assembly

The assembly uses data from the 2006-2008 American Time Use Surveys (ATUS) and the 2009 National Household Travel Survey (NHTS). Both ATUS and NHTS collected detailed socio-economic, demographic, and one-day activity-travel information for a large sample of persons. Initial sample includes one household member from each household that (a) has no missing trip purpose, gaps or overlap in their individual trip schedule, (b) is not reported as being out of the country during the travel survey day, (c) is 15 years old or older. Individuals who stay at home (no out-of-home activities) are included in the sample.

Household selection: For the NHTS, all Florida households with known location and with weekday travel surveys were selected (close to 9,000). For the ATUS, all available cases were selected, or about 35,000 households.

Person selection: To meet ATUS limitations, all individuals under age 15 or with unknown age were removed from NHTS sample and only one person per household was selected. Selection of that individual included all one-person household members and random selection of household member for multiple-member households. For households with more than one individual, uniform random values were assigned to all and the person with smallest value was selected.

The NHTS trip files reports one-day out-of-home travel from 4:00 am to 4:00 am (of the following day) on a specified travel day. Trip duration and reported mode are used to determine out-of-home utilitarian walking. Walking trips that did not end at home by the end of the day, would be assigned the last known trip purpose and duration calculated to 4 a.m. of the following day.

Sample

For this study, both socio-economic and walking data are available for 5,168 (Florida NHTS) and 35,599 (ATUS) individuals. A comparison of shared sample characteristics is presented in Table 5-1, additional description of ATUS dataset activities can be found in Chapter 3 Tables 3-1 to 3-2. The ATUS and NHTS samples have similar attributes, with age being a discerning factor for Florida. To address the Florida age gap, the final NHTS sample includes adults under the age of 70. The final sample also excludes individuals surveyed on weekend (about 20% of the cases), as the model application focused on weekday travel behavior. The represented samples are largely Caucasian and there are more females than males (55% versus 45%). The survey respondents had some secondary-degree education and were significantly likely to be employed full-time. The most common household lives in an owned property within a Metropolitan Statistical Area according to 2000 Census (have at least one urbanized area of 50,000 or more population, plus adjacent territory that has a high degree of social and economic integration with the core).

Data

This analysis uses a representative subset of weekday behavior for 5,168 people that participated in the 2009 Florida add-on of the NHTS. The NHTS collected detailed socio-economic, demographic, and one-day travel information for a large sample of persons. A comparison of walking pattern yields higher shares of utilitarian walking reported in the NHTS (16.6% vs 5%). Employing the 10-min and 15-min bout recommendations described in Chapter 2, 12.1% and 9.4% of the entire weekday sample employed walking mode for at least one leg of their trips, respectively. The add-on component enhances walking information by identifying the spatial location of the household as a set of coordinates, allowing for supporting information on the surrounding neighborhood, a common trip producer and attractor.

The ATUS-Florida NHTS match followed the steps recreated in Chapter 4. A successful record linkage allowed for 100% of the NHTS households to be matched and follow similar demographics of ATUS (Table 5-1). As Figure 5-1 demonstrates, matching occurred over a wide range of BMI values. Table 5-2 shows a BMI and observed utilitarian walking breakdown by walking guidelines.

Built environment information is derived from the U.S. Environmental Protection Agency Smart Location Database (EPA SLD) *(70)*, a nationwide geographic dataset at Census block group-level of detail. The dataset included more than 90 attributes summarizing characteristics such as housing density, diversity of land use, neighborhood design, destination accessibility, transit service, employment, and demographics from the 2010 decennial Census, the 2010 American Community Survey, 2011 InfoUSA, the Census Longitudinal Employer-Household Dynamics (LEHD) survey, 2010 Census TIGER/Line, Navteq Navstreets, and the 2011 General Transit Feed Specification (GTFS).

Land Use Descriptors

For this chapter, an individual's household was geolocated within the state of Florida using the provided Florida add-on latitude and longitude in ArcGIS. Weekday trips and activities were reported for a period of 24-hours by all household members. Census block-level information, taken from the EPA SLD dataset was superimposed to aggregate and apportion population, employment, and transportation information at ½-mile and 1-mile buffer radii from household location (see Table 5-3).

Employment was measured through the employment mix variable with the entropy denominator set to account observed existing employment types (retail, office, service, industrial, entertainment, education, healthcare, and public administration) within each block group. The SLD classified all street links as either auto-oriented (pedestrian restricted or burdensome), multi-modal (urban, arterials), or pedestrian-oriented (low-speed arterials, locals, trails) and calculated the facility-miles for each. Counts of intersections for each classification accounted for three- and four-link intersections within the group. Transit stop information was not available for the entire state and therefore not directly accounted.

