Built-environment Variables Influencing Aggregate Walking:

a Multivariate Analysis of Halifax Neighbourhoods

By
Kevin Neatt

A Thesis Submitted to
Saint Mary's University, Halifax, Nova Scotia
in Partial Fulfillment of the Requirements for
the Degree of Master of Arts in Geography

February 2016, Halifax, Nova Scotia

Approved: Dr. Hugh Millward
(Supervisor)

Approved: Dr. Jamie Spinney
(Reader)

Approved: Dr. Michael Fox
(External Examiner)

Date: February, 2016

Abstract

Built-environment Variables Influencing Aggregate Walking: a Multivariate Analysis of Halifax Neighbourhoods

By Kevin Neatt

This study examines neighbourhood characteristics affecting the incidence of walking trips in urban and suburban areas of Halifax, Canada, using data from the Space-Time Activity Research (STAR) survey, which was conducted in 2007-2008. Primary respondents completed a two-day time-diary survey, and their movements were tracked using a GPS data logger.

Based on mapped walking tracks, hypotheses were developed regarding variations in walking density. To test these, walking distances were aggregated by census tracts, and expressed as walking densities (per resident, per metre of road, and per developed area). Multivariate regression was used to examine which neighbourhood variables are most useful as estimators of walking densities. Contrary to much of the planning literature, built-environment measures of road connectivity and dwelling density were found to have little estimating power. Office and institutional land uses are more useful estimators, as are the income and age characteristics of the resident population.

February, 2016

Acknowledgments

This thesis would not have come to fruition without the guidance and support of several individuals. First and foremost, I would like to offer my sincere and heartfelt thank you to both Dr. Hugh Millward and Dr. Jamie Spinney who have generously shared their collective knowledge. They have provided the much needed guidance and facilitated my growth and understanding of walking behaviour throughout this endeavour.

A special thank you is also deserved for those in the faculty of Geography and Environmental Studies at Saint Mary's University who have contributed to shaping my academic career. Thank you to Dr. Robert McCalla for shepherding me through my geography studies as well as for honing my critical thinking. Thank you to Dr. Jason Grek-Martin for expanding my geographic vocabulary. Thank you to Mr. Greg Baker for navigating my way through the technical GIS aspects of this project.

I am also grateful to have received tremendous support outside of the academic setting. Firstly, to my family, my wife Nancy Neatt, thank you for encouraging me to succeed and for being understanding of the time commitments required to complete this challenge, and to my mother Gail Neatt for the unwavering support and encouragement of all of my academic goals. Finally, my employer deserves a special thank you for their support and flexibility while I undertook both undergraduate and graduate degrees; their investment in my education is truly appreciated.

TABLE OF CONTENTS

List of Figures

List of Tables

CHAPTER 1

Introduction

Neighbourhood walkability is particularly important to urban designers, planners, policy makers, and those in the public health, environmental, and transportation fields (Saelens et al., 2008; Li et al., 2015). The concern for neighbourhood walkability stems from a desire to improve one or more of the following three objectives: improve public health, reduce infrastructure costs, and reduce environmental impacts of transportation (e.g. greenhouse gas emissions).

The lack of physical activity and high obesity rates have been, and continue to be, of concern for governments in the western world. Ewing et al. (2013, p.118) contends that "physical inactivity is the fourth leading risk factor for global mortality". Health and medical researchers have promoted walking as a good form of physical activity and report that even moderate amounts can have positive impact on public health (Frank et al., 2004; Ewing et al., 2013). Researchers have been making the connection between public health and the built environment for some time (Frank et al., 2004), and Mowatt (2015) argues the relationship between the built environment and health has "entered the mainstream of public health practice". The built environment is a rather broad term that incorporates, as Handy et al. (2002, p.65) describe, "urban design, land use, and the transportation system, and encompasses patterns of human activity within the physical environment". Improvements in the built environment, such as an interconnected pedestrian system (of sidewalks, pathways, multi-use trails), mixed land uses that are within walking distance, transit friendly design, and pleasing neighbourhood aesthetics can encourage physical activity, particularly walking (Saelens et al., 2003; Leslie et al., 2005; Cerin et al., 2006).

A reduction in the reliance on automobile transport generally coincides with an increase in public transit ridership and active transportation (AT), both of which require pedestrian access and walking. Greater transit ridership and AT participation would also reduce the need for costly infrastructure improvements and future transportation investments (Cervero, 1988; Gordon, et al., 1997; Frank, 2000). Governments and policy makers are seeking to improve public policy that encourages walking in an effort to reduce pressure on public coffers by delaying or eliminating road improvements, and creating non-subsidized public transit.

Concerns surrounding climate change and greenhouse gas emissions have encouraged urban designers, planners, and policy makers to find ways of reducing reliance on automobile travel and thus encourage the use of both public transit and AT (Boarnet et al., 2011). Researchers have determined that limited or non-existent public transit and AT options strengthen the reliance on automobile dependency (Wilson et al., 2013).

Based on the concerns noted above, considerable efforts are being put into determining if neighbourhoods enable and encourage walking, often termed neighbourhood walkability. Marshal et al. (2009, p.1752) define walkability as “a measure of how conducive the built environment is to walking and that predicts physical activity and active transportation”. Researchers suggest that neighbourhood walkability can be measured by scoring several objective physical characteristics of the built environment, thereby creating an index of walkability (Saelens et al., 2003; Leslie et al., 2007; Frank et al., 2009; Mayne et al., 2013). Walkability indices can be used to evaluate neighbourhood designs and to either estimate or better understand the likelihood of physical activity of residents.

The objective of this research is to identify built environment characteristics that promote walking behaviour at the neighbourhood level, in order to test the highly-cited index of walkability (e.g. Frank et al., 2005; Lee and Moudon, 2006) against objective and verified walking data, and ultimately to improve neighbourhood design. This research uniquely contributes to the knowledge of walking behaviour in the following three ways: 1) the use of both objective and self-report time diaries to record walking activity; 2) walking data is aggregated based on census tracts (CT), enabling the CTs themselves to be considered the unit of measurement, rather than the individual respondents; and, 3) walking activity includes both active transport and recreational walking so that total walking can be examined. Examining walking behaviour using these three unique considerations provides further insights as to the spatial incidence of walking in a medium-sized city.

This study is organised into seven chapters, beginning with a comprehensive review of literature in chapter two that outlines the seminal research conducted on neighbourhood walkability and then highlights relevant walking theories that relate to three unique considerations of this study noted above. Chapter three discusses the methods employed in the study, including both the study area and the data sets. The study area is the medium-sized, metropolitan city of Halifax, Canada. This research employs data derived from the innovative Halifax STAR (Space-Time Activity Research) time-use and transport survey conducted in 2007-08. Respondents from the STAR survey were provided GPS-enabled (global positioning system) personal data loggers that tracked walking activity of respondents aged 15 years and older for a 48-hour reporting period. Walking distances were aggregated by Census Tract, and expressed as three walking densities. Chapter four – Spatial Analysis – examines the spatial patterns of the three walking densities. This chapter provides an opportunity to discuss walking patterns and complements the statistical component of this research. Chapters five and six focus on the statistical analysis of the aggregated walking densities. Chapter five employs a two-tailed, bivariate, Pearson correlation analysis to assess the statistical significance of any linear relationships among the individual walking index variables and two composite walkability indices against the three walking densities. The correlation analysis also allows us to investigate potential issues with multicollinearity. Chapter 6 reports on multiple linear regression models, which were run for each of the three measures of walking density (dependent variables). The models tested walking densities against built environment variables, socio-demographic control variables, and the walkability indices, with the goal of determining which set of variables have the strongest ability to explain walking behaviour. The final chapter concludes the study by highlighting important insights gained from the research, providing suggestions for future research directions, and discussing implications of the research findings for practical policy improvements.

CHAPTER 2

Literature Review

A substantial amount of research has been conducted on walkability over the past ten years. A search conducted in the Web of Science database (February, 2016) using the term “walkability” resulted in 845 articles to the end of 2015. Figure 2.1 illustrates the number of published items per year. There were only a small number of articles prior to 2006, and a steep and steady increase after 2006.

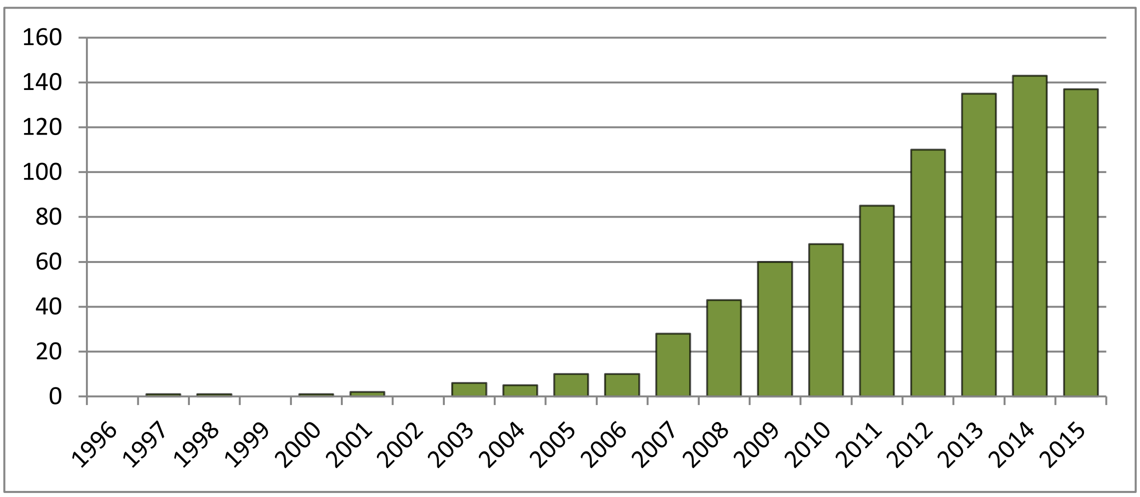


Figure 2.1. Published Walkability Articles per Year (2015) (Source: Web of Science)

A review of the literature suggests that walkability is of interest to two primary fields of academic enquiry, transportation planning and public health. Of the twenty-one articles that were cited over 100 times, all but one article was published in a health-related journal. The one exception was published in the Journal of the American Planning Association. Although there has been an explosion of interest in walkability in the last ten years, seminal research was conducted from 1988 – 2006. Figure 2.2 illustrates the timeline associated with these articles. This literature review will focus on these seminal articles and will address literature relating to three unique considerations that this research incorporates.

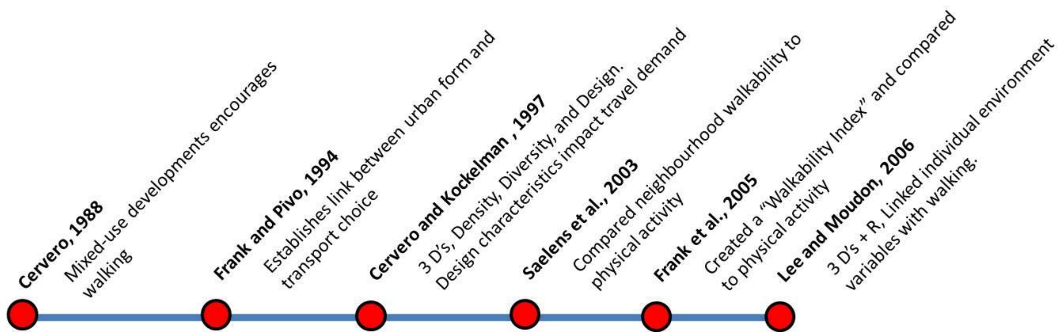


Figure 2.2. Timeline of Seminal Walkability Articles 1988 – 2006

The planning field is interested in the study of walking as a form of active transportation (Handy et al., 2002). Researchers have established a link between urban form and the modes of transport people choose (Cervero et al., 1997). Many years of single land use, residential suburbanization, as well as single-use office parks, has led to unprecedented traffic congestion and automobile dependence. Early work by Cervero (1988, p.429) suggests that the suburbanization of jobs in the 1980s led to "fundamentally altered commuting patterns" that resulted in many employees driving their own vehicle to work. This is in contrast to traditional commuting patterns where employment was primarily located in the downtown core and employees had transport options (Cervero, 1988). Cervero (1988) contends that mixed-use development has beneficial impacts on travel behaviour and traffic conditions and that appropriate planning policies could be used as a tool to encourage mixed-use developments that would host a variety of shops, restaurants, entertainment venues, offices and residential units. Once established, this mixed-use core would create a rich mixture of activities that would be utilized around the clock and would decrease peak hour traffic congestion by allowing people to walk to their destinations of interest (Cervero, 1988).

Frank and Pivo (1994) built upon Cervero's notion that there is a link between urban form and transport choice, and conducted tests to determine the impacts of urban form (land use mix, population density, and employment density) on transport mode (e.g. single-occupant vehicle, transit, and walking) at the census tract scale. Frank and Pivo (1994) used empirical data from a variety of sources to develop and then test hypothetical relationships between urban form and transport modes. As illustrated in Figure 2.3 below, Frank and Pivo noted that both urban form factors, such as density and land use mix, as well as non-urban form factors, such as income, gender, age, and level of service have significant impacts on travel behaviour. Since urban form factors were the focus of their research, they controlled for non-urban form factors through the analysis of descriptive statistics, correlation, and regression (Frank and Pivo, 1994). Controlling for non-urban form factors allowed for comparisons to be made between trip makers with similar socio-economic attributes (Frank and Pivo, 1994).

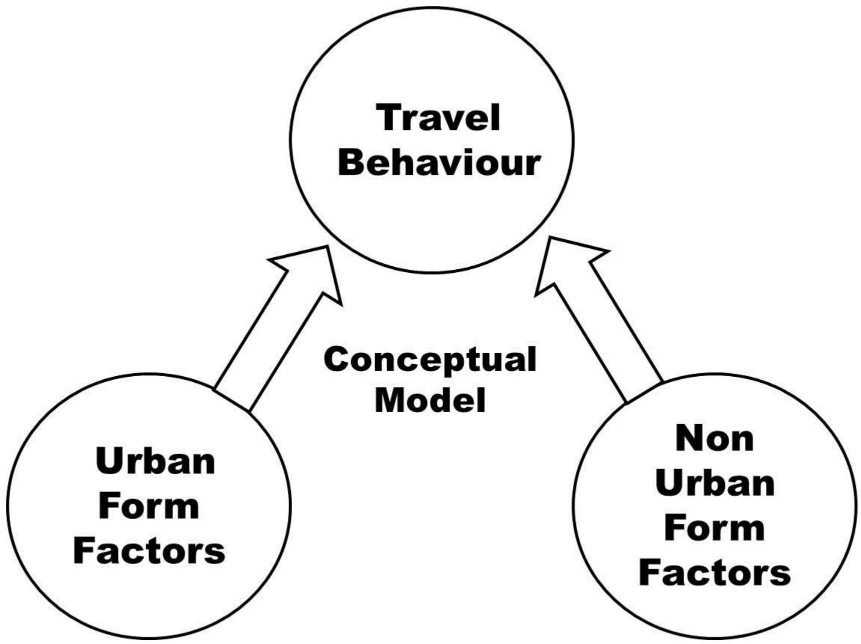


Figure 2.3. Relationship between travel behaviour and factors that affect it. (Source: Frank and Pivo, 1994, p.45)

The seminal conclusion of Frank and Pivo's research was that density and land use mix are related to choice of transportation mode. Specifically, the choice of both public transit and walking as a means of transport increases with an increase in density and land use mix; conversely, single-occupant vehicle as a transport choice decreases as density and land use increase (Frank and Pivo, 1994).

Research into the link between urban form and transportation choice continued with the work of Cervero and Kockelman (1997). Cervero and Kockelman expanded the research on urban form and introduced "design" as another element of urban form. In their article titled "Travel Demand and the 3Ds: Density, Diversity, and Design", they suggest that urban design characteristics can have an impact on travel demand (Cervero and Kockelman, 1997). Unique to their research, Cervero and Kockelman introduce the notion that not only does density and land use diversity impact travel choices, but neighbourhood design elements can have an impact on travel mode as well (Cervero and Kockelman, 1997). They went as far as providing specific examples of design elements that could enhance the pedestrian experience and would thus encourage walking. Specifically, they suggested pedestrian-oriented designs ought to have storefronts that are orientated to the sidewalk and any associated parking would be located behind the store, creating an inviting pedestrian storefront that would not hinder pedestrian access by forcing people to walk across vast parking lots in order to enter the store. Additionally, they suggested enhancing the pedestrian environment by including shade trees and benches along the sidewalk. It was intended that these design considerations would work to enhance the pedestrian *milieu* (Cervero and Kockelman, 1997). Their research concludes that density, land-use diversity, and pedestrian-oriented designs generally reduce car trip rates and also encourage non-auto travel (Cervero and Kockelman, 1997). Of particular interest to this research, Cervero and Kockelman contend that higher density, mixed land use, and pedestrian-friendly designs must be collectively incorporated into neighbourhood design in order to achieve the desired outcome of lower auto-dependence and higher rates of walking. For example, it is unlikely that a predominantly single-family neighbourhood with only sidewalks and attractive landscaping will prompt people to walk to shopping areas (Cervero and Kockelman, 1997).