Model

At first glance, BMI and reported walking behavior may not offer much differentiation, therefore BMI is modeled using a linear regression model as part of a two-stage approach. Given that two definitions of "walking" have been adopted (based on minimum duration thresholds of 10 and 15 minutes), all the models are estimated using each of these definitions of walking. Because the short-term (daily) decision of walking on a single weekday may not be a truly "exogenous" predictor of health outcome measures that are generally descriptors of longer-term condition, an instrumental-variables approach is adopted with a “predicted probability of walking on a weekday” being used as the instrument. Recognizing that leisure walking/exercise could play a substitute or complementary role, a “first-stage” OLS model was developed to determine the probability of utilitarian walking as a function of the socio-economic characteristics of the individual and his/her household and other characteristics of time-use on that day (Tables 5-4 and 5-5). These models use each of the two time definitions of “walking,” and are then used to calculate the predicted probability of walking for each person in the analysis sample. The Nagelkerke R^2 goodness of fit value is reported for comparison, as the calculated value is normalized to be between 0 and 1. The predicted probabilities of walking are used as the explanatory variables in the “second stage” models for health. These models employ the mean BMI matching methodology developed from ATUS sample in the previous chapter. The 35,599 cases available for the ATUS sample were matched in five instances to 5,168 available Florida add-on person-households. The mean BMI value was then assigned to each person for subsequent analysis.

Findings

As Figure 5-1 indicates, disaggregate matching of ATUS health values to NHTS respondents occurred over a wide range of BMIs, particularly those in the overweight and normal range. The results from the “second stage” comprehensive models for health outcomes are presented in Table 5-6. All effects reported are statistically significant at 95% confidence or higher. The predicted probability of weekday walking (10 minutes or more) showed no impact on BMI after controlling for respondent and household characteristics as well as a land-use attribute. A similar finding occurs with the predicted probability of weekday walking 15 minute-recommendation. This is not consistent with the national sample trend. The availability of land-use variables and household location showed discernible impact using the half-mile radius. Miles of pedestrian-oriented facilities were associated with lower BMIs. No other population or entropy measure remained significant after accounting for personal demographics and household characteristics. Among these characteristics, increased education and higher household incomes were associated with lower BMIs in the Florida sample.

Overall, this analysis was able to show an application of a utilitarian walking, land use and health measurement model by linking the NHTS (travel), ATUS (health), and SLD (land-use) datasets. The use of a statewide Florida travel survey and a land use database reflected the impact of pedestrian facilities up to a half mile from the household location on walking for transportation, even after controlling for socioeconomic indicators. In particular, greater pedestrian facilities were associated with desirable BMIs, but the strength of this association remains uncertain. Practitioners acquainted with the NHTS or other travel surveys with discrete household location could benefit from this methodology when enhancing their models. Outcomes from the Florida models continues to support the positive impact of a strong pedestrian-oriented network around residential locations (low-speed roads, sidewalks, trails and paths).

Table 5-1. Comparison of Florida NHTS sample and the ATUS sample.

	Matched ATUS donor attributes					NHTS
	Run 1	Run 2	Run 3	Run 4	Run 5	Receiver
Sample size	35,599	35,599	35,599	35,599	35,599	5,168
% matched	100%	100%	100%	100%	100%	n/a
Gender						
Male	46.5%	46.5%	46.5%	46.5%	46.5%	46.5%
Female	53.5%	53.5%	53.5%	53.5%	53%	53.5%
Age						
<20	2.2%	2.2%	2.2%	2.2%	2.2%	2.1%
20-39	16.2%	16.2%	16.2%	16.2%	16.2%	16.4%
40-59	44.1%	44.2%	44.2%	44.1%	44.1%	44.1%
60-69	37.5%	37.4%	37.4%	37.5%	37.5%	37.5%
Race						
Hispanic	7.4%	7.5%	7.5%	7.4%	7.5%	9.2%
Black	7.2%	7.2%	7.5%	7.4%	7.5%	6.0%
Asian	0.4%	0.4%	0.4%	0.5%	0.4%	0.9%
Caucasian	84.2%	84.1%	83.9%	83.9%	83.8%	81.3%
Education						
College (some/completed)	23.1%	23.1%	23.1%	23.1%	23.1%	21.2%
Employment						
Partial/multiple jobs	26.4%	26.3%	28.8%	29.6%	28.5%	12.5%
Full-time job	34.5%	34.5%	32.1%	31.3%	32.3%	48.4%
Household						
Members	2.09	2.09	2.09	2.09	2.09	2.10
Owned HH	90.8%	90.8%	90.8%	90.8%	90.8%	90%
HH Income <$35k	22.2%	23.9%	23.0%	21.1%	26.9%	26.9%
HH Income $35k-$75k	40.0%	39.4%	39.7%	39.4%	38.0%	38.0%
HH Income $75k+	28.6%	28.5%	28.7%	28.7%	28.0%	28.0%
Location						
Within Metropolitan region	81.2%	81.2%	81.2%	81.2%	81.2%	79%
Western US Region	0.2%	0.3%	0.2%	0.3%	0.1%	n/a
Northeast US Region	0.2%	0.2%	0.2%	0.1%	0.25%	n/a
Midwest US Region	0.1%	0.1%	0.1%	0.1%	0.12%	n/a
Southern US Region	99.5%	99.5%	99.5%	99.5%	99.50%	100%

Note: n/a (not applicable)

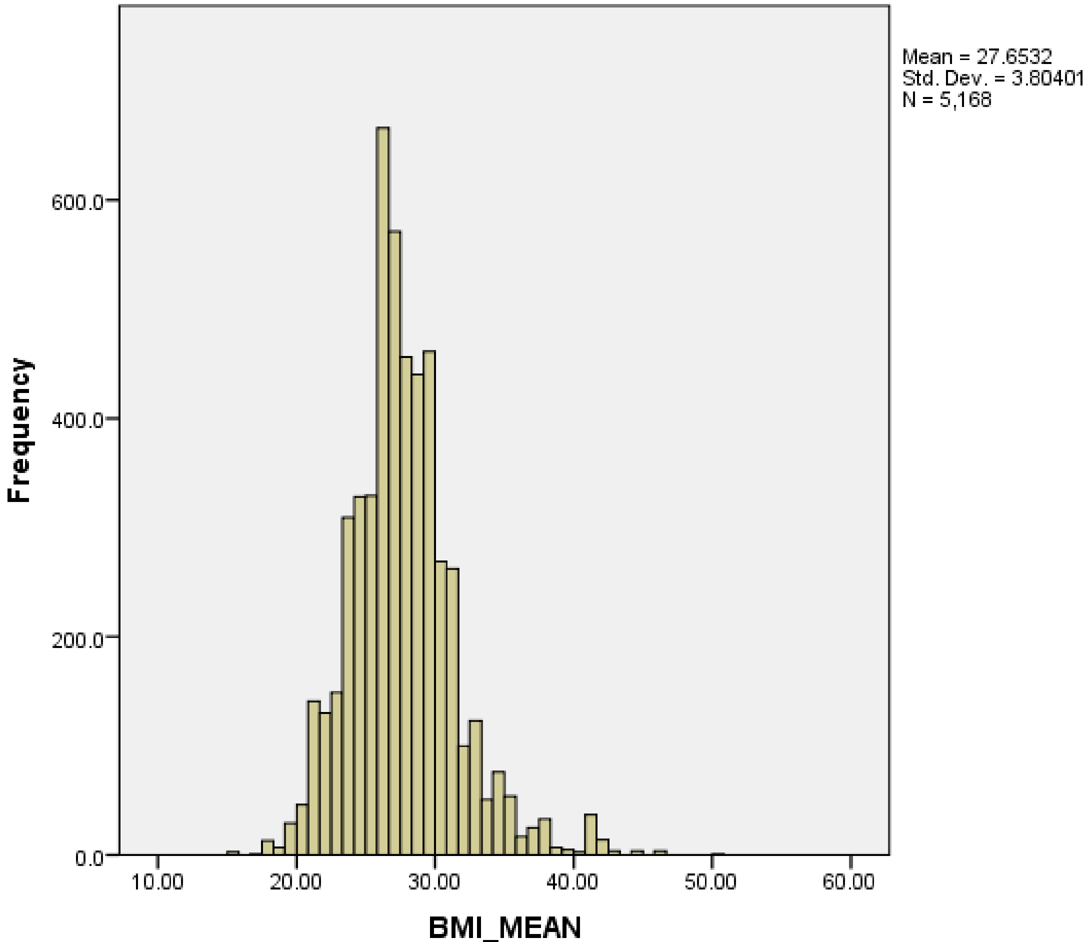

Figure 5-1. Histogram of BMIs matched to Florida sample.

Table 5-2. BMI and walking participation in Florida sample.

	Duration	
BMI range	10-min	15-min
Underweight (all)	15.0%	5.0%
Normal	11.9%	9.2%
Overweight	12.3%	9.5%
Obese (I,II,III)	11.8%	9.4%

Table 5-3. Land use descriptors

	1/2-mi buffer			1-mi buffer		
	Min	Max	Mean	Min	Max	Mean
Total population within nearby CBSA	no CBSA	5,564,635	2,155,637	no CBSA	5,564,635	2,151,154
Total workers within nearby CBSA	no CBSA	2,118,833	837,350	no CBSA	2,118,833	835,791
Total employment within nearby CBSA	no CBSA	2,088,064	822,581	no CBSA	2,088,064	820,990
Population	2.3	2895	406	7.40	4855	441
Households	1.6	1247	188	3.48	2081	207
Workers	0.9	1060	157	2.72	1694	166
Total jobs	0.0	2787	125	0.00	5540	144
Retail jobs	0.0	554	20	0.00	1186	21
Office jobs	0.0	967	12	0.00	388	13
Industrial jobs	0.0	1068	20	0.00	2345	27
Service-related jobs	0.0	863	22	0.00	867	25
Entertainment jobs	0.0	690	18	0.00	739	19
Education jobs	0.0	1197	8	0.00	1948	12
Health-affiliated jobs	0.0	845	18	0.00	911	21
Public-service jobs	0.0	743	6	0.00	458	6
8-tier employment entropy [0-1]	0.0	1.0	0.6	0.00	0.6	0.1
Miles of auto-oriented links	0.0	10.0	1.2	0.00	33665.9	35.7
Miles of multimodal-oriented links	0.0	7.3	1.3	0.00	25.3	5.3
Miles of pedestrian-oriented links	0.3	28.9	11.3	1.30	93.4	42.7
Auto-oriented intersections	0.0	12.0	0.9	0.00	40.7	3.5
Pedestrian oriented intersections having four or more legs	0.0	25.0	10.4	0.00	405.9	49.8