Researchers began to examine the relationship between the built environment and the residents' physical activity at the neighbourhood scale (Saelens et al., 2003; Frank et al., 2005). Saelens et al. (2003) was the first to document the correlation between neighbourhood design and physical activity. The study examined residents' levels of physical activity for two neighbourhoods with differing degrees of walkability. Residents of the two neighbourhoods reported their physical activity by self-reported surveys and accelerometers. Both methods of physical activity tracking were used for verification purposes. The seminal conclusion was that residents of high-walkability neighbourhoods (i.e. neighbourhoods that had higher residential density, land use mix, street connectivity, aesthetics, and safety) reported higher levels/amounts of physical activity than did residents of lower walkability neighbourhoods (Saelens et al., 2003).

Frank et al. (2005) built on the research by Saelens et al. and introduced objective measures of both the built environment and physical activity, rather than the subjective measures employed by Saelens et al. (2003). The researchers created an Index of Walkability that could be used as a way of objectively measuring urban form and could also be used as a predictor of walking behaviour. Three neighbourhood design variables were selected as inputs to the index based on their independent relationships with walking. Frank et al. (2005) define the three variables as follows:

- Net Residential Density, which is defined as the number of residential units per unit of residential area. Areas with residential density greater than six per acre (i.e., 15/ha) are considered as more walkable.

- Street Connectivity, which is measured as the number of intersections per square kilometer. More intersections result in more direct walking route to destinations. Areas with greater than 30 intersections per square kilometer are considered more walkable.
- Land-use Mix, which can be expressed as the evenness of distribution of square footage of residential, commercial, and office space. This is measured by an entropy index.

Using these three variables Frank et al. developed a formula to create a walkability index as follows:

Walkability Index = (z-score of land-use mix)

+ (z-score of net residential density)

+ (z-score of intersection density)

Accelerometers were then provided to 357 adults and data were collected over a two day period. Neighbourhood walkability could then be related to physical activity. Their results concluded that neighbourhoods with a high degree of walkability were positively (though weakly) related with moderate amounts of daily physical activity ($R^2 = 0.107$) (Frank et al., 2005). This article is of particular importance and interest to this study as a walkability index was created with the same formula derived by Frank et al. (2005).

Other researchers have provided alternative approaches to using composite walkability indices that employ highly inter-correlated variables that are often difficult to gather, in order to measure the built environment and thus the “walkability” of neighbourhoods

(Lee et al., 2006). Lee and Moudon (2006), based on the 3Ds (density, diversity, and design) framework developed by Cervero and Kockelman (1997), created a notion of 3Ds + R (route). The addition of "route" to Cervero and Kockelman's framework recognized that distance measures between home and walking destinations were also important indicators of walkability. Their research focused on determining built environment variables that were correlates of walking. The study employed 900 built environment variables in a multiple regression analysis to identify those with significant effects on walking activity. The "shotgun" approach to the inclusion of independent variables helped to identify and understand broad groupings of walking determinants. In general, the broad groupings of determinants fell under the 3D's + R categories of destination, distance, density, and route (Lee et al., 2006). The research concluded that many previously proposed walking determinants were not statistically significant, such as width of road, traffic volume, and the presence of parks and fitness centers (Lee et al., 2006). Additionally, by grouping walking determinants under the 3D's+R formula, future data collection efforts could be better directed. The researchers admit that this study is still subject to the shortfall of relying only on a telephone-based self-reported survey rather than a quantitative method to determine the amount of walking (Lee et al., 2006).

An often cited criticism of walking behaviour research is the lack of objective walking data (Handy et al., 2002; Hoehner et al., 2005; Kang et al., 2013). In the absence of objective walking data, many studies utilize self-reported walking data that were collected via telephone surveys or questionnaires (Berke et al., 2007; Owen et al., 2007; Adams et al., 2011; Mayne et al., 2013). Telephone surveys or questionnaires have the benefit of being able to survey a larger sample with fewer expenses compared to objective reporting methods that may use an accelerometer or GPS. These self-reported methods of data collection have two common accuracy issues. Firstly, respondents of self-reported walking surveys or questionnaires tend to over-estimate their reported physical activity, both in terms of duration and frequency (Dewulf et al., 2012; Kang et al., 2013; Van Holle et al. 2015). Dewulf et al. (2012) explain that the volume of over estimation is not uniform across respondents and that those with less physical activity tend to report an even greater over-estimation of physical activity compared to those who may exercise more. The second concern with using self-reported data is that shorter walking trips tend to be under-reported (Kang et al., 2013). While researchers recommend using objective data when available, it is recognized that objective reporting of physical activity is not without its limitations. Objective walking data are collected, generally, by using accelerometers or GPS. While accelerometers can accurately calculate walking activity, location information or purpose of the walking trip is not as easily obtainable. The use of GPS to calculate walking activity provides location context as well as distance and speed, but it also has constraints. GPS datasets often lack data completeness due to lost satellite signals or otherwise erroneous data (Kang et al., 2013). Utilizing the combination of self-report travel diary, accelerometer, and GPS data may result in a more complete and accurate assessment of walking activity (Kang et al., 2013). The Halifax STAR dataset used in this study utilized both objective and self-reported methods of calculating physical activity by using a combination of time diary, GPS, and verified telephone recall survey.

Many studies make the assumption that walking trips begin or end at the respondents' home. Therefore, when examining the built environment in relation to walking behaviour, it is common for researchers to define the built environment "neighbourhood" as a set distance or buffer, around the home location of the respondent (Carlson et al., 2015; Frank et al., 2005; Grasser et al., 2013). Brownson et al. (2009) completed a review article that examined 38 studies in terms of how the built environment was measured. They determined that 25 of the 38 studies, or 65%, were based on a neighbourhood buffer surrounding the home location. Previous research has indicated that examining only the neighbourhood surrounding the home location may greatly underestimate the amount of walking, particularly for walking trips associated with active transportation (Spinney et al., 2012; Millward et al., 2013). Additional research has supported this notion, suggesting that walking is the most dominant form of reported physical activity (within the USA) and takes place both at home and from the work place (Williams et al., 2008). Researchers have also delineated the geographic bounds of the built environment by other methods, such as by both county (Doyle et al., 2006) and by commuting route (Rodriguez et al., 2004). Studies that delineate the built environment by buffer, county, or commuting route are often concerned with the individual respondent and the environment they inhabit. This research considers each census tract (CT) as a "case" rather than using the individual respondent. The concern is where actual walking is taking place, irrespective of the respondents' home or work location. The aggregation of walking distance by CT facilitates the correlation of walking activity with the built environment of the neighbourhood where the walking occurs.

It has been well documented that the built environment encourages the use of automobiles and thus discourages walking activity (Adams et al., 2011; Frank et al., 2004; Glazier et al., 2014). Walkability research has typically divided walking activity into two categories, recreational walking and active transportation (AT), with more research focusing on recreational (Spinney et al., 2012). However, Ulmer et al. (2014) contend that built environment features affect walking behaviour for both AT and recreational purposes. Recent research has suggested a gap in the literature that examines walking for recreation and utilitarian purposes as an aggregate (Hajna et al., 2015). Hajna et al., (2015) have completed a systematic review and meta-analysis of walkability literature that investigated the association of neighbourhood walkability and walking behaviour. The researchers concluded that the collective results of the examined studies support the notion that high neighbourhood walkability is associated with higher levels of walking behaviour. Hajna et al. (2015, p.2) explain that while active transport and recreational walking, separately, have been well studied, “our understanding of the association of walkability and total walking is limited”. The goal of this study is to improve neighbourhood design in an effort to increase walking activity, regardless of purpose.

This literature review has identified the seminal work that continues to be actively cited in most of the recent papers. While a substantial amount of walkability research has been conducted over the past ten years, this literature review focused on three areas that are uniquely considered in this dissertation: 1) the use of both objective GPS and self-report time diaries to record walking activity; 2) the use of census tracts as the case instead of the individual respondent; 3) the examination of walking as an aggregate without distinguishing between specific purposes.

CHAPTER 3

Methods

3.1 Study Area

As described in Chapter 1, Halifax Regional Municipality, Nova Scotia, Canada was chosen as the study area. Halifax is the capital city of Nova Scotia and the largest urban centre in the province, consisting of approximately 400,000 people. Halifax is a mixture of urban, suburban, and rural development. The Halifax Space-Time Activity Research (STAR) dataset was collected during 2007-08 for all urban and suburban areas in HRM, and most parts of the commuter-shed. It provided an opportunity to study neighbourhood design and walkability in a medium sized city.

The scale of this study is at the census tract (CT) level. Census tracts are defined by Statistics Canada (2006) as small stable areas that have a population range of approximately 2,500 – 8,000 people with an intended average population of 4,000 people. Census tracts are divided by physical attributes of the built environment such as roads and railway lines. They are intended to be relatively homogenous in terms of socio-economic characteristics and generally represent defined neighbourhoods (Booth et al., 2005). This study includes all 87 census tracts within Halifax (Figure 3.1).

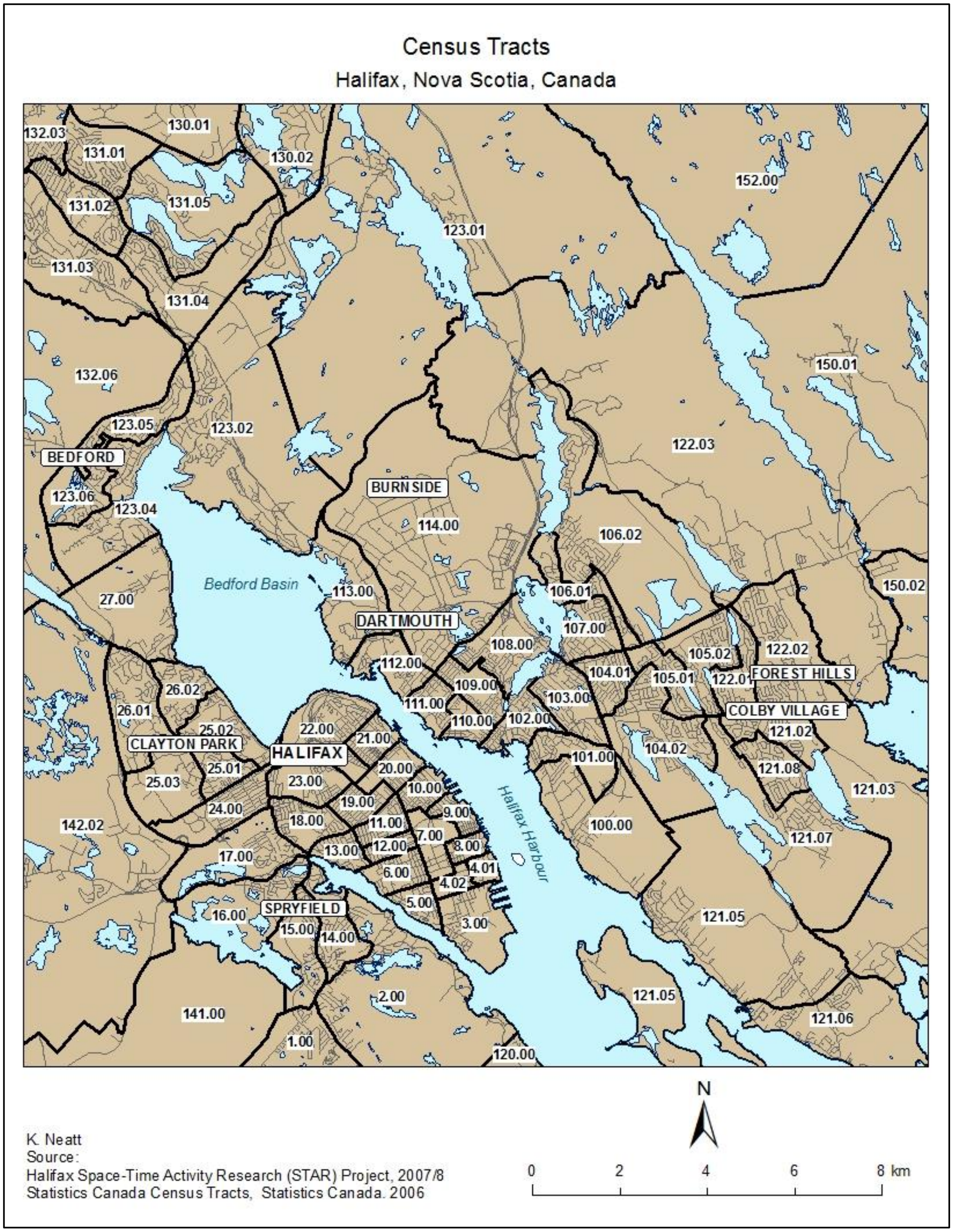


Figure 3.1. Census Tracts, Halifax, Nova Scotia

The modifiable areal unit problem (MAUP) is inherent in spatial analysis of data that are aggregated into zones (Jelinski, 1996). Coined by Openshaw and Taylor (1979), MAUP identifies the issues that result when analysing data that are based on specific, yet relatively arbitrary boundaries. The issues stem from the fact that different results may be observed if boundaries or scales of the study are varied (also see Fotheringham and Rogerson, 2009). There does not seem to be one clear approach to address MAUP. Clarke et al. (2014) suggest selecting a boundary or scale that is appropriate for the problem being investigated. Ultimately, the intention of this research is to provide insights into characteristics that influence walking behaviour in an effort to improve neighbourhood design. Therefore, following Clark's notion, a census tract scale is an appropriate scale of study to further this goal of improved neighbourhood design, as municipal policy closely relates to neighbourhoods that are defined at the census tract scale. Fortunately, many neighbourhood characteristics of both socio-demographic and built environments are collected at the census tract scale, making data collection and analysis much more viable.

3.2 Data Sources

Walking data were collected as part of the Halifax STAR Project, a joint project between Saint Mary's University and Halifax Regional Municipality in 2007 - 2008. The Halifax STAR project is a unique survey that collected information from households regarding travel activity and time-use. The STAR survey distributed GPS data loggers (HP iPAQ) to 1,971 randomly-selected primary respondents of survey households for a 48-hour collection period. These geo-referenced data were used to verify detailed time diaries, through telephone prompted-recall from 1,971 primary respondents. The GPS data loggers recorded positions at a resolution of three recordings every two seconds and had a spatial accuracy of sub-ten meters, with many positions having accuracy of within three meters (TURP, 2008, p.10).

Of the 1,971 primary respondents, 1,189 recorded 5,005 walking episodes (trips). These episodes, which are single acts "engaged in by an individual at a specified place and time under certain conditions" (Harvey, 1990), include walking for both active transportation as well as recreation purposes. Their locations were fully geo-referenced. The STAR survey classified activities based primarily on the 2005 General Social Survey of Canada (GSS) Time Use Survey (Statistics Canada, 2006). Although the STAR data include information on trip purpose, the present study does not distinguish between active transportation and recreation walking purposes. It was felt that neighbourhood design ought to encourage walking for both purposes and that within the confines of this research, it was more appropriate to investigate all walking behaviour rather than focusing on the specific purpose.

Built environment data, including census tract boundaries, waterbodies and roads (2006), were obtained from Statistics Canada. Using data provided by HRM, several built environment variables were generated and subsequently provided by Dr. Darren Scott, who is at McMaster University. Specifically, Dr. Scott provided the following built environment variables: intersection density, areas for each of the six land uses (retail, commercial, industrial, office, institutional and parkland) and retail lot coverage ratio. These variables are explained further in subsection 3.5 (independent variables).

Two census tract-level socio-demographic variables were used in this study: numbers of residents by age cohorts; and, household income. These variables are available in Statistics Canada's (2006) Public Use Microdata Files that were accessed through the Statistics Canada Data Liberation Initiative. Further explanation can be found in subsection 3.5.

3.3 Data Cleaning

Data cleaning or "weeding" refers to the removal of unwarranted or inaccurate information in the dataset. The raw walking data file, consisting of 5,005 walking trips and 781,205 individual GPS points (cases), required weeding. Weeding was a four-step process to remove redundant or inaccurate points. The raw data file (.csv) was imported into IBM SPSS Statistics version 21. The first weeding step was to delete all GPS points that had coordinates based on fewer than six satellites; this resulted in the removal of 2,214 points. HDOP is one quantifiable representation of GPS accuracy. A high HDOP value equals lower GPS accuracy (Wagner, 2011). The second weeding step was to eliminate all GPS points that had a horizontal dilution of precision (HDOP) of greater than or equal to eight. A figure of less than eight was selected as an appropriate cut-off value based on industry standards which suggest that GPS positions with a HDOP exceeding eight are likely invalid**;** this step resulted in the removal of 47,968 points. The third weeding step was to exclude all points with a speed of greater than 14 km/h. Fourteen kilometers per hour was chosen as the maximum speed based on an assumption that most people are not able to run faster than that speed; therefore it would eliminate any errors the GPS may have inadvertently logged as walking (e.g. interpolation between missing points due to signal issues). This led to a reduction of 7,613 points. The final weeding step was to remove all points with a minimum speed threshold of zero km/h. Many cases had a speed equal to zero due to inactivity while the GPS unit continued to take 3 recordings every two seconds. These readings were redundant to the entire dataset; 621,506 points were removed from the dataset. As a result of the four-step weeding process, 159,699 GPS points remained in the dataset.