Table 5-4. Predicted weekday 10-min walk

Parameter	**B**	**S.E.**	**Sig.**
Past week walk trips	0.057	0.005	<0.001
Part-time job	-0.350	0.145	0.016
Full-time job	-0.428	0.097	<0.001
HH income >75k	0.289	0.099	0.004
Shared HH vehicle	-0.995	0.221	<0.001
Non-shared vehicle HH	-1.065	0.200	<0.001
Total Educational jobs [1-mi]	0.001	0.000	0.028
Total pedestrian links [1-mi]	0.008	0.002	<0.001
Constant	-1.526	0.224	<0.001
Goodness of fit			
Nagelkerke R^2	0.080		

Table 5-5. Predicted weekday 15-min walk

Parameter	**B**	**S.E.**	**Sig.**
Past week walk trips	0.050	0.005	<0.001
Part-time job	-0.380	0.162	0.019
Full-time job	-0.562	0.109	<0.001
HH income >75k	0.327	0.114	0.004
Shared HH vehicle	-0.929	0.243	<0.001
Non-shared vehicle HH	-1.129	0.209	<0.001
Number of HH adults	-0.188	0.082	0.023
Total pedestrian links [1-mi]	0.006	0.002	0.012
Constant	-1.22	0.264	<0.001
Goodness of fit			
Nagelkerke R^2	0.074		

Table 5-6. Effect of land use and predicted walking on health-BMI

Parameter	B	S.E.	Sig.
Predicted 10-min walking probability	n/s		
Predicted 15-min walking probability	n/s		
Total pedestrian links (mi) [1/2-mi]	-0.025	0.009	0.006
Age	0.027	0.004	<0.001
Black	0.780	0.221	<0.001
Hispanic	0.823	0.186	<0.001
College education	-0.690	0.128	<0.001
Part-time job	-0.401	0.170	0.018
Full-time job	-0.481	0.117	<0.001
HH children (total)	0.192	0.097	0.048
Medium-income HH	-0.312	0.125	0.013
High-income HH	-1.296	0.139	<0.001
Constant	27.273	0.279	<0.001
Goodness of fit			
Adjusted R^2	0.054		

Note: n/s (not significant at p=0.05)

CHAPTER 6
CONCLUSION: WHERE DO WE GO FROM HERE?

Utilitarian walking can be part of a set of strategies employed to facilitate recommended physical activity levels for healthy adults. Despite the fact the qualitative impact of walking on health can be inferred from the energy-balance model, empirical studies that capture the effects of land-use patterns are necessary to ascertain the magnitude of the impact of walking on health. While studies on built environment, walking and physical health exists in the fields of public health, urban planning, and transportation, comprehensive models addressing all their aspects are limited by lack of health data, in particular transportation surveys. Health surveys often do no collect travel behavior data and/or do not provide detailed spatial resolution to construct land-use descriptors. This research supports the incorporation of health measures such as Body Mass Index and Self-Assessed Physical Activity Score (SAPHS) in future transportation studies and household travel surveys.

The availability of one reported (BMI) and one perceived measure (SAPHS), allowed for their contrast as well. While BMI exemplifies a longer-term variable, SAPHS can represent shorter durations. Higher BMI does constrict utilitarian walking; a lower SAPHS (poor health perception) did not seem to carry this pattern in a national sample after controlling for exercise, nutrition, and socioeconomic status. Utilitarian walking has often been measured in inconsistent ways. The review and adoption of two threshold values already part of health policy guidelines (continuous bouts of 10 and 15 minutes) reinforce the expected relationship, while supporting a diminishing magnitude after controlling for multiple indicators and separate weekend effects. It is evident that exercise and nutrition continue to play a substantial role.

This research also addressed the travel diary and health gap in the literature by merging transportation and health surveys using data-fusion techniques to develop comprehensive models of land-use, utilitarian walking, and health. The use of probabilistic matching to populate a statewide Florida travel survey with BMI values proved a practical application. Record linkage allowed for the production of extreme cases and an aggregate distribution similar to the general population.

The national sample showed some impact of utilitarian walking on health, while the Florida sample, with an older population, did not show such correlation. The presence of pedestrian facilities within a half-mile radius of the household and dense locations continue to be associated with lower BMIs for the Florida statewide sample.

The work conducted continues to support much of the knowledge gained from local, scoped studies through a large, national survey. It also offers an alternative to unknown health information through the use of probabilistic matching of common parameters to link NHTS person files with ATUS records. Lastly, the work highlighted the need to continuously support pedestrian facilities and cautions how increasing U.S. overweight and obesity rates could be detrimental to future mode shifts to active transportation.

Future research should aim at examining the relationships between long-term walking and health as well as incorporating users' perception of nearby infrastructure. The availability of individual nutrition patterns could also enhance current research. Moving forward, researchers and policy makers can re-examine the impact and projections on active transportation and relevant transportation facilities as the U.S. population experiences changes in its health profile. Similarly, practitioners can incorporate available health measures at earlier stages of their planning and modeling processes as opposed to final-stage outputs of health and air-quality. From a Florida perspective, the unique composition of the state population may warrant a closer inspection of transportation facilities among the elderly and young adults.

LIST OF REFERENCES

1. Blackwell, D. L., J. W. Lucas, and T. C. Clarke. Summary Health Statistics for U.S. Adults: National Health Interview Survey, 2012. *Vital and Health Statistics. Series 10, Data from the National Health Survey*, Vol. 260, No. 260, 2014, pp. 1-171.

2. Iacono, M., K. J. Krizek, and A. El-Geneidy. Measuring Non-Motorized Accessibility: Issues, Alternatives, and Execution. *Journal of Transport Geography*, Vol. 18, No. 1, 2010, pp. 133-140.

3. U.S. Department of Transportation, Federal Highway Administration. 2009 National Household Travel Survey, 2014. http://nhts.ornl.gov. Accessed Jul. 31, 2014.

4. U.S. Census Bureau. 2011 American Community Survey, 2014. http://www.census.gov/acs/www/data_documentation/2011_release/. Accessed Jul. 30, 2014.

5. Milne, A., and M. Melin. Bicycling and Walking in the United States: 2014 Benchmarking Report. 2014, pp. 268.

6. U.S. Environmental Protection Agency. Transportation Conformity Rule Restructuring Amendments. 77 FR 14979, 2012.

7. Elsom, D. Smog Alert: Managing Urban Air Quality. Routledge, 1996.

8. Bernardo, C., and C. Bhat. Non-Motorized Travel as a Sustainable Travel Option. *Handbook of Sustainable Travel*, Springer Netherlands, 2014, pp. 277-291.

9. Sallis, J., A. Bauman, and M. Pratt. Environmental and Policy Interventions to Promote Physical Activity. *American Journal of Preventive Medicine*, Vol. 15, No. 4, 1998, pp. 379-397.

10. NHLBI Obesity Education Initiative Expert Panel. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults. 1998.

11. Tremblay, M. S., R. C. Colley, T. J. Saunders, G. N. Healy, and N. Owen. Physiological and Health Implications of a Sedentary Lifestyle. *Applied Physiology, Nutrition, and Metabolism*, Vol. 35, No. 6, 2010, pp. 725-740.

12. Slentz, C. A., B. D. Duscha, J. L. Johnson, K. Ketchum, L. B. Aiken, G. P. Samsa, J. A. Houmard, C. W. Bales, and W. E. Kraus. Effects of the Amount of Exercise on Body Weight, Body Composition, and Measures of Central Obesity: STRRIDE—a Randomized Controlled Study. *Archives of Internal Medicine*, Vol. 164, No. 1, 2004, pp. 31-39.

13. Physical Activity Guidelines Advisory Committee. Physical Activity Guidelines for Americans. *Washington, DC: US Department of Health and Human Services*, 2008, pp. 15-34.

14. United States. Department of Health. *Physical Activity and Health: A Report of the Surgeon General.* DIANE Publishing, 1996.

15. United States. Public Health Service. Office of the Surgeon General, National Center for Chronic Disease Prevention, Health Promotion (US), President's Council on Physical Fitness, and Sports (US). *Physical Activity and Health: A Report of the Surgeon General.* Jones & Bartlett Learning, 1996.

16. Sallis, J. F., L. D. Frank, B. E. Saelens, and M. K. Kraft. Active Transportation and Physical Activity: Opportunities for Collaboration on Transportation and Public Health Research. *Transportation Research Part A: Policy and Practice*, Vol. 38, No. 4, 2004, pp. 249-268.

17. Lawrence, D. L., and S. M. Low. The Built Environment and Spatial Form. *Annual Review of Anthropology*, 1990, pp. 453-505.

18. Handy, S.L., M.G. Boarnet, R. Ewing, and R.E. Killingsworth. How the built environment affects physical activity: Views from urban planning. *American Journal of Preventive Medicine*, Vol 23, No. 2, Supplement 1, 2002, pp. 64-73.

19. Sallis, J. F., N. Owen, and E. B. Fisher. Ecological Models of Health Behavior. *Health Behavior and Health Education: Theory, Research, and Practice*, Vol. 4, 2008, pp. 465-485.