Additional data cleaning was conducted once the data points were imported into ArcGIS version 10. A .csv file was created in SPSS and imported into ArcGIS, enabling the data to be mapped based on the geographic coordinates for each GPS point. Once the points were imported into ArcGIS, they were converted into continuous lines, using the *Point to Line* command, based on the unique walking event ID attribute. Walking track lines were then broken at each census tract boundary to facilitate the future aggregation of total line distance, which would represent total walking distance per census tract. A manual, judgement-based, editing process was then performed to improve accuracy. Three general scenarios were encountered when performing manual, judgement-based editing of walking tracks: 1) deficient data yet predictable routes; 2) irrational walking routes; 3) illogical yet verifiable walking routes.

In this study it was preferable that walking tracks were edited rather than deleted. Figure 3.2 provides an example of where there were sufficient GPS data to accurately depict the walking track at the beginning and the end of the journey, but insufficient data within the middle section. A manual procedure was performed to create additional vertices and align the walking track with the road system. The result represents a more logical and accurate walking route that is supported with GPS accuracy at each end of the route.

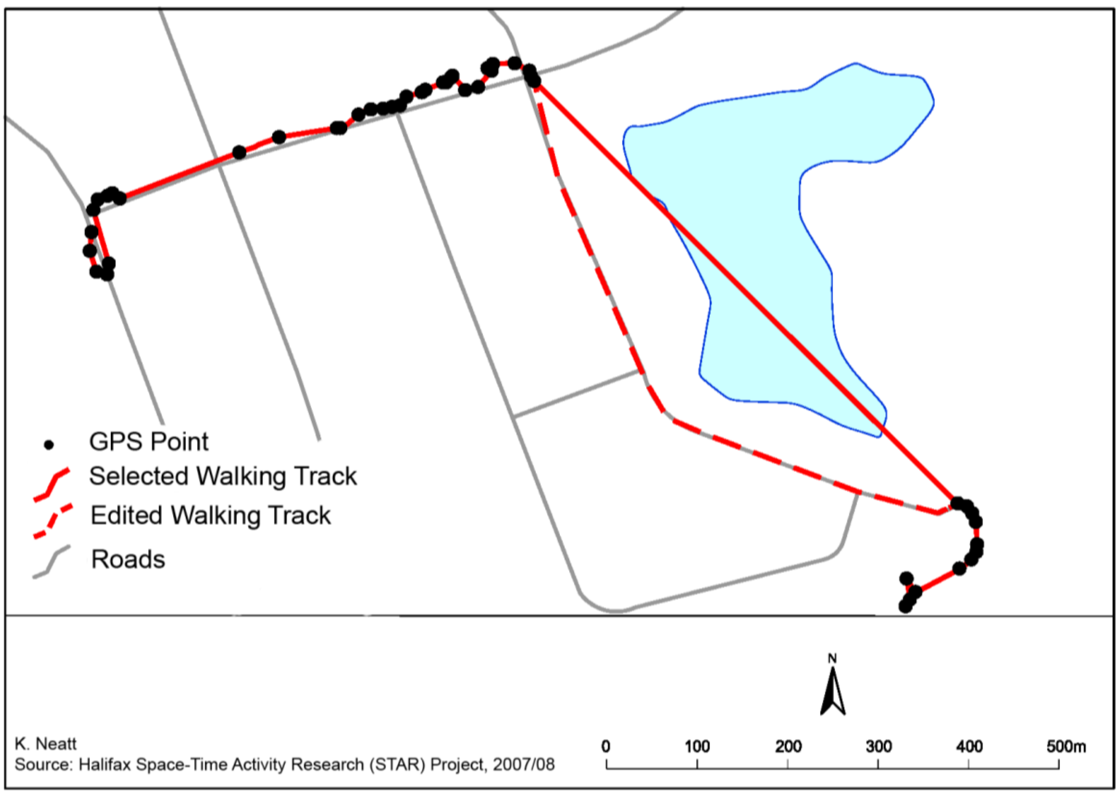


Figure 3.2. Revised Walking Route Example

Figure 3.3 provides a second example of an illogical walking track. Figure 3.3 depicts a walking track crossing water and properties in a very straight line. Although technically possible in the winter, it is unlikely that the walking track would be that straight and not follow the general road patterns once reaching land. There was insufficient GPS data to suggest a reasonable alternative route; therefore this walking track was deleted.

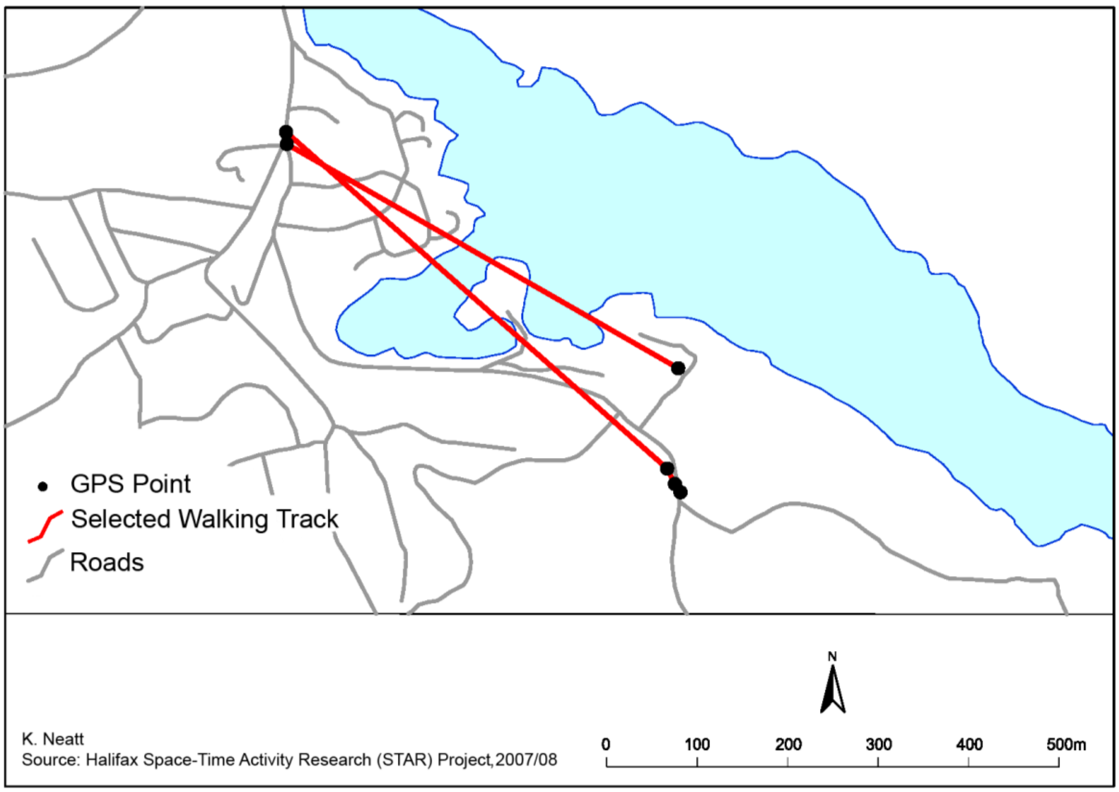


Figure 3.3. Example of a walking track error.

Walking tracks were not simply deleted on the basis of the road network. Tracks were also visually compared with air photos to determine if the walking track is reasonable. One such example is illustrated in Figure 3.4, which depicts a walking track extending out into a harbour. When the corresponding location was visible on the air photo, it was determined that the walking track was on a trail (a former rail causeway). Further validation of the walking track was based on the continuous number of GPS points that were displayed when the track was selected. This walking track remained in the dataset.

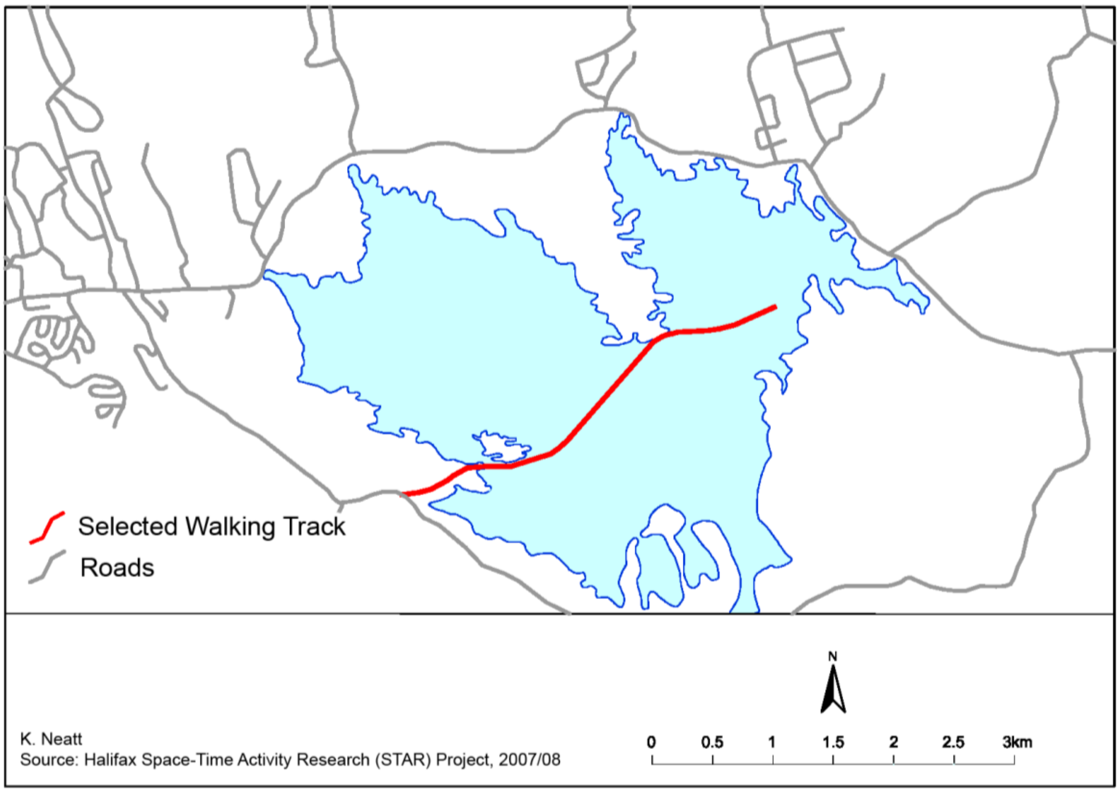


Figure 3.4. Verifiable Walking Route

A single ArcGIS file was created that contained the walking tracks, roads, water bodies, and census tracts. Each walking track was then associated with the particular census tract within which it was located. An aggregate measure of walked distance per census tract was calculated by summing the total lengths of all walking tracks per individual census tract.

Due to variations in area and population of census tracts, three walking densities were developed to measure walked distance per census tract in formats that facilitated comparison between census tracts. The first walking density variable measured walked distance per capita for each census tract. This variable, “walked distance by population” (W/P), was calculated by dividing the total walked distance per census tract by its total resident population. The second walking density variable measured the walked distance per meter of road in each census tract. This variable, “walked distance by road” (W/R), was calculated by dividing the aggregate walked distance per census tract by the total length of all paved roads in each census tract. The third walking density, “walked distance by developed area” (W/DA) was calculated by dividing the total walked distance per census tract by the total developed area (in square metres) of each census tract. The developed area was derived from the aggregate area of six land uses (residential, commercial, office, park and recreation, institutional and industrial).

The three walking densities (W/P, W/R and W/DA) all reflect different walking attributes. W/P is a measure of the propensity for people to walk; while, W/R and W/DA measure the use of the built environment in terms of roads and built environment area. It would be expected that changes in variables that typically influence walking may not necessarily be correlated equally with all three densities. For example, if a change occurred in residential dwelling density it would be expected to have an impact on both the amount of walking per meter of road and amount of walking per developed area, assuming that additional residents would result in increased walking, *ceteris paribus*. One would expect that an increase in dwelling density would have much less influence on the W/P, since dwelling density and population are directly related. There may, however, be some influence on residents' walking propensities, since higher residential densities typically imply a closer arrangement of walkable destinations.

Regarding the W/P measure, one should bear in mind that not all walking within a census tract will be performed by residents of that neighbourhood, and indeed perhaps only a small portion of walking will be by residents (for example, in downtown areas, or in areas with many retail, office, or institutional destinations). For this reason, there will not necessarily be a strong correlation between resident population and aggregate walking behaviour.

3.5 Independent Variables

As explained in the literature review, two walkability index formulas are used in this study, a three-variable index and a four-variable index. The three-variable walkability index (WI-3) employs dwelling density, intersection density, and entropy (land-use mix), while the four-variable walkability index (WI-4) employs a fourth variable, retail lot coverage ratio. Although the components of the walkability indices vary, the basic formula remains as the summation of the Z-scores for each component in each census tract.

WI-3 = dwelling density (Z-score) + intersection density (Z-score) + entropy (Z-score)

WI-4 = dwelling density (Z-score) + intersection density (Z-score) + entropy (Z-score) + retail lot coverage ratio (Z-score).

The individual walkability components are defined as follows:

- Dwelling density (Z-score) was calculated by dividing the total number of dwellings per census tract by the total amount of developed area within the census tract. The figure was then normalized as a Z-score. The number of dwellings was derived from the 2006 Statistics Canada dataset. The developed area was derived from the aggregate area of six land uses (residential, commercial, office, park and recreation, institutional and industrial). The land use dataset was developed as part of the STAR project.

- Intersection density (Z-score) was calculated by dividing the total number of intersections within each census tract by the census tract area. This figure was provided by Dr. Darren Scott from McMaster University. The figure was then normalised as a Z-score.

- Entropy (Z-score) is a measure of land-use mixture within a prescribed area. The literature lacks consistency in the formula used to calculate entropy. This research utilizes an entropy formula based on the work of Frank et al. (2006) as follows:

Entropy = −A/lnN

- A=(b1/a)*ln(b1/a) + (b2/a)*ln(b2/a) + (b3/a)*ln(b3/a) + (b4/a)*ln(b4/a) + (b5/a)*ln(b5/a) + (b6/a)*ln(b6/a)
- a = total square feet of land for each of six land uses within census tract

b1through b6 measure areas of land use for:

- b1= residential area
- b2= commercial area
- b3= institutional area
- b4= office area
- b5= park and recreation area
- b6= industrial area
- N= number of six land uses with area > 0.

- Retail lot coverage ratio (Z-score) was calculated by dividing the total retail building footprint area by total retail parcel area per census tract.

An expanded walkability index, including additional components that influence walking, would have been interesting to include in this study. Specifically, the inclusion of sidewalks as a built environment variable within the walkability index tends to provide valuable insights into walking behaviour (Booth et al., 2005). Sidewalk data were obtained from the HRM's GIS department. Unfortunately, however, it was discovered that the dataset was incomplete and vector lines depicting sidewalks were contiguous polygons that outlined the edge of the concrete sidewalk and included driveway cuts. The inclusion of driveway cuts meant that an accurate distance of sidewalks for each census tract was not possible, because the number of intersections and driveway cuts would have greatly skewed the aggregate distance. Ideally, sidewalks would have been represented in a similar pattern as roads; as single vector lines with a distance. As this study was not focusing specifically on sidewalks, it was felt that the effort to digitize the sidewalk coverage into single lines was not warranted and better left for specific studies relating to sidewalks in the future.

Two socio-demographic control variables were included in the study; age and income. Specifically, age was categorized into three cohorts: young adult (age 15 – 39), middle-aged (age 40 – 64) and older adult (age 65 plus). These figures were expressed as a proportion of the census tract population of each cohort by census tract. The second control variable was average (mean) household income, expressed in thousands of dollars, per census tract. These two control variables were included in the study as the literature suggests that both age and income tend to influence walking behavior (Berke et al., 2007; Sallis et al., 2009), and this is substantiated in a study using the STAR data (Spinney et al., 2012).

A third socio-demographic variable, a dissimilarity index, was investigated but not included in the study. The literature suggests that those living closer to employment locations have a greater propensity to select active transportation as a mode of travel (Cerin et al., 2007). The intention was to include a variable that would address this notion by providing a measure of *spatial mismatch* between employment and residential location. A reasonable effort was made to collect the information required to calculate this measure. While Statistics Canada collects data on the number of jobs, unfortunately the data are only published for census metropolitan areas and not at the required census tract scale.

The developed land within the census tract was categorized into six land uses: residential, commercial, institutional, park / recreation, office, and industrial. The total area for each land use was available through the STAR dataset. These six land-use variables were included in the study as it was felt that entropy alone was not sufficient to capture the impact of land use on walking behaviour. Although the literature supports the notion that a high degree of entropy creates a greater propensity to walk (Cervero et al., 1997), it was felt that individual, specific land uses such as park / recreation and commercial lands may attract more walking than other land uses. The proportions of each of the six land uses per census tract (CT) were also included as independent variables. A figure was also calculated for each land use and expressed as a percentage of developed land per census tract. As an example, the formula for percentage of residential land in a CT is provided as follows.

$$\% \, Residential = \frac{Area \; of \; residential \; land}{Total \; area \; of \; developed \; land}$$

Where: Total area of developed land = residential area + commercial area + institutional area + park / recreation area + office area + industrial area, per CT.

3.6 Statistical Analysis in SPSS

The final datasets used for analysis contained the following set of variables, which were calculated for each CT.

Dependent Variables:

- W/P (walked distance by population per census tract)
- W/R (walked distance by aggregate road distance per census tract)
- W/DA (walked distance by developed area per census tract)

Independent Variables:

- Three variable walkability index (Z-score)
- Four-variable walkability index (Z-score)
- Dwelling density (Z-score)
- Intersection density (Z-score)

- Entropy (Z-score)
- Retail lot coverage ratio (Z-score)

Control Variables:

- % Young adults
- % Middle-aged
- % Older adults
- Average (mean) household income
- % residential land
- % commercial land
- % park and recreation land
- % institutional land
- % office land
- % industrial land

The first task was to create scatter plots that depict the relationship between each dependent variable (W/P, W/R and W/DA) against each independent variable. Six scatter plots were created for each dependent variable. The creation of scatter plots enables confirmation of the type of relationship that exists between the variables. Additionally, the scatter plots enable the identification of any outliers in the dataset.

The second task was to create a two-tailed, bivariate, Pearson correlation table. A correlation table measures the degree of linear relationship between two variables. The correlation value, ranging between -1 and +1, expressed by the *r* value, shows whether the linear relationship between the two variables is positive, negative, or non-existent. The correlation output also includes the significance or *p*-value, which is the likelihood that an equal or greater correlation value could occur by chance. This study conforms to generally-accepted statistical practice that states if $p \leq 0.05$ then the correlation is significant.

The final statistical step in this study was to create eighteen multiple regression models (six for each W/P, W/R and W/DA), individually testing the dependent variables against the various independent and control variables. The eighteen models are described as follows:

Model 1:

- Dependent Variable: W/P
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density

Model 2:

- Dependent Variable: W/P
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density, Z-Retail Lot Coverage Ratio

Model 3:

- Dependent Variable: W/P
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density, Z-Retail Lot Coverage Ratio
- Control Variables: Average Income, % Young Adult, % Middle-aged, % Older Adult

Model 4:

- Dependent Variable: W/P
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density, Z-Retail Lot Coverage Ratio
- Control Variables: Average Income, % Young Adult, % Middle-aged, % Older Adult, % Residential, % Commercial, % Parkland, % Institutional, % Industrial, % Office

Model 5:

- Dependent Variable: W/P
- Independent Variables: Walkability Index (4 variable)
- Control Variables: Average Income (in thousands), % Young Adult, % Middle-aged, % Older Adult

Model 6:

- Dependent Variable: W/P

- Independent Variables: Walkability Index (4 variable)
- Control Variables: Average Income (in thousands), % Young Adult, % Middle-aged, % Older Adult, % Residential, % Commercial, % Parkland, % Institutional, % Industrial, % Office

Model 7:

- Dependent Variable: W/R
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density

Model 8:

- Dependent Variable: W/R
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density, Z- Retail Lot Coverage Ratio

Model 9:

- Dependent Variable: W/R
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density, Z-Retail Lot Coverage Ratio
- Control Variables: Average Income (in thousands), % Young Adult, % Middle-aged, % Older Adult

Model 10:

- Dependent Variable: W/R

- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density, Z-Retail Lot Coverage Ratio
- Control Variables: Average Income, % Young Adult, % Middle-aged, % Older Adult, % Residential, % Commercial, % Parkland, % Institutional, % Industrial, % Office

Model 11:

- Dependent Variable: W/R
- Independent Variables: Walkability Index (4 variable)
- Control Variables: Average Income, % Young Adult, % Middle-aged, % Older Adult

Model 12:

- Dependent Variable: W/R
- Independent Variables: Walkability Index (4 variable)
- Control Variables: Average Income (in thousands), % Young Adult, % Middle-aged, % Older Adult, % Residential, % Commercial, % Parkland, % Institutional, % Industrial, % Office

Model 13:

- Dependent Variable: W/DA
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density

Model 14:

- Dependent Variable: W/DA
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density, Z-Retail Lot Coverage Ratio

Model 15:

- Dependent Variable: W/DA
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density, Z-Retail Lot Coverage Ratio
- Control Variables: Average Income, % Young Adult, % Middle-aged, % Older Adult

Model 16:

- Dependent Variable: W/DA
- Independent Variables: Z-Entropy, Z-Dwelling Density, Z-Intersection Density, Z-Retail Lot Coverage Ratio
- Control Variables: Average Income, % Young Adult, % Middle-aged, % Older Adult, % Residential, % Commercial, % Parkland, % Institutional, % Industrial, % Office

Model 17:

- Dependent Variable: W/DA
- Independent Variables: Walkability Index (4 variable)

- Control Variables: Average Income (in thousands), % Young Adult, % Middle-aged, % Older Adult

Model 18:

- Dependent Variable: W/DA
- Independent Variables: Walkability Index (4 variable)
- Control Variables: Average Income (in thousands), % Young Adult, % Middle-aged, % Older Adult, % Residential, % Commercial, % Parkland, % Institutional, % Industrial, % Office

In an effort to verify the robustness of the forward stepwise results, all 18 regression models were also completed using the Enter multiple regression method. Both procedures produced similar results, but only the stepwise results will be reported and discussed in detail.

CHAPTER 4

Spatial Analysis

Four maps were created to assist in the spatial analysis component of this research. Examining spatial patterns is important for gaining further insights into walking behaviour, and compliments the statistical analysis.

The locations of the individual GPS tracks for each walking episode are illustrated in Figure 4.1. Individual walking tracks have been collectively displayed, in order to illustrate the spatial distribution of walking tracks across the study area. In general, walking more frequently occurs in urban areas compared to suburban areas. Specifically, walking occurs on the Halifax Peninsula within the central business district (CBD), which includes lands along the waterfront southeast of the Dockyard (Canadian naval yards) and lands surrounding the south and east sides of Citadel Hill. The CBD of Dartmouth, surrounding Alderney Gate and the ferry terminal, also exhibits multiple walking tracks.

Upon review of the walking locations outside the two CBD's, several areas of clustered walking tracks become apparent. Specifically, walking tracks are clustered throughout the Halifax Peninsula in the West End abutting Quinpool Road, and in the South End, focused on Dalhousie University and the adjacent hospital district. In addition, walking episodes are clustered within Point Pleasant Park in the South End of Halifax. Although less pronounced than the clustered walking episodes located on the Halifax Peninsula, there are two additional discernible areas with clustered walking tracks worth noting. In

Dartmouth, clustered walking episodes are noticeable extending from Alderney Gate, northwards adjacent to Lake Micmac and Lake Banook, and within Shubie Park. This corridor is an attractive walking area that includes both recreation and commercial land uses. The second area encompasses recreation opportunities surrounding the Bedford waterfront (Mill Cove) and extends north toward the commercial district in Bedford (Sunnyside Mall). These two areas of clustered walking episodes suggest that both commercial and recreational land uses play a role in attracting walkers.

Equally important to examining concentrations of walking tracks, consideration was also given to where walking did not occur. Specifically, the inner-city areas with the least evidence of walking behaviour are in the North Ends of both Halifax and Dartmouth. These are areas of lower income, and may be less inviting for walkers due to both fear of crime and less attractive streetscapes (McCormack et al., 2009). Similarly, lower-income areas of Spryfield also exhibit little walking. Additionally, Burnside Business Park also exhibits very few walking episodes. This may seem surprising, given the amount of employment and access to public transit in the area. However, employment densities are very low in much of Burnside, and destinations tend to be far apart, making vehicular access the preferred mode. In addition, perhaps less attractive streetscapes or the absence of sidewalks contribute to less walking activity.

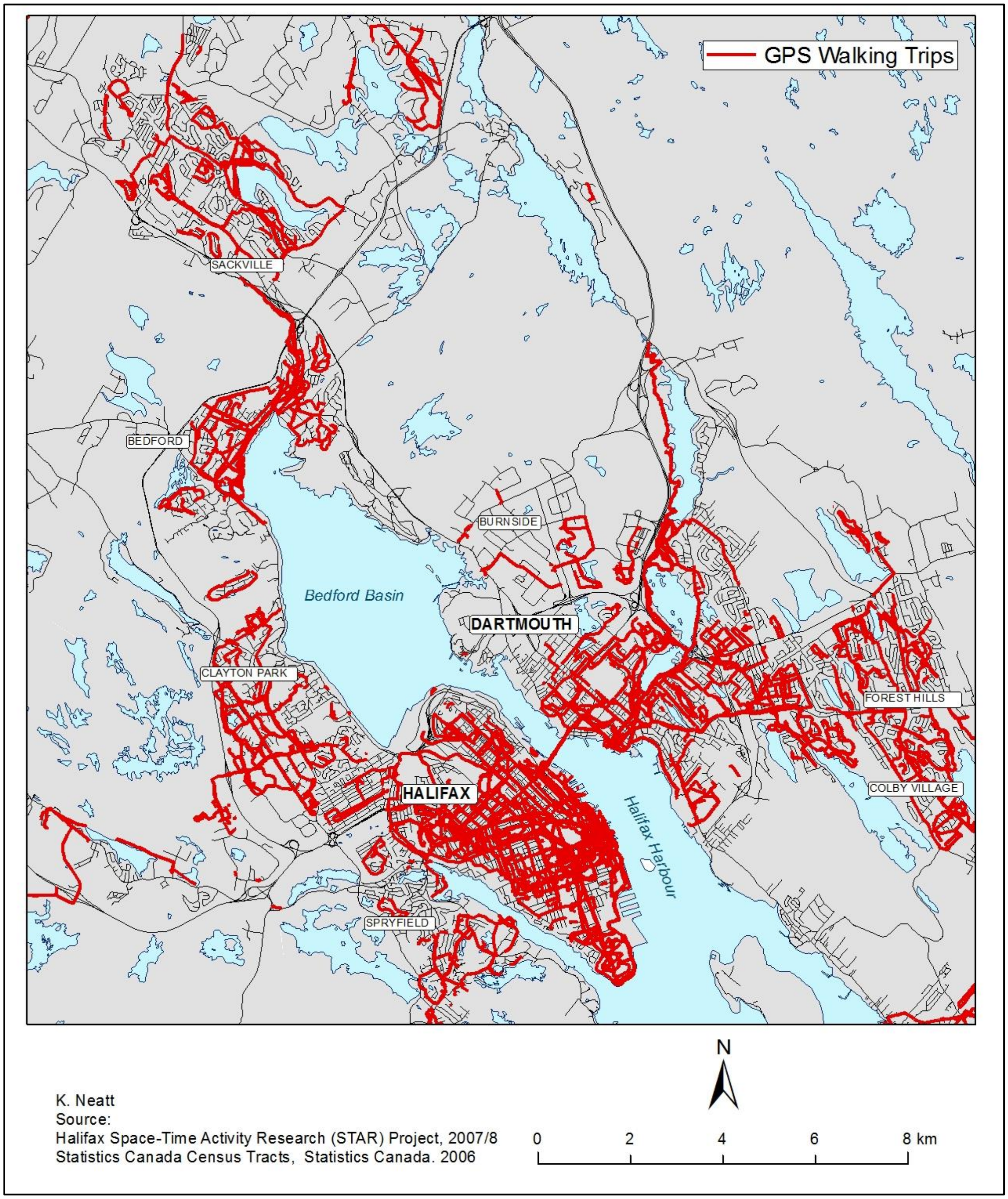


Figure 4.1. Walking Tracks, Halifax, Nova Scotia

One measure of walking density used in this study, which relates to the available street network, is the aggregate walked distance per meter of road by census tract (W/R). This measure of walking density is illustrated in Figure 4.2 and confirms most of the observations noted for Figure 4.1. The areas of highest walking density values are located on the Halifax Peninsula surrounding the Halifax Commons and extend to the waterfront (census tracts 7, 8, and 9). The high concentration of walking compared to road length in this area could be a result of the mixture of land uses in the CBD. Several dominant land uses are located in this area: recreation uses such as the public gardens, Citadel Hill, and the waterfront boardwalk; retail uses such as those located on Spring Garden Road; institutional uses such as Dalhousie University (Sexton campus) and the public library; employment uses such as the QEII and IWK hospitals and the Maritime Center office building; and, high density residential uses such as those located near Spring Garden Road.

A little farther to the south, the high walking density in census tract 3 is partially attributable to Point Pleasant Park, which is a popular destination for walking activity and no vehicular roads. This census tract also contains Saint Mary's University. Much of the rest of Peninsular Halifax has moderate values for walking density. Areas containing both commercial and residential land uses exhibit a greater amount of walking, particularly areas adjacent to Quinpool Road and census tract 18, which surrounds Halifax Shopping Centre.

The CBD in Dartmouth surrounding Alderney Gate (census tracts 102 and 110) and the tract surrounding Mic Mall Mall (census tract 108) have moderate to high walking densities. These areas encompass shopping, transit facilities, employment, and recreation land uses, as well as high density residential buildings. These areas with moderate to high walking density values coincide with the clustered walking patterns illustrated in Figure 4.1 described above.

Overall, suburban areas tend to exhibit lower densities of walking per road meter, but some have moderately high values. Areas of particularly high values include Clayton Park, Colby Village, Bedford, and the First Lake area of Sackville. Perhaps the moderate walking density values result from somewhat higher residential densities (most of these areas contain apartment complexes in addition to single-family housing), positive pedestrian infrastructure such as sidewalks and walking trails, and the location of commercial nodes along transportation corridors. In the case of First Lake, however, the attractive element is a large regional park.

Areas with low walking density have lower population density (typically lacking apartment housing) and lack both commercial areas and large parks. Perhaps co-incidentally, most are also areas of lower income: North End Halifax, Spryfield, North End Dartmouth, Woodside, and Eastern Passage all exhibit lower walking density values. This is somewhat surprising given that areas with lower income generally have high levels of public transit ridership, and therefore would have a greater number of AT walking trips to and from bus stops (Pucher et al., 2003).

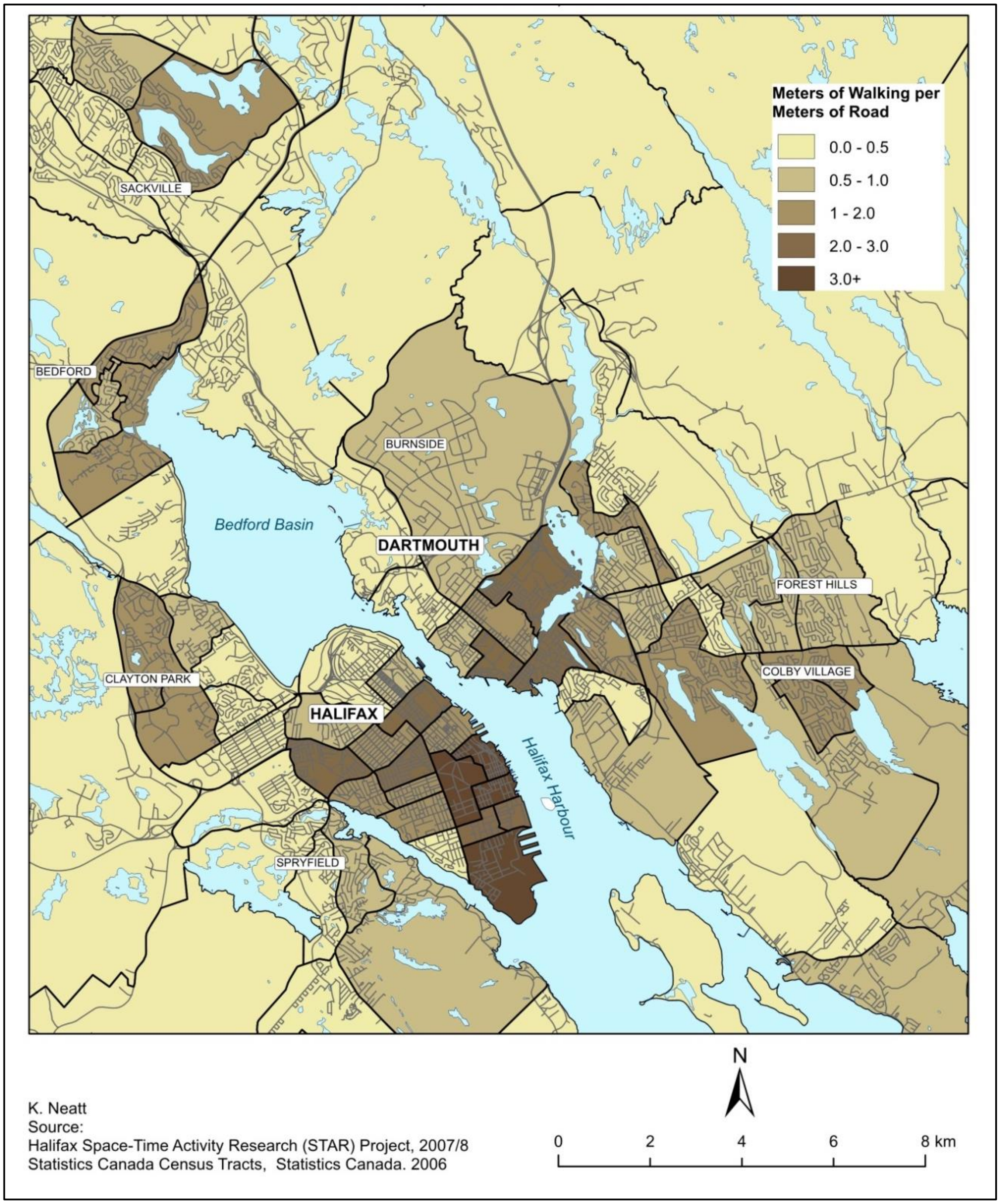


Figure 4.2. Aggregate Walked Distance per Meter of Road by Census Tract

Walking density is also measured using aggregate walked distance per census tract population by census tract area (W/P), which is illustrated in Figure 4.3. While W/P may be viewed as a measure of the propensity for CT residents to walk, one should bear in mind that much walking, particularly in central CT's, is by non-residents.