20. Corbin, C. B., R. P. Pangrazi, and B. D. Franks. Definitions: Health, Fitness, and Physical Activity. *President's Council on Physical Fitness and Sports Research Digest*, 2000.

21. Frank, L.,D, J. Kerr, J. Sallis F, R. Miles, and J. Chapman. A Hierarchy of Sociodemographic and Environmental Correlates of Walking and Obesity. *Preventive Medicine*, Vol. 47, No. 2, 2008, pp. 172-178.

22. Wanner, M., T. Götschi, E. Martin-Diener, S. Kahlmeier, and B. W. Martin. Active Transport, Physical Activity, and Body Weight in Adults: A Systematic Review. *American Journal of Preventive Medicine*, Vol. 42, No. 5, 2012, pp. 493-502.

23. Pucher, J., R. Buehler, D. R. Bassett, and A. L. Dannenberg. Walking and Cycling to Health: A Comparative Analysis of City, State, and International Data. *American Journal of Public Health*, Vol. 100, No. 10, 2010, pp. 1986-1992.

24. Van Dyck, D., E. Cerin, G. Cardon, B. Deforche, J. F. Sallis, N. Owen, and I. de Bourdeaudhuij. Physical Activity as a Mediator of the Associations between Neighborhood Walkability and Adiposity in Belgian Adults. *Health & Place*, Vol. 16, No. 5, 2010, pp. 952-960.

25. Cleland, V., K. Ball, C. Hume, A. Timperio, A. C. King, and D. Crawford. Individual, Social and Environmental Correlates of Physical Activity among Women Living in Socioeconomically Disadvantaged Neighbourhoods. *Social Science & Medicine*, Vol. 70, No. 12, 2010, pp. 2011-2018.

26. Bostock, L. Pathways of Disadvantage? Walking as a Mode of Transport among Low-income Mothers. *Health & Social Care in the Community*, Vol. 9, No. 1, 2001, pp. 11-18.

27. Saunders, L. E., J. M. Green, M. P. Petticrew, R. Steinbach, and H. Roberts. What are the Health Benefits of Active Travel? A Systematic Review of Trials and Cohort Studies. *PLoS ONE*, Vol. 8, No. 8, 2013, pp. e69912.

28. Andersen, L. B., P. Schnohr, M. Schroll, and H. O. Hein. All-Cause Mortality Associated with Physical Activity during Leisure Time, Work, Sports, and Cycling to Work. *Archives of Internal Medicine*, Vol. 160, No. 11, 2000, pp. 1621-1628.

29. Hayashi, T., K. Tsumura, C. Suematsu, K. Okada, S. Fujii, and G. Endo. Walking to Work and the Risk for Hypertension in Men: The Osaka Health Survey. *Annals of Internal Medicine*, Vol. 131, No. 1, 1999, pp. 21-26.

30. Besson, H., U. Ekelund, S. Brage, R. Luben, S. Bingham, K. Khaw, and N. J. Wareham. Relationship between Subdomains of Total Physical Activity and Mortality. *Medicine and Science in Sports and Exercise*, Vol. 40, No. 11, 2008, pp. 1909-1909-1915.

31. Sallis, J. F., and N. G. Owen. *Physical Activity and Behavioral Medicine.* Sage Publications, 1998.

32. Trost, S. G., N. Owen, A. E. Bauman, J. F. Sallis, and W. Brown. Correlates of Adults' Participation in Physical Activity: Review and Update. *Medicine & Science in Sports & Exercise*, Vol. 34, No. 12, 2002, pp. 1996-2001.

33. Harris, J. K., J. Lecy, J. A. Hipp, R. C. Brownson, and D. C. Parra. Mapping the Development of Research on Physical Activity and the Built Environment. *Preventive Medicine*, Vol. 57, No. 5, 2013, pp. 533-540.

34. Ding, D., and K. Gebel. Built Environment, Physical Activity, and Obesity: What have we Learned from Reviewing the Literature? *Health & Place*, Vol. 18, No. 1, 2012, pp. 100-105.

35. Ferdinand, A. O., B. Sen, S. Rahurkar, S. Engler, and N. Menachemi. The Relationship between Built Environments and Physical Activity: A Systematic Review. *American Journal of Public Health*, Vol. 102, No. 10, 2012, pp. e7-e13.

36. Heath, G. W., R. C. Brownson, J. Kruger, R. Miles, K. E. Powell, L. T. Ramsey, and Task Force on Community Preventive Services. The Effectiveness of Urban Design and Land use and Transport Policies and Practices to Increase Physical Activity: A Systematic Review. *Journal of Physical Activity & Health*, Vol. 3, 2006, pp. S55.

37. Durand, C. P., M. Andalib, G. F. Dunton, J. Wolch, and M. A. Pentz. A Systematic Review of Built Environment Factors Related to Physical Activity and Obesity Risk: Implications for Smart Growth Urban Planning. *Obesity Reviews*, Vol. 12, No. 5, 2011, pp. e173-e182.