The overall pattern is similar to that for walked distance per road meter. For example, both peninsular Halifax and the CBD in Dartmouth have higher walking density values. These areas experience greater amounts of walking in relation to the number of residents. This supports the notion that people are walking in areas outside of their resident neighbourhood (Spinney et al., 2012). The suburban areas of Dartmouth, Forest Hills, Clayton Park, Bedford, and Sackville all exhibit low to moderate walking density values, except for census tracts with greater amounts of commercial land uses (particularly census tracts 104.01 and 108 in Dartmouth, and 123.05 in Bedford).

A third measure of walking density is the aggregate walked distance by developed area by census tract (W/DA), which measures walking in relation to the built environment and is illustrated in Figure 4.4. This measure of walking density exhibits similar patterns to the previous two densities. For example, both peninsular Halifax and the CBD of Dartmouth, exhibit high walking density values, while the suburban areas exhibit much lower values. One particular difference noticed when comparing W/DA to the previous two walking densities is that there appears to be a greater contrast between the urban and suburban areas. This is likely due to the mixture of land uses that attract walking. Specifically, walking is associated with the downtowns and traditional inner-city retail shopping areas, with the main university campuses, and to a lesser extent with some suburban commercial areas.

All three walking density maps (W/R, W/P, and W/DA) illustrate similar spatial patterns. High walking density values are evident in peninsular Halifax and the CBD of Dartmouth, while the suburbs are generally characterised as exhibiting low walking density values, except for areas with commercial land uses located along principal transportation routes

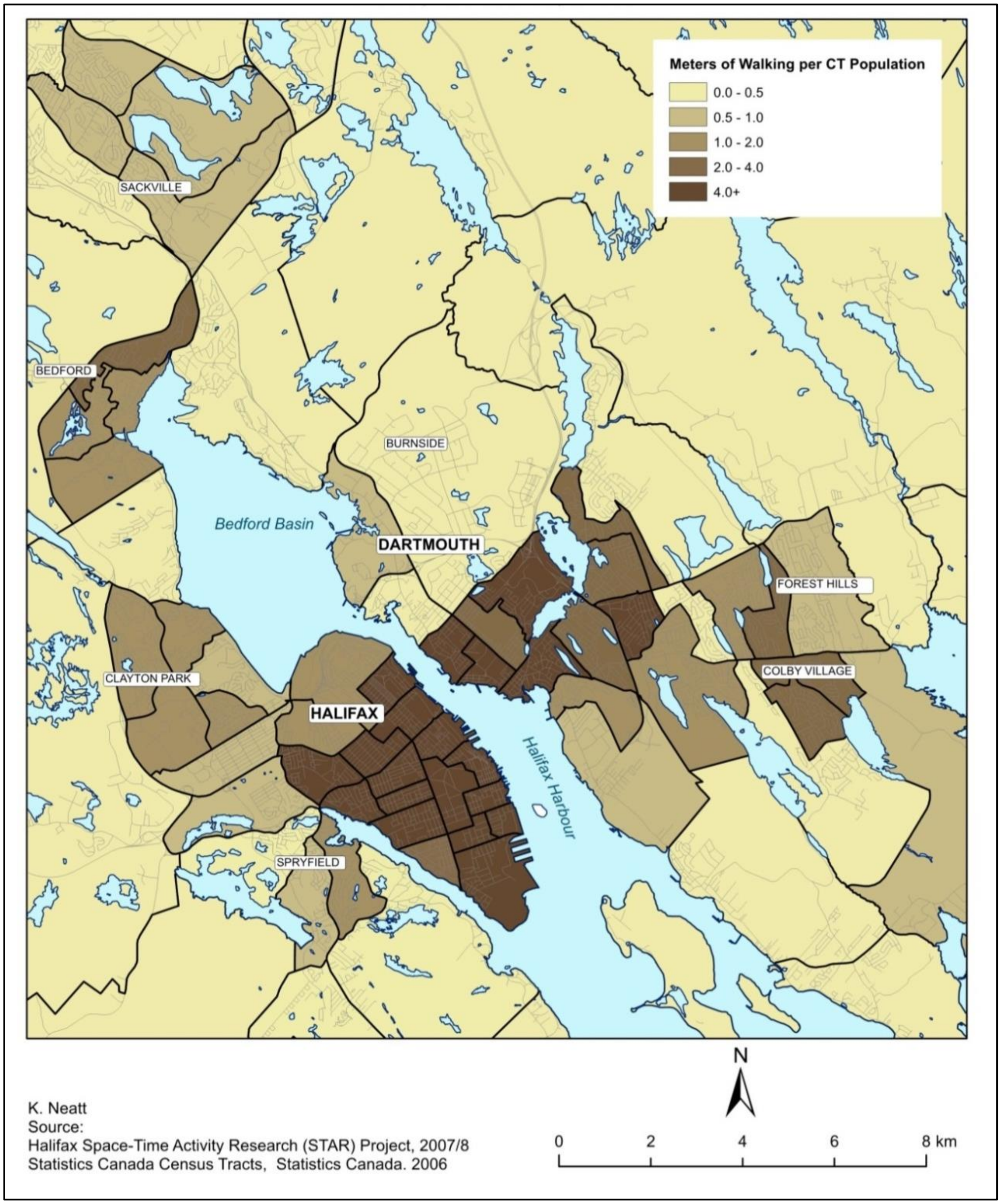


Figure 4.3. Aggregate Walked Distance per Census Tract Population by Census Tract Area

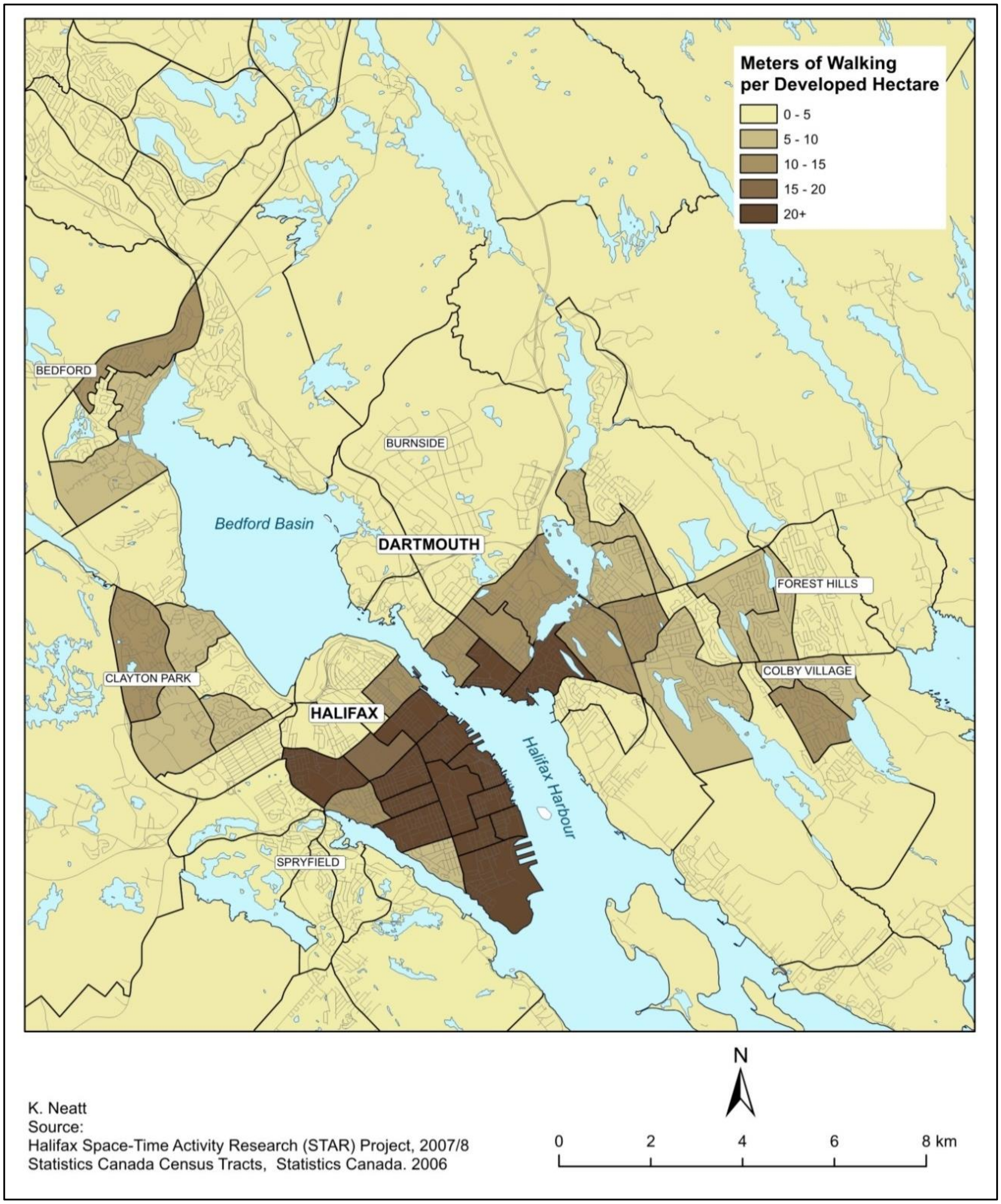


Figure 4.4. Aggregate Walked Distance by Developed Area by Census Tract

CHAPTER 5

Modeling: Correlations

A two-tailed, bivariate, Pearson Correlation analysis was used to assess the significance of any statistical linear relationships among the individual walking index variables, and the two composite walkability indices against the three walking densities – walked distance per person (W/P), walked distance by meters of road (W/R) and walked distance per developed area (W/DA).

		Total Walking Distance (Meters)	W/P	W/R	W/DA	Z-score Intersection Density	Z-score Entropy	Z-score Dwelling Density	Z-score Retail Lot Coverage Ratio	Z-score Walkability Index_3 Variable	Z-score Walkability Index_4 Variable
Total Walking Distance (Meters)	Correlation	1	**.851**	**.803**	**.700**	.289	.416	.192	.446	.360	.406
W/P	Correlation		1	**.904**	**.787**	.335	.391	.177	.489	.363	.420
W/R	Correlation			1	**.926**	.474	.447	.454	.617	.552	.602
W/DA	Correlation				1	.616	.458	.532	**.724**	.645	**.704**
Z-score Intersection Density	Correlation					1	.462	**.718**	**.743**	**.876**	**.883**
Z-score Entropy	Correlation						1	.418	.428	**.755**	.697
Z-score Dwelling Density	Correlation							1	**.709**	**.858**	**.859**
Z-score Retail Lot Coverage Ratio	Correlation								1	**.755****	**.870**
Z-score Walkability Index_3 Variable	Correlation									1	**.980**
Z-score Walkability Index_4 Variable	Correlation										1

Table 5.1. Correlations
All r-values exceeding 0.25 are significant at p=0.01

In the following discussion, Pearson correlations, r values, are described as three distinct strengths; “weak” values ranging from 0.0 to 0.39, “moderate” 0.4 to 0.69 and “strong” from 0.7 to 1.0. A weak relationship exists between the three ‘basic’ walkability index variables (dwelling density, intersection density and entropy) and W/P (Table 5.1). Dwelling density is the least correlated with W/P ($r = 0.177$) and is not significant ($p = 0.102$). The weak correlation between dwelling density and W/P is expected as W/P is based on walking per capita; it is unexpected that changes in the number of dwellings would be linearly associated with the amount of walking per capita. Figures 5.1, 5.2, and 5.3 indicate these weak positive linear associations between the three walkability variables (dwelling density, intersection density, and entropy) and W/P. The fourth walkability variable, retail lot coverage ratio, exhibits the strongest, yet still moderate, correlation to W/P ($r = 0.489$), and Figure 5.4 illustrates the positive relationship.

There are, however, some notable outliers in Figures 5.1 to 5.4. The census tracts identified as 3, 7, 8 and 9 are located in the inner city area of the Halifax Peninsula (Figure 3.1). Point Pleasant Park, located in census tract 3, is a prominent walking attraction, while census tracts 7, 8, and 9 include a mixture of walking attractions such as, the business district of Spring Garden Road, which has the highest pedestrian volumes east of Montreal (Terrain, 2009), three major hospitals, and popular park and recreation facilities that include the Halifax Commons and the waterfront.

The values depicted in Figures 5.1 through 5.4 are Z-scores, which facilitates comparison among the different variables. Z-scores represent the number of standard deviations away from the mean for a particular observation and can be positive or negative. Therefore, Z-scores can be used to compare variables that are initially in differing units: for example, dwelling density, which is expressed as number of units per unit of area, and retail lot coverage ratio, which is a ratio of the total retail building footprint area by total retail parcel area per census tract.

The outliers, identified as census tracts 3, 4.01, 4.02, 7, 8, 9, and 11, are all located within the core of the city on the Halifax Peninsula. The outlier census tracts are areas of high walking volumes and, proportionately, have much less road surface than many of the other census tracts. Retail lot coverage ratio exhibits the strongest correlation of the four walkability index variables with W/R ($r = 0.617$, $p = 0.000$). The associated scatterplot (Figure 5.8) illustrates a positive linear relationship, and it also appears to have the largest effect on Y estimates, as determined by the larger coefficient ($b = 0.89$).

The three basic walkability index (WI) variables (dwelling density, intersection density, and entropy) all have moderate correlations with W/DA, with r values ranging from 0.532 to 0.616 (Table 5.1). Notably, the fourth walkability variable (retail lot coverage ratio) has a much stronger correlation with W/DA ($r = 0.724$) than the other three walkability variables. Figures 5.9 to 5.12 depict the scatter-plots for each individual variable of the walkability index against W/DA.

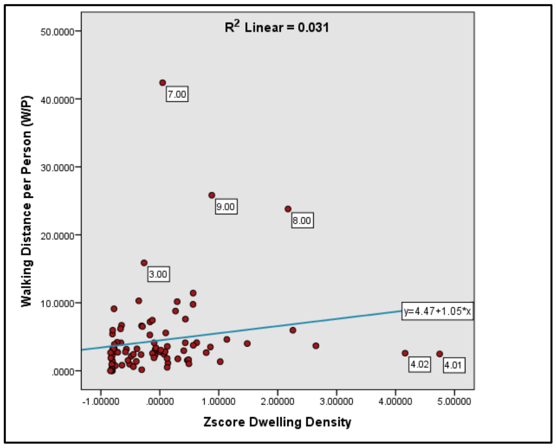


Figure 5.1. W/P: Z-score Dwelling Density

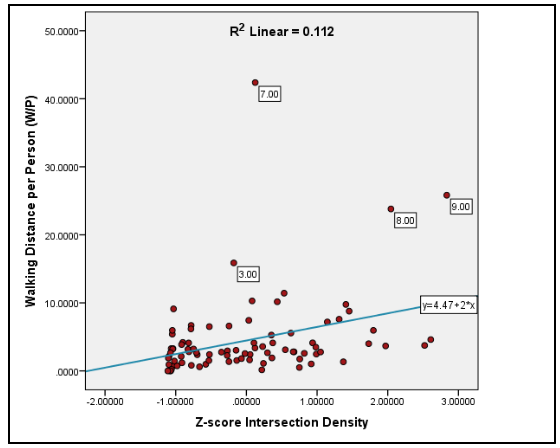


Figure 5.2. W/P: Z-score Intersection Density

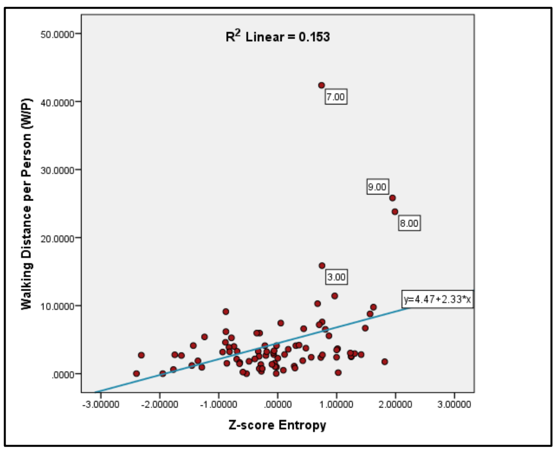


Figure 5.3.W/P: Z-score Entropy

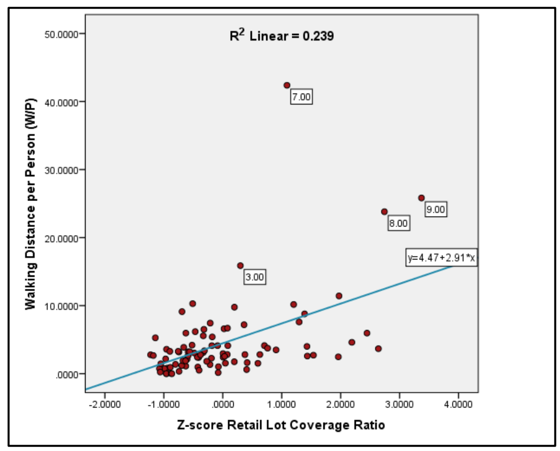


Figure 5.4. W/P: Z-score Retail Lot Coverage Ratio

Figure 5.5. W/R: Z-score Dwelling Density

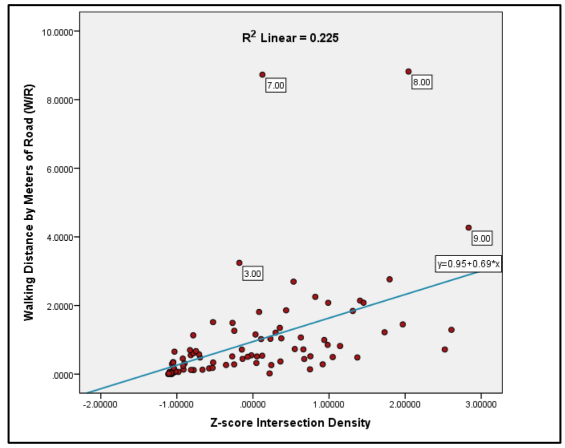


Figure 5.6. W/R: Z-score Intersection Density

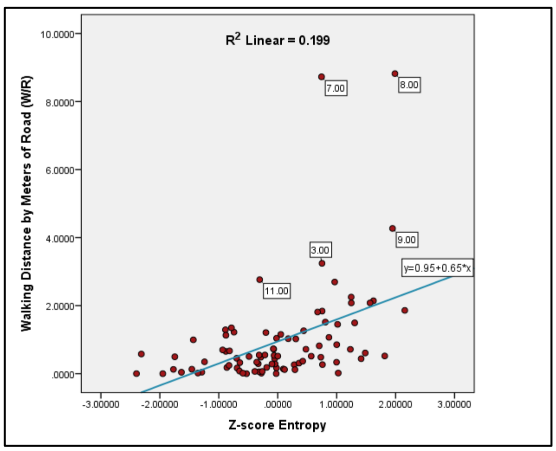


Figure 5.7. W/R: Z-score Entropy

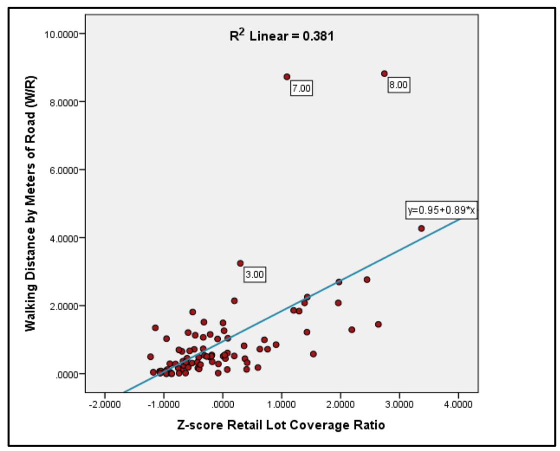
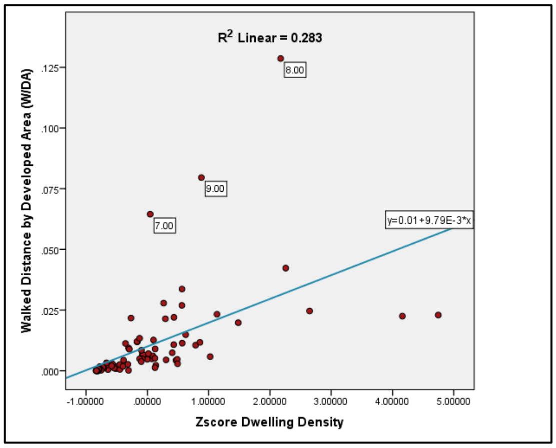


Figure 5.8. W/R: Z-score Retail Lot Coverage Ratio

All four variables have a positive, linear relationship with W/DA. Additionally, all four scatter-plots have outliers (census tracts 7, 8 and 9) similar to those represented in the previous scatter-plots relating to both W/P and W/R.



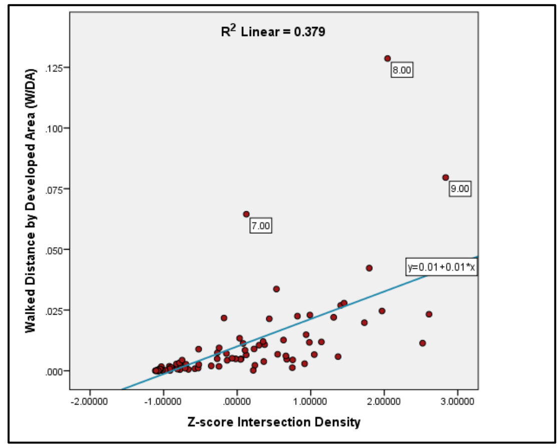
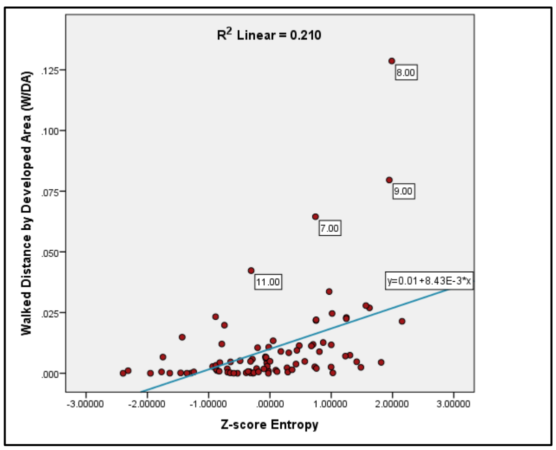


Figure 5.9. Z-score Dwelling Density

Figure 5.10. Z-score Intersection Density



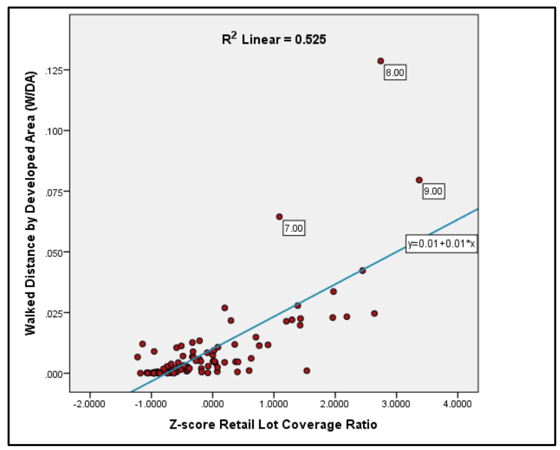


Figure 5.11. Z-score Entropy

Figure 5.12. Z-score Entropy

As explained in previous chapters, Z-score walkability variables were combined into a three-variable walkability index (WI-3) and a four-variable walkability index (WI-4). The WI-3, when tested against W/P, resulted in a weak, but significantly linear association with W/P (r = 0.363, p = 0.001) (Table 5.1). The scatterplot for WI-3 and W/P is illustrated in Figure 5.13. The results are comparable in strength to correlations between the three individual index variables and entropy alone (r = 0.391), but entropy is a slightly better estimator of W/P than the composite WI-3.

The WI-4 index (r = 0.420) is a slightly better estimator of W/P than WI-3. Figure 5.14 illustrates the fitted regression line and associated model, which shows that WI-4 is better able to describe the walking distance per person. The b coefficient is 2.51 in the WI-4 compared to 2.16 in the WI-3 regression line (Figure 5.13). Notably, of all four separate variables and two composite indices, retail lot coverage ratio remains the best estimator of W/P with a moderate correlation (r = 0.489), and also has the greatest effect on Y (b = 2.91) (Figure 5.4).

WI-3 and WI-4 both have moderate correlation with W/R. Figures 5.15 and 5.16 illustrate the relationship of both walkability indices against W/R. The regression coefficients indicate similar effects on Y for both indices. WI-3 was found to be moderately correlated with W/DA (r = 0.645, p = 0.000), while WI-4 was found to be strongly correlated (r = 0.704, p = 0.000). The strong correlation of WI-4 with W/DA is heavily influenced by the inclusion of retail lot coverage ratio, as retail lot coverage ratio alone had a higher correlation with W/DA (r =0.724) than the composite WI-4 (r = 0.704).

The individual walkability variables are all moderately or strongly correlated with both walking indices. As shown in Table 5.1, all of these correlation values exceed 0.697 and reach as high as 0.883. Intersection density is most highly correlated with WI-3 and WI-4 (r = 0.876, r = 0.883, respectively). These results indicate that intersection density by itself is a strong estimator of the walkability indices in that intersection density is indicative of shorter neighbourhood blocks (Cervero et al., 1997), and is associated with decreased car use due to travel delays (Peiravian et al., 2014), and facilitates more efficient walking routes (Jacobs, 1961).

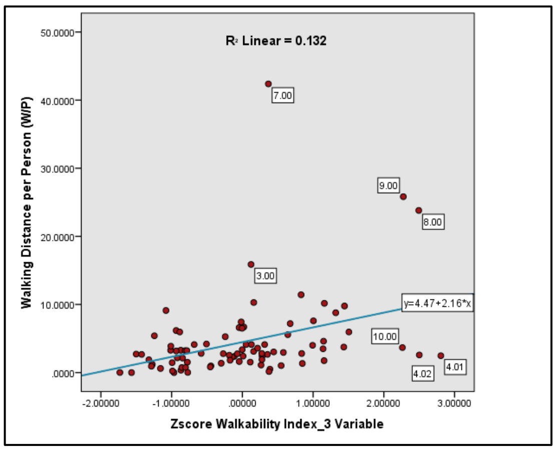


Figure 5.13. W/P by WI-3

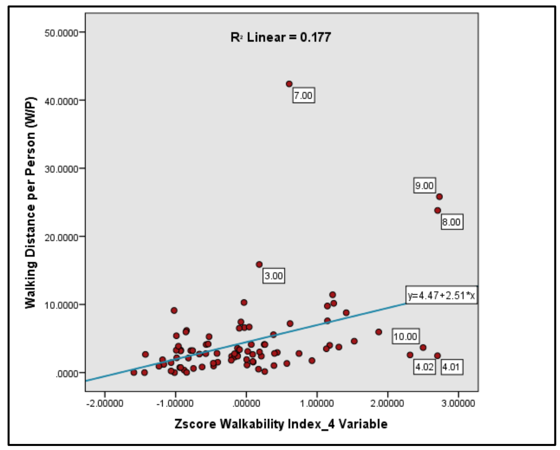


Figure 5.14. W/P by WI-4

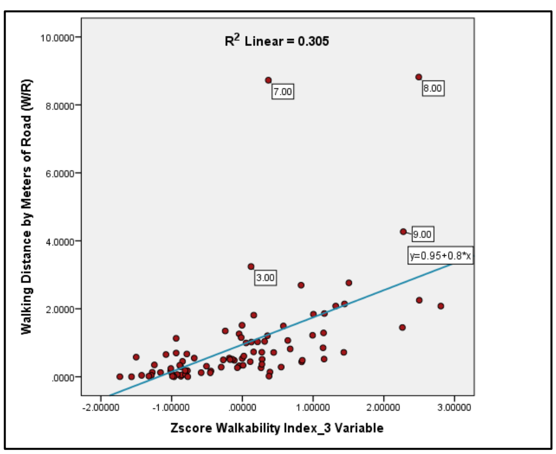


Figure 5.15. W/R by WI-3

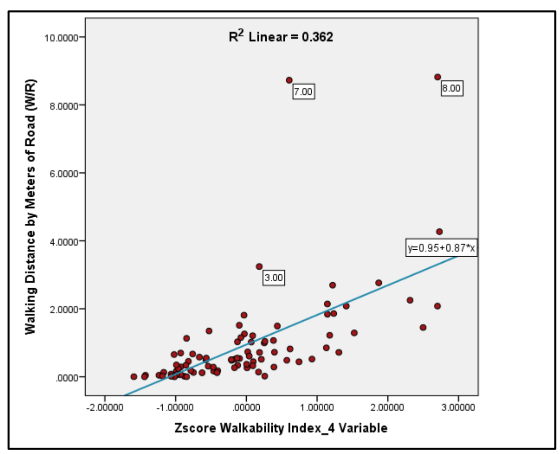


Figure 5.16. W/R by WI-4

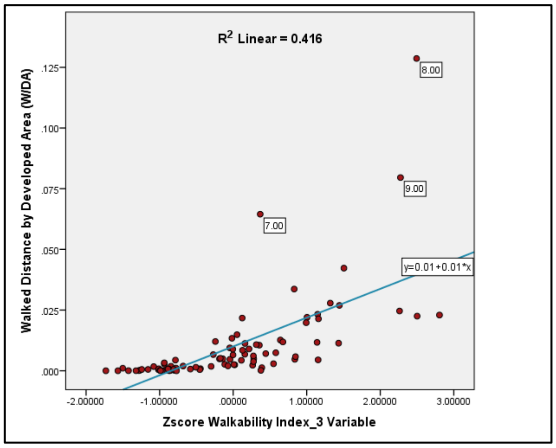


Figure 5.17. W/DA by WI-3

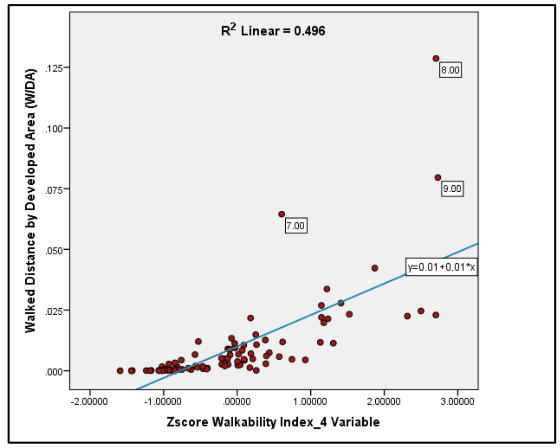


Figure 5.18. W/DA by WI-4

Results in Table 5.1 also highlight multicollinearity among the variables. Dwelling density, intersection density, and retail lot coverage ratio are all strongly correlated with each other. This multicollinearity brings into question the rationale for creating a composite index. Interestingly, one would expect that dwelling density would have a strong correlation with entropy, as typically areas of higher land use mix are associated with areas of high dwelling density. However, the data yield only a moderate correlation of r = 0.418 (Table 5.1). Also, entropy is only moderately correlated with the three other walkability variables, suggesting that entropy is measuring an aspect of the built environment that is not replicated in the other variables. A composite index should include separate variables that have little or no correlation to ensure the index is measuring distinct characteristics of the subject matter; otherwise, it seems that one could simply use a single variable and would have similar explanatory power. This research suggests that the use of a single variable (retail lot coverage ratio) would provide greater estimating potential than either composite index.

CHAPTER 6

Modeling: Multiple Regression

Six separate stepwise multiple linear regression models were estimated for each dependent variable (W/P, W/R, and W/DA), to gauge their associations with the walkability indices, built environment variables, and socio-demographic control variables. The software used was IBM SPSS Version 21. The purpose of estimating the regression models was to determine which walkability indices and built environment variables have the strongest ability to explain walking behaviour. Although a regression formula can be derived from each model, the primary objective was not to quantitatively predict walking behaviour, but to explain the most useful estimators of walking behaviour. This is not to suggest that correlation necessarily or directly implies causality; however, both common sense and the literature suggest there should be at least partial causality between the built environment and socio-demographic variables and the amount of walking (Cervero, 1988; Handy et al., 2002; Frank at al., 2007; Mayne et al., 2013; Villanueva et al., 2014).

Stepwise forward regression was used in this analysis, recognizing that this method has both advantages and disadvantages compared with the enter method (Thompson et al., 1995; Whittingham et al., 2006). The forward stepwise method excludes more of the entered variables, thus providing a multiple regression equation which is simpler to interpret and easier to employ, but which may lead to spurious interpretations when multicollinearity is high. The enter method retains all of the entered variables, and can better handle multicollinearity, but may result in an inflated R^2 value and a more complex regression equation. To verify the robustness of the forward stepwise results, a duplicate enter method regression analysis was completed. The enter method provided results that were very similar to those found with stepwise. The adjusted R^2 was indeed slightly inflated for many of the enter models, but the significant variables and the β-weights were similar to the stepwise results. The full enter method results can be found in Appendix A. As previously discussed, this study uses three measures of walking density to represent the amount of walking: walked distance per person (W/P), walked distance per meter of road (W/R), and walked distance per developed area (W/DA). Table 6.1 consolidates the results from the six regression models using W/P as the dependent variable.

Model 1 tests the standard (basic) three components that comprise the WI-3 walkability index (dwelling density, intersection density, and entropy) against walked distance per person (W/P). The model determined that only entropy was significant (p = 0.000) and both dwelling density and intersection density variables were excluded from the model due to p-values exceeding 0.05 (Table 6.1). The coefficient of multiple determination (adjusted R^2), indicates that Model 1 explains 14.3% of the variation in W/P accounted for by entropy (aR^2 = 0.143). The significance of the F test confirms this model as significant (p = 0.000).

The regression analysis performed in Model 2 tests the three variables included in Model 1 (dwelling density, intersection density, entropy) plus retail lot coverage ratio. Model 2 results in three variables (retail lot coverage ratio, dwelling density, entropy) being included in the model and only one variable being excluded (intersection density) (Table 6.1).