38. Lachowycz, K., and A. Jones. Greenspace and Obesity: A Systematic Review of the Evidence. *Obesity Reviews*, Vol. 12, No. 5, 2011, pp. 183-189.

39. McCormack, G. R., and A. Shiell. In Search of Causality: A Systematic Review of the Relationship between the Built Environment and Physical Activity among Adults. *Int J Behav Nutr Phys Act*, Vol. 8, No. 125, 2011, pp. 101-186.

40. Pulleyblank Patrick, S., M. Raimi, R. Ewing, L. D. Frank, J. Chapman, and R. Kreutzer. *Understanding the Relationship between Public Health and the Built Environment: A Report Prepared for the LEED-ND Core Committee.* U.S. Green Building Council, 2006.

41. Saelens, B. E., and S. L. Handy. Built Environment Correlates of Walking: A Review. *Medicine and Science in Sports and Exercise*, Vol. 40, No. 7 Suppl, 2008, pp. 550-566.

42. Handy, S., X. Cao, and P. Mokhtarian. Correlation or Causality between the Built Environment and Travel Behavior? Evidence from Northern California. *Transportation Research Part D: Transport and Environment*, Vol. 10, No. 6, 2005, pp. 427-444.

43. Frank, L. D., B. E. Saelens, K. E. Powell, and J. E. Chapman. Stepping Towards Causation: Do Built Environments Or Neighborhood and Travel Preferences Explain Physical Activity, Driving, and Obesity? *Social Science & Medicine*, Vol. 65, No. 9, 2007, pp. 1898-1914.

44. Knuiman, M. W., H. E. Christian, M. L. Divitini, S. A. Foster, F. C. Bull, H. M. Badland, and B. Giles-Corti. A Longitudinal Analysis of the Influence of the Neighborhood Built Environment on Walking for Transportation: The RESIDE Study. *American Journal of Epidemiology*, Vol. 180, No. 5, 2014, pp. 453-461.

45. Mackenbach, J. D., H. Rutter, S. Compernolle, K. Glonti, J. Oppert, H. Charreire, I. De Bourdeaudhuij, J. Brug, G. Nijpels, and J. Lakerveld. Obesogenic Environments: A Systematic Review of the Association between the Physical Environment and Adult Weight Status, the SPOTLIGHT Project. *BMC Public Health*, Vol. 14, No. 1, 2014, pp. 233.

46. Feng, J., T. A. Glass, F. C. Curriero, W. F. Stewart, and B. S. Schwartz. The Built Environment and Obesity: A Systematic Review of the Epidemiologic Evidence. *Health & Place*, Vol. 16, No. 2, 2010, pp. 175-190.

47. Hess, D. B., and J. K. Russell. Influence of Built Environment and Transportation Access on Body Mass Index of Older Adults: Survey Results from Erie County, New York. *Transport Policy*, Vol. 20, March, 2012, pp. 128-137.

48. Wen, M., and L. Kowaleski-Jones. The Built Environment and Risk of Obesity in the United States: Racial–ethnic Disparities. *Health & Place*, Vol. 18, No. 6, 2012, pp. 1314-1322.

49. Rundle, A., A. V. Diez Roux, L. M. Free, D. Miller, K. M. Neckerman, and C. C. Weiss. The Urban Built Environment and Obesity in New York City: A Multilevel Analysis. *American Journal of Health Promotion*, Vol. 21, No. 4 Suppl, 2007, pp. 326-334.

50. Bauman, A., and F. Bull. Environmental Correlates of Physical Activity and Walking in Adults and Children: A Review of Reviews. *London: National Institute of Health and Clinical Excellence*, 2007.

51. Frank, L. D., J. F. Sallis, T. L. Conway, J. E. Chapman, B. E. Saelens, and W. Bachman. Many Pathways from Land use to Health: Associations between Neighborhood Walkability and Active Transportation, Body Mass Index, and Air Quality. *Journal of the American Planning Association*, Vol. 72, No. 1, 2006, pp. 75-87.

52. Doyle, S., A. Kelly-Schwartz, M. Schlossberg, and J. Stockard. Active Community Environments and Health: The Relationship of Walkable and Safe Communities to Individual Health. *Journal of the American Planning Association*, Vol. 72, No. 1, 2006, pp. 19-31.

53. Mackenbach, J., H. Rutter, S. Compernolle, K. Glonti, J. Oppert, H. Charreire, I. De Bourdeaudhuij, J. Brug, G. Nijpels, and J. Lakerveld. Obesogenic Environments: A Systematic Review of the Association between the Physical Environment and Adult Weight Status, the SPOTLIGHT Project. *BMC Public Health*, Vol. 14, No. 1, 2014, pp. 233.