Dependent Variable: Walked Distance by Population (W/P)										
(STEPWISE Method)			**Unstandardized**	**Standardized**						
		(order of entry)	**Coefficients**	**Coefficients**				**Adjusted**		
	Excluded Variables	**Included Variables**	**B**	**Beta**	**Sig.**	**t**	**R^2**	**R^2**	**F**	**Sig.**
Model 1							0.153	0.140	15.378	0.000
IV:	Dwelling Density	Constant	4.470							
	Intersection Density	Entropy	2.330	0.391	0.000	3.920				
Model 2							0.359	0.336	15.492	0.000
IV:	Intersection Density	Constant	4.470							
		Retail Lot Coverage Ratio	3.920	0.657	0.000	5.159				
		Dwelling Density	-2.420	-0.406	0.002	-3.206				
		Entropy	1.670	0.280	0.006	2.830				
Model 3							0.503	0.472	16.395	0.000
IV:	Intersection Density	Constant	18.860							
	% Senior	% Middle Aged	-0.462	-0.581	0.000	-4.092				
	% Young Adult	Dwelling Density	-3.430	-0.575	0.000	-4.633				
		Retail Lot Coverage Ratio	2.650	0.445	0.001	3.438				
		Average Income	0.082	0.303	0.001	3.355				
		Entropy	1.670	0.281	0.003	3.038				
Model 4							0.593	0.562	19.411	0.000
IV:	Intersection Density	Constant	7.760							
	Entropy	% Institutional Land Use	0.330	0.381	0.000	4.269				
	% Young Adult	% Office Land use	0.900	0.323	0.000	4.006				
	% Residential Land Use	Average Income	0.080	0.319	0.000	3.914				
	% Parkland Land Use	Retail Lot Coverage Ratio	2.780	0.360	0.004	3.002				
	% Commercial Land Use	Dwelling Density	-1.720	-0.413	0.001	-3.559				
	% Industrial Land Use	% Middle Aged	-0.270	-0.336	0.020	-2.377				
	% Senior									
Model 5							0.332	0.308	13.75	0.000
IV:	WI-4*	Constant	35.370							
	% Senior	% Middle Aged	-0.634	-0.797	0.000	-6.052				
		Average Income	0.068	0.252	0.014	2.497				
		% Young Adult	-0.193	-0.278	0.027	-2.258				
Model 6							0.536	0.507	18.685	0.000
IV:	WI-4*	Constant	6.605							
	% Parkland Land Use	% Institutional Land Use	0.280	0.393	0.000	4.126				
	% Commercial Land Use	% Office Land use	0.903	0.371	0.000	4.405				
	% Industrial Land Use	Average Income	0.098	0.363	0.000	4.172				
	% Young Adult	% Middle Aged	-0.209	-0.262	0.013	-2.533				
	% Senior	% Residential Land Use	-0.059	-0.169	0.040	-2.085				
	* WI-4: four-variable walkability index									

Table 6.1. Walked Distance by Population, Consolidated Regression Results

With the inclusion of the retail lot coverage ratio, Model 2 explains a much higher amount of the variation ($aR^2 = 0.336$) in W/P than Model 1. Out of the three variables included in the model, retail lot coverage ratio contributes the greatest to the model (β = 0.657). Dwelling density is the second largest contributor to the explanatory power of the model (β = -0.406). Unexpectedly, the coefficient for dwelling density is negative; suggesting that as dwelling density increases the resulting walking per resident person decreases. Upon reflection, this makes sense: as residents are added in a given area, naturally there would be a greater amount of walking overall due to the additional people, but it may not result in a greater percentage of people walking, or a greater distance walked per person.

Regression Model 3 included all variables in Model 2 (dwelling density, intersection density, entropy, retail lot coverage ratio) and then added the socio-demographic control variables (percent young adult, percent middle-aged, percent older adult, and average income) (Table 6.1). Model 3 is statistically significant (p = 0.000) and explains 47.2% of the variation in W/P. By including the control variables, Model 3 becomes an improved estimator of the variation in W/P compared to Model 2. Model 3 excluded three variables from the regression analysis (intersection density, percent older adults, and percent young adults). Of the included variables, percent middle-aged was found to have the greatest effect on the estimating potential of the model (β = -0.581), but a built environment variable was of almost equal importance (dwelling density, β = -0.575). The negative value of the coefficient suggests that as the percentage of middle-age people in a census tract increases, the amount of walking per person would decrease. Intuitively, this result makes sense, in that middle-aged people (40-64) tend to be busier with family and career obligations as compared to the two other age cohorts, and are perhaps also more reliant on automobile transport. There is evidence in the literature that both older adults and younger adults have a greater propensity to walk (Spinney et al., 2012), though the two age groups tend to walk for different purposes. Younger adults tend to walk for active transportation, while older adults tend to walk for recreation (Spinney et al., 2012). Of the three additional explanatory variables, retail lot coverage contributes most, followed by entropy, and average income (Table 6.1).

The regression models continue to build in complexity. Model 4 includes the four built environment variables and the control variables, and then adds six land use variables expressed as a percent (percent residential, percent commercial, percent industrial, percent institutional, percent park and recreation, and percent office) (Table 6.1). Model 4 has the greatest estimating potential of the six regression models in this study, explaining 56.2% of the variation in W/P ($aR^2 = 0.562$). Eight variables were excluded from the model, leaving six variables remaining. Notably, entropy was excluded from this model where the previous three models included it, and Model 1 selected entropy as the only significant variable out of the three built environment variables. Only two built environment variables (retail lot coverage ratio and dwelling density) were included in the model, the highest coefficient of all variables being dwelling density (β = -0.413). Of the six land-use variables entered in the model, only two (percent institutional and percent office) were found to be significant, both having a moderate effect on the explanation of W/P (β = 0.382 and 0.323, respectively). The inclusion of institutional and office land uses suggests that the presence of employment land uses may be a better estimator of walking than those land uses associated with recreation or habitation. This notion is supported in the literature, as Millward, Spinney, and Scott's (2013) results suggest that the majority of walking does not take place in areas around people's homes.

As discussed in previous chapters, four variables are combined into a frequently-used version of the walkability index. Model 5 tests the four-variable walkability index (WI-4) and the socio-demographic control variables against W/P. WI-4 was found to be not statistically significant, having a *p*-value greater than 0.05, and was thus excluded from the model. This model creates an interesting comparison with model 3, where model 3 tests the four individual walkability components against W/P, as opposed to a composite index in this model. The regression analysis of Model 3 showed that three (dwelling density, retail lot coverage ratio, and entropy) of the four individual walkability variables were statistically significant. However, when the individual components of the walkability index were combined into a walkability index and then tested against W/P, the index was found to be insignificant and was excluded from the model. The exclusion of the walkability index from the regression analysis is contrary to much of the literature regarding the walkability index. Many studies weight specific components of the walkability index (Frank et al., 2005; Frank et al., 2006; Kerr et al., 2006), although specific details of weighting procedures are often lacking. It would seem that researchers could use the Beta values as an appropriate weighting system.

Three socio-demographic variables were statistically significant in Model 5; percent middle-aged, average income, and percent young adult. The model weakly explains 30.2% of the variation in W/P ($aR^2 = 0.308$). Percent middle-aged is the dominant variable of the three, with a strong coefficient of -0.797. This suggests that middle-aged people tend to walk less than other cohorts. So, as the percentage of middle-aged population increases within a census tract, the model suggests that walked distance per person will decrease.

Model 6 tests the four-variable walkability index (WI-4), socio-demographic variables, and the percentage of land use against W/P. This model explains a moderate 50.7% of the variation in W/P. As with Model 5, WI-4 was found to be not significant and was excluded from the model. Institutional and office land uses were the two most influential variables in the model ($\beta = 0.393$ and 0.371, respectively). This seems consistent with individual scatterplots that point to census tracts with high amounts of W/P also exhibiting a greater percentage of both institutional and office land uses. A third land use was included in the model (residential), albeit with a weak impact ($\beta = -0.169$). This is the only model that found residential land use to be significant. The residential Beta value is negative, supporting the notion that walking is occurring around places of employment or business more so than surrounding the home location. As previously referenced, this notion is supported by the findings of Millward, Spinney, and Scott (2013).

Average income and middle age were the final two variables included in Model 6. Consistent with previous models, middle age is found to have a negative effect on W/P with a Beta of -0.262. Income is positively associated with W/P, when other variables are held constant. This is also not unexpected, though the relationship is probably not linear: Spinney et al. (2012) provide evidence that walking participation levels are higher for both low and high income groups.

Table 6.2 consolidates the relevant information from the following six regression models (Models 7 through 12) using W/R as the dependent variable. These six models follow the same independent variable sequence as the previous six models tested against W/P. All of the models show better estimation of walked distance per meter of road (W/R) than do the equivalent models for walked distance per person (W/P). This accords with the fact that the independent variables tend to have higher Pearson correlations with W/R, and underscores the fact that much of the walking does not occur in respondents' home neighbourhoods (Millward et al., 2013).

Model 7 tests the three walkability variables against W/R. The adjusted R^2 is 0.274, which suggests that the model has weak explanatory power (Table 2). Both intersection density and entropy are significant and, therefore, were included in the model. Intersection density appears to have a slightly stronger impact on W/R than entropy (β = 0.340 and 0.289, respectively).

Model 8 builds on Model 1 and uses the same three variables and includes retail lot coverage ratio, which is the fourth characteristic of walkability (Table 6.2). This model accounts for a moderate 40.8% of the variability in W/R. Retail lot coverage ratio and entropy were found to be significant and included in the model. Retail lot coverage ratio has more than double the influence on the model than entropy with a coefficient of 0.522 compared to entropy's coefficient of 0.223. Intersection density and dwelling density were both found to be insignificant and were thus excluded from the model

Model 9 tests the four walkability variables and introduces the socio-demographic control variables (percent young adult, percent middle-aged, percent older adult, and average income) (Table 6.2). The explanatory power of this model continues to become stronger as additional control variables were added. This model explains 50.5% of the variation in W/R. Four variables were included in the model (percent middle age, retail lot coverage ratio, average income, and entropy). Two of these four variables were socio-demographic control variables (percent middle age and average income) and two were walkability indicator variables (retail lot coverage ratio and entropy). Four variables were excluded from the model (intersection density, percent older adult, percent young adult, and dwelling density). Interestingly, as additional variables are added to the sequential models, entropy has a reduced importance in each. This suggests that entropy does not have sufficient explanatory influence to overcome the addition of socio-demographic variables. Model 9 includes entropy but it has a very weak coefficient (β = 0.219). Percent middle-aged is the variable with the greatest effect on the estimating potential of the model (β = -0.447). The coefficient for percent middle-aged continues to be negative, as observed when testing this variable against W/P. A negative coefficient suggests that as the percentage of middle-aged population in a particular census tract increases, the amount of walking per meter of road will decrease. As previously described, a negative coefficient for the percent middle-aged variable is anticipated as the population in this cohort are generally busy with family and careers and have less time for walking (Spinney et al., 2012). Retail lot coverage ratio continues to have relatively moderate strength when compared to the three other variables included in the model (β = 0.285).

Model 10 builds on Model 9 by adding the six percentage of land use variables (percent residential, percent commercial, percent industrial, percent institutional, percent park and recreation, and percent office) to the four walkability characteristics (dwelling density, intersection density, entropy, and retail lot coverage ratio) and the socio-demographic variables (percent young adult, percent middle-aged, and percent older adult). This model has the most effective explanatory power of the six models tested against W/R; with an ability to explain 60.7% of the variation in W/R. The percentage of institutional land use is the most significant variable in the model with a Beta coefficient of 0.358, followed by percent office land use (β = 0.305). These two land-use variables are clearly more effective explanatory variables than the standard components of the walkability index. Only one walkability index variable is included in the model (retail lot coverage ratio), while the other three components are excluded from the model.

Model 11 tests the composite, four-variable walkability index and the socio-demographic variables (percent young adult, percent middle-aged, percent older adult, and average income). This model explains 46.5% of the variability in W/R. In contrast to the corresponding Model 5 for W/P, the walkability index is now found to be significant and thus included in the model. However, percent middle-aged is still the most influential variable in this model.

Dependent Variable: Walked Distance by Road Length (W/R)										
			Unstandardized	Standardized						
		(order of entry)	Coefficients	Coefficients				Adjusted		
	Excluded Variables	Included Variables	B	Beta	Sig.	t	R^2	R^2	F	Sig.
Model 7							0.290	0.274	17.196	0.000
IV:	Dwelling Density	Constant	0.947							
		Intersection Density	0.493	0.340	0.001	3.283				
		Entropy	0.419	0.289	0.007	2.790				
Model 8							0.422	0.408	30.633	0.000
IV:	Intersection Density	Constant	0.947							
	Dwelling Density	Retail Lot Coverage Ratio	0.756	0.522	0.000	0.568				
		Entropy	0.323	0.223	0.017	2.431				
Model 9							0.528	0.505	22.955	0.000
IV:	Intersection Density	Constant	3.436							
	% Senior	% Middle Aged	-0.086	-0.447	0.001	-3.549				
	% Young Adult	Average Income	0.018	0.278	0.002	3.178				
	Dwelling Density	Entropy	0.317	0.219	0.016	2.449				
		Retail Lot Coverage Ratio	0.413	0.285	0.019	2.401				
Model 10							0.630	0.607	27.615	0.000
IV:	Intersection Density	Constant	1.370							
	Entropy	% Middle Aged	-0.053	-0.275	0.024	-2.306				
	% Young Adult	% Office Land use	0.180	0.305	0.000	4.096				
	% Residential Land Use	% Institutional Land Use	0.062	0.358	0.000	4.266				
	% Parkland Land Use	Average Income	0.020	0.299	0.000	3.876				
	% Commercial Land Use	Retail Lot Coverage Ratio	0.364	0.251	0.021	2.361				
	% Industrial Land Use									
	% Senior									
	Dwelling Density									
Model 11							0.483	0.465	25.898	0.000
IV:	% Senior	Constant	4.307							
	% Young Adult	% Middle Aged	-0.103	-0.531	0.000	-4.029				
		Average Income	0.016	0.240	0.008	2.699				
		WI-4*	0.410	0.283	0.031	2.201				
Model 12							0.605	0.586	31.367	0.000
IV:	WI-4*	Constant	2.862							
	% Parkland Land Use	% Middle Aged	-0.088	-0.454	0.000	-4.781				
	% Residential Land Use	% Office Land use	0.194	0.329	0.000	4.344				
	% Commercial Land Use	% Institutional Land Use	0.063	0.364	0.000	4.221				
	% Industrial Land Use	Average Income	0.019	0.294	0.007	3.719				
	% Young Adult									
	% Senior									
	* WI-4: four-variable walkability index									

Table 6.2. Walked Distance by Road Length, Consolidated Regression Results

There is value in comparing regression results of the model that includes the walkability index components separately (Model 9) to the model that includes the walkability index components combined in one composite index (Model 11). When Model 11 is compared with Model 9, it illustrates that separate walkability components provide better estimating potential than the walkability indices, as observed by the adjusted R^2 of each (Model 9 $aR^2 = 0.505$, Model 11 $aR^2 = 0.465$). This result brings into question the necessity of creating a composite index when the separate components are better estimators than the composite index.

Model 12 again builds upon Model 11 by adding the percentages of six land uses. Only four variables were found to be significant and remained in the model, two socio-demographic variables (percent middle-aged and average income) and two land uses (percent office and percent institutional). In contrast to Model 11, the walkability index was excluded from the model. The percentage of middle-aged population within a census tract continued to be the most dominant factor impacting the estimating potential of the model, but not in a positive manner, as the coefficient of percent middle-aged is negative as observed in previous models ($\beta = -0.454$).

Model 12 should be compared with Model 10 for the same rationale as described for comparing Model 9 with Model 11. The intention is to evaluate the effectiveness of the walkability index. Comparison of the adjusted R^2 confirms that it is more effective to test individual components of the walkability index against W/R than to use the composite walkability index (Model 10 $aR^2 = 0.607$, Model 12 $aR^2 = 0.586$).

The final walking density tested in this study is W/DA. Table 6.3 consolidates the relevant information from six regression models (Models 13 through 18) using W/DA as the dependent variable. These six models follow the same independent variable sequence as the previous six models tested for each walking density variable.

The three basic walkability components (entropy, intersection density and residential density) were tested in Model 13, explaining 40.3% of the variability in W/DA. Only intersection density and entropy were found to be significant, with intersection density being a more important indicator than entropy (β = 0.514 and 0.221, respectively).

Model 14 builds on the three basic components of the walkability index (WI) as tested in Model 13 and includes retail lot coverage ratio. With the inclusion of the fourth component of the WI, this model creates a better estimation of the variation in W/DA ($aR^2 = 0.541$) than Model 13 ($aR^2 = 0.403$). Retail lot coverage ratio had much more influence on the model than entropy, as indicated by comparing the Beta values (β = 0.647 and 0.181, respectively).

Model 15 tests the four walkability variables and introduces the socio-demographic control variables (percent young adult, percent middle-aged, percent older adult, and average income) (Table 6.3). As additional variables are added to the model, the ability to explain the variation in W/DA improves. Model 15 explains 55.8% of the variation in W/DA. Retail lot coverage ratio is a substantially more influential variable (β = 0.688) than either entropy or average income (β = 0.229 and 0.165, respectively).

Model 16 is the best estimator of the variation in W/DA out of the six models (aR^2=

0.683). This model includes the four individual walkability variables and the control variables, and introduces the six land use variables into the model. In keeping with the previous two models, retail lot coverage ratio continues to be the most influential variable out of the included set of variables (β = 0.550).