54. Gebel, K., A. E. Bauman, T. Sugiyama, and N. Owen. Mismatch between Perceived and Objectively Assessed Neighborhood Walkability Attributes: Prospective Relationships with Walking and Weight Gain. *Health & Place*, Vol. 17, No. 2, 2011, pp. 519-524.

55. Ullmann, S. H., N. Goldman, and A. R. Pebley. Contextual Factors and Weight Change Over Time: A Comparison between U.S. Hispanics and Other Population Sub-Groups. *Social Science & Medicine*, Vol. 90, No. 0, 2013, pp. 40-48.

56. Barry, A. *An Analysis of the Perception of the Built Environment, Physical Activity, and Body Mass Index in Allegheny County, Pennsylvania*. Doctoral Dissertation, University of Pittsburgh, 2013.

57. Glazier, R. H., M. I. Creatore, J. T. Weyman, G. Fazli, F. I. Matheson, P. Gozdyra, R. Moineddin, V. K. Shriqui, and G. L. Booth. Density, Destinations or Both? A Comparison of Measures of Walkability in Relation to Transportation Behaviors, Obesity and Diabetes in Toronto, Canada. *PloS One*, Vol. 9, No. 1, 2014, pp. e85295.

58. Gu, L., R. Baxter, D. Vickers, and C. Rainsford. Record Linkage: Current Practice and Future Directions. *CSIRO Mathematical and Information Sciences Technical Report*, Vol. 3, No. 83, 2003.

59. Jaro, M. A. Advances in Record-Linkage Methodology as Applied to Matching the 1985 Census of Tampa, Florida. *Journal of the American Statistical Association*, Vol. 84, No. 406, 1989, pp. 414-420.

60. Winkler, W. E. Frequency-Based Matching in Fellegi-Sunter Model of Record Linkage. *Bureau of the Census Statistical Research Division*, Vol. 14, 2000.

61. Baker, K., P. Harris, and J. O'Brien. Data Fusion: An Appraisal and Experimental Evaluation. *Journal of the Market Research Society*, 1989, pp. 153–212.

62. Kusakabe, T., and Y. Asakura. Behavioural Data Mining of Transit Smart Card Data: A Data Fusion Approach. *Transportation Research Part C: Emerging Technologies*, Vol. 46, 2014, pp. 179-191.

63. Amorim, M., S. Ferreira, and A. Couto. Linking Police and Hospital Road Accident Records. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2432, No. 1, 2014, pp. 10-16.

64. Rosman, D. L. The Western Australian Road Injury Database (1987–1996): Ten Years of Linked Police, Hospital and Death Records of Road Crashes and Injuries. *Accident Analysis & Prevention*, Vol. 33, No. 1, 2001, pp. 81-88.

65. Williamson, A., and S. Boufous. A Data-Matching Study of the Role of Fatigue in Work-Related Crashes. *Transportation Research Part F: Traffic Psychology and Behaviour*, Vol. 10, No. 3, 2007, pp. 242-253.

66. Thakuriah, V., J. Lee, and S. Niumpradit. Motor Carrier Safety Databases: Strategies for Developing Linkages, Assessing Completeness and Making Imputations. In *International Truck and Bus Safety Research and Policy Symposium*, 2002.

67. Pawlak, J., J. W. Polak, and A. Sivakumar. An Imputation Approach to the Fusion of Travel Diary and Lifestyle Data: Application to the Analysis of the Interaction of ICT and Physical Mobility. In *New Techniques and Technologies for Statistics Conference.,* Brussels, Belgium, 2013.

68. Kressner, J. D., and L. A. Garrow. Using Third-Party Data for Travel Demand Modeling: Comparison of Targeted Marketing, Census, and Household Travel Survey Data. In *Transportation Research Board 93rd Annual Meeting*, 2014.

69. Dunton, G. F., D. Berrigan, R. Ballard-Barbash, F. M. Perna, B. I. Graubard, and A. A. Atienza. Adolescents' Sports and Exercise Environments in a U.S. Time use Survey. *American Journal of Preventive Medicine*, Vol. 39, No. 2, 2010, pp. 122-129.

70. Ramsey, Kevin, and Alexander Bell. The smart location database: A nationwide data resource characterizing the built environment and destination accessibility at the neighborhood fiscalement and destination accessibility at the neighborhood scale. *Cityscape: A Journal of Policy Development and Research*, Vol 16, No. 2, 2014.

BIOGRAPHICAL SKETCH

Miguel Lugo graduated from the transportation engineering doctoral program within the Department of Civil and Coastal Engineering at the University of Florida Engineering School of Sustainable Infrastructure & Environment. He received a Master of Engineering in Civil Engineering with a Minor in Urban and Regional Planning from the University of Florida and a Bachelor of Science in Civil Engineering from the University of Puerto Rico at Mayagüez. During his time at Florida, he has was recognized with a pre-doctoral fellowship from the Louis Stokes Alliance for Minority Participation Program, as well as service and leadership awards through various professional and community leadership roles. His research interests include public infrastructure policy, active transportation, and travel behavior.