Model 17 tests the walkability index and the socio-demographic control variables. This model explains 49.0% of the variability in W/DA. The walkability index variable was the only variable found to be statistically significant, and thus included in the model, while all other variables were excluded. Model 17 can be compared with Model 15, as Model 15 tests each walkability component separately, while Model 17 includes those same variables as a composite walkability index. Model 15 is a better estimator or W/DA (aR^2= 0.558) compared to Model 17 (aR^2= 0.490). In the same manner, Model 18 can be compared with Model 16. It was found that when the four components of walkability are included separately (Model 16), as opposed to as a composite index (Model 18), the model with separate components was a moderately better estimator of W/DA than the walkability index (Model 16 aR^2 = 0.683, Model 18 aR^2 = 0.644).

The three regression tables display the same walkability, socio-demographic, and land-use variables against the three measures of walking (W/P, W/R and W/DA), and all three produced similar results. The fourth models of all three walking densities (i.e. Models 4, 10, and 16) were consistently better estimators of the respective walking density than the other five models (see adjusted R^2 for each table).

Dependent Variable: Walked Distance by Developed Area (W/DA)										
(STEPWISE Method)			Unstandardized	Standardized						
		(order of entry)	Coefficients	Coefficients				Adjusted		
	Excluded Variables	**Included Variables**	**B**	**Beta**	**Sig.**	**t**	**R^2**	**R^2**	**F**	**Sig.**
Model 13							0.417	0.403	30.076	0.000
IV:	Dwelling Density	Constant	0.010							
		Intersection Density	0.009	0.514	0.000	5.467				
		Entropy	0.004	0.221	0.000	2.349				
Model 14							0.552	0.541	51.686	0.000
IV:	Intersection Density	Constant	0.010							
	Dwelling Density	Retail Lot Coverage Ratio	0.012	0.647	0.000	8.003				
		Entropy	0.003	0.181	0.028	2.241				
Model 15							0.573	0.558	37.152	0.000
IV:	Intersection Density	Constant	0.001							
	% Senior	Retail Lot Coverage Ratio	0.013	0.688	0.000	8.405				
	% Young Adult	Entropy	0.004	0.229	0.007	2.764				
	% Middle Aged	Average Income	0.000	0.165	0.044	2.043				
	Dwelling Density									
Model 16							0.698	0.683	47.342	0.000
IV:	Intersection Density	Constant	-0.008							
	Entropy	Retail Lot Coverage Ratio	0.010	0.550	0.000	7.446				
	% Young Adult	% Office Land use	0.003	0.398	0.000	6.036				
	% Residential Land Use	% Institutional Land Use	0.001	0.246	0.001	3.532				
	% Parkland Land Use	Average Income	0.000	0.205	0.003	3.030				
	% Commercial Land Use									
	% Industrial Land Use									
	% Senior									
	% Middle Aged									
	Dwelling Density									
Model 17							0.496	0.490	83.556	0.000
IV:	Average Income	Constant	0.010							
	% Senior	WI-4*	0.013	0.704	0.000	9.141				
	% Middle Aged									
	% Young Adult									
Model 18							0.665	0.644	32.165	0.000
IV:	% Middle Aged	Constant	-0.014							
	% Residential Land Use	WI-4*	0.012	0.670	0.000	6.301				
	% Commercial Land Use	% Office Land use	0.003	0.364	0.000	4.975				
	% Industrial Land Use	% Institutional Land Use	0.001	0.245	0.002	3.180				
	% Young Adult	Average Income	0.000	0.228	0.002	3.127				
	% Senior	% Parkland Land Use	0.000	0.231	0.010	2.632				
	* WI-4: four-variable walkability index									

Table 6.3. Walked Distance by Developed Area, Consolidated Regression Results

This "fourth" model tests the individual walkability index variables (entropy, residential density, intersection density, and retail lot coverage ratio), along with the socio-economic variables (income and age) and the percentage of six developed land uses for each CT. This regression analysis determined that land-use variables, such as residential, commercial, industrial, parkland, institutional, and office, in combination with socio-demographic control variables, were consistently the dominant variables. Specifically, both institutional and office land uses were found to be the two most influential variables against three walking densities. Average income was also identified as significant against all three walking densities as it was included in all three models that utilized separate variables (Models 4, 10, and 16). Retail lot coverage ratio was the only separate walkability component that was significant against W/P, W/R and W/DA. On the other hand, dwelling density was included in Model 4 against W/P, but was found to be insignificant and was excluded from Models 10 (W/R) and 16 (W/DA).

When the walkability variables were combined into a composite index, W/P, W/R and W/DA were most highly influenced by the same three variables (percent office, percent institutional, and average income) only varying by the individual strength of each. Interestingly, three out of the six regression models that included the composite walkability index excluded it from the respective models as it was not statistically significant. As indicated above, this result suggests that it is more effective to regress individual components of the walkability index against the three walking densities than to use the composite walkability index.

In general, of four built environment variables associated with walkability, retail lot coverage ratio was found to be the only influential variable on W/P, W/R, and W/DA. The land use and socio-demographic variables were found to be better estimators of walking densities. The above analysis suggests that it may be more beneficial to investigate individual components that influence walking rather than combining walking influences into an index that only includes built environment variables. The consistent re-occurrence of both office and industrial land uses in the regression models indicates that walking is associated with areas of employment more so than with residential land uses. This notion is consistent with the work of Spinney, Millward and Scott (2012) that suggests walking for transport typically occurs outside of the home neighbourhood.

There is still value in creating a walkability index that can function as a consistent and objective evaluation tool for measuring the propensity of people to walk in any particular neighbourhood. Currently, as discussed in the literature review, much of the research focuses on walkability components surrounding the built environment (residential density, entropy, intersection density, and retail lot coverage ratio), but does not account for how these individual components function in terms of influencing walking. For example, entropy is a measure of land use mixture, but simply because land use is mixed, does not mean it necessarily correlates with increased walking. Rather it depends on the type of land use that is mixed. This notion could be considered an ecological fallacy (Freedom, 1999), whereby the aggregate walkability index may be a reasonable estimator of walking in some scenarios; however, it would be incorrect to infer that each component of the index also has the same correlation with walking. Based on this research there is greater likelihood that a more robust and better estimator of the propensity for walking could be created by including alternative, or at least additional, variables particularly associated with land uses that pertain to employment functions and income.

CHAPTER 7

Conclusions

The purpose of this research was to identify the location of walking activity in a medium-sized North American city and to identify built environment characteristics associated with walking activity aggregated at the neighbourhood level. This research uniquely contributes to the knowledge of walking behaviour in the following three ways: (1) the use of both objective and self-report time diaries to record walking activity; (2) walking data are aggregated based on census tracts (CT), enabling the CTs themselves to be considered the unit of measurement, rather than the individual respondents; and, (3) walking activity includes both active transport and recreational walking so that total walking can be examined. The research is particularly concerned with evaluating the highly-cited index of walkability (e.g. Frank et al., 2005; Lee and Moudon, 2006) against objective and verified walking data, and ultimately to provide insights that will lead to improved neighbourhood design. Planners, urban designers, and other professionals interested in walking behaviour will benefit from this research by gaining further insights into how built environment characteristics influence the location of walking in a medium-sized North American city.

Much of the current research on walking behaviour focuses on individual respondents and establishes a neighbourhood buffer surrounding their home location. The assumption is that most walking is home-based, and that walking frequency is thus highly influenced by the built environment of the home neighbourhood. This assumption is unwarranted, however, as previous researchers have contended; a large portion of walking activity does not take place surrounding the respondents' home locations (Spinney et al., 2012; Millward et al., 2013).

This research focussed on where walking occurs, rather than who walks. The aggregation of walking by census tracts enabled the CTs themselves to be considered the unit of measurement, rather than the individual respondents. Inherent variation in CTs prompted the creation of three walking densities to reflect different aggregates of walked distance. These are: walking density per resident population (W/P); walking density per road length of road network (W/R); and, walking density per developed area (W/DA).

The spatial analysis component of this research provided several insights into walking behaviour. The walking tracks and associated walking density measures all confirm that a significant portion of walking activity takes place in the urban centres; specifically in the CBDs of both Halifax Peninsula and Dartmouth, where there are areas of high employment (e.g. universities, hospitals, and office towers). Although walking more frequently occurs in inner-urban areas compared to suburban areas, a considerable amount of walking still occurs in the suburbs. Specifically, suburban walking occurs along major transportation routes and areas dominated by commercial or recreational land uses. Insights can also be gained by understanding where walking does not occur. The data suggest that inner-city areas such as the North Ends of both Halifax and Dartmouth, and also suburban Spryfield, experience very little walking. This may be due to factors that have negative influences that impact the pedestrian milieu, such as perceived lack of safety and poor neighbourhood aesthetics. Also, little walking activity occurs in the large Burnside Business Park. This could be a result of both low employment density and the absence of sidewalks, which both tend to encourage vehicular travel.

The multiple regression component of this research provides valuable insights into the role of the built environment on walking activity. The three walking densities were modeled against two variations of the walkability index (WI) as well as against the individual components of the WI, six land uses, and several socio-demographic variables. Six models were tested for each walking density; 18 separate models in total. Multiple regression models consistently indicate that built environment variables (intersection density, dwelling density, entropy, and retail lot coverage ratio), as well as land-use and socio-demographic variables, are the best predictors of walking density. Two of the three walking density models indicate the walkability index variable was not a statistically significant predictor of walking density. This suggests that the component variables of the walkability index are better estimators than the composite index.

Employment-related land uses (institutional and office), retail lot coverage ratio, and income were consistently included in each of the best predictive models and they all exhibited moderate estimating potential. These findings are contrary to much of the existing literature that suggests amalgamating built environment variables into a single index creates a more accurate predictive model. This research suggests that creating a composite index is unnecessary. Moreover, focusing on only four variables – office and institutional land uses, retail lot coverage, and income – creates a model that more accurately estimates walking behaviour. This notion supports Lee and Moudon's (2006) contention that complex indices are time consuming to create and expensive field data collection may not be necessary to estimate or explain walking behaviour.

Further research is required to explore walking behaviour in association with the four walking determinants noted above. The two land uses (institutional and office), although represented in the regression models as unique variables, could be examined collectively as "employment land uses". Based on a review of the literature, it appears that little attention has been given to walking activity in relation to employment location. Other researchers have identified this gap in knowledge (Spinney et al., 2012; Millward et al., 2013) and this research supports their notion that a large portion of walking activity occurs outside of the home neighbourhood. In fact, Kwan (2013) acknowledges the importance of the workplace and school as daily space-time anchors for individuals. Further research could explore walking activity specifically around employment centres. Based on this research, planners and urban designer should be attentive to the pedestrian environment surrounding not only residential neighbourhoods, but particularly neighbourhoods of dense employment land uses. Better still, planners and urban designers should consider the integration of these residential and employment land uses, such as in traditional neighbourhood design and in New Urbanist neighbourhoods (seminal works on these include Lynch, 1960, Katz, 1993, and Handy 2008, respectively).

Retail lot coverage, often termed commercial floor area ratio or just floor area ratio

(FAR) has been frequently used as a component of walkability indices (Handy et al., 2002; Cerin et al., 2006; Frank et al., 2007; Adams et al., 2011; Wood et al., 2010; Mayne et al., 2013). Retail lot overage ratio can be thought of as a measure of pedestrian friendliness. Land parcels with high retail lot coverage ratios will have more ground floor building area (retail space) and, thus, less space allocated to parking. A site design with a high retail lot coverage ratio would have less expansive parking lots and, therefore, would be more advantageous to pedestrian access. Retail lot coverage could be thought of as one measure of the quality of the pedestrian environment. There are several other site design elements which enhance the pedestrian environment that coincide with increasing retail lot coverage, such as widening sidewalks, and re-locating parking (to the side, rear, or underground) in relation to pedestrian access (Lund, 2003). In response to this research, municipal planners could adjust applicable minimum parking standards for retail commercial uses. Creating a reduced parking requirement would encourage retail sites to have smaller parking lots, thus providing greater flexibility in parking location and site design, thereby improving the pedestrian environment.

Income is often included as a control variable in walking activity research (Kerr et al., 2006; Sallis et al., 2009). This study did not specifically investigate the potential impact of income variation on walking behaviour. Further research could explore the impact of income on walking behaviour to determine if income alone affects walking activity, perhaps due to lower income groups having reduced access to vehicles, or whether income is correlated with safety concerns that impact walking activity. In fact, this raises an important question for future research regarding model specification: do all independent variables contribute equally to the heuristics behind a person's decision to walk? Perhaps multi-level modelling, such as structural equation modelling, would be better suited to represent the impacts of control variables and built environment variables on walking behaviour.

In addition to the suggested further research noted above, there are additional research directions that could be explored. For example, it would be beneficial to separate walking episodes into active transportation (AT) and recreational trips and then duplicate this study's methods for each trip type. This research could provide valuable insights into built environment characteristics that influence AT and recreational walking in different manners. For example, it is expected that much of the recreational walking occurs in parks near a person's home, whereas much of the AT walking is in commercial areas and to and from transit nodes.

Another potential research direction could be to refine the walkability index (WI), as typically employed by researchers, to create a more accurate WI. Other researchers have attempted to create various alternative walkability indices, such as Giles-Corti et al., (2014) and Peiravian (2014); however, these indices include variables that may have little impact on walking behaviour. As an example, entropy is often cited as a walking determinant; however, as Brown et al. (2009) noted, a high entropy score may not necessarily be reflective of a built environment that encourages walking. Brown (2009, p.4) illustrates this point using the example of a study area that consisted of: one-third large single unit dwellings plus one-third big box retail stores located adjacent a highway plus one-third office park. This site would have a perfect entropy score, but simply because land uses are mixed does not necessarily result in a built environment conducive to walking.

Researchers often create a WI with varying weights of individual components (Frank et al., 2005; Frank et al., 2006; Sallis et al., 2009). One could weight the specific components of the WI based on the standardised beta coefficients (β) of a model that tested each component individually (as performed in model four of each density measure). This research could provide further refinement and appropriate rationale to the WI. Building on the theme of creating an index measure of neighbourhood walkability based on the 5Ds – density, diversity, design, distance to transit, and destination accessibility (Cervero et al., 2009) – researchers could develop an entirely new walkability index based on the four attributes found in this research that most strongly influence walking behaviour. Arguably, a new index based on this research would include data that are readily available and, thus, would create an index that may be less complicated and more easily replicable than those proposed by others.

This study has limitations that ought to be noted. Although the Halifax STAR survey attempted to stratify the respondents both in terms of age and geographic location, both those in younger age cohorts and rural households are under-represented in the survey (Spinney et al., 2011). It would certainly create a more representative study if younger age respondents had been adequately sampled; however, the fact that rural households are under-represented is of little consequence to this study, as the focus of this research is on neighbourhood form in urban and suburban contexts. From a statistical perspective, the R-square values in this study were well below 100%; therefore, there are additional factors not considered in this analysis that impact walking behaviour. Further, the Halifax STAR survey is only a cross-sectional sample (although robust) and may not be representative of the entire population. An additional limitation relates to the process of aggregation: whenever spatial geographic research is conducted, thought must be given to the modifiable aerial unit problem (MAUP). The rationale for dealing with MAUP has been addressed in Chapter three. However, it is worth noting that if the scale of the aggregation units in this study was altered, the analysis would likely produce different results. Perhaps the final limitation that should be discussed is that fact that this research is quantitative, and therefore lacks qualitative insights into how people feel about the built environment in which they walk.

The knowledge gained from this research will assist academic researchers and urban design / planning practitioners. These findings should provide academic researchers with further insights into the potential limitations of the existing walkability index as a tool and a concept, and highlight alternative built environment characteristics that strongly influence walking behaviour. Urban designers and planners may also benefit from these findings by gaining a further understanding of built environment variables that promote walking, thus enabling them to make more informed decisions related to neighbourhood design.

References

Adams, M.A., Sallis, J.F., Kerr, J., Conway, T.L., Saelens, B.E., Frank, L.D., Norman, G.J., Cain, K.L., 2011. Neighbourhood environment profiles related to physical activity and weight status: A latent profile analysis. *Preventive Medicine* **52**, 326–331.

Berke Ethan M EM, K., Thomas D., Moudon, Anne Vernez, Hoskins, Richard E., Larson, Eric B., 2007,. Association of the built environment with physical activity and obesity in older persons. *American Journal of Public Health* **97**, 486–92.

Booth, K.M., Pinkston, M.M., Poston, W.S.C., 2005. Obesity and the Built Environment. *Journal of the American Dietetic Association* **105**, 110–117.

Boarnet, M.G., Joh, K., Siembab, W., Fulton, W., Mai Thi Nguyen, 2011. Retrofitting the Suburbs to Increase Walking: Evidence from a Land-use-Travel Study. *Urban Studies* **48**, 129–159.

Brown, B.B., Yamada, I., Smith, K.R., Zick, C.D., Kowaleski-Jones, L., Fan, J.X., 2009. Mixed land use and walkability: Variations in land use measures and relationships with BMI, overweight, and obesity. *Health & Place* **15**, 1130–1141.

Brownson, R.C., Hoehner, C.M., Day, K., Forsyth, A., Sallis, J.F., 2009. Measuring the Built Environment for Physical Activity. *American Journal of Preventive Medicine* **36**, S99–S123.

Carlson, J.A., Saelens, B.E., Kerr, J., Schipperijn, J., Conway, T.L., Frank, L.D., Chapman, J.E., Glanz, K., Cain, K.L., Sallis, J.F., 2015. Association between neighbourhood walkability and GPS-measured walking, bicycling and vehicle time in adolescents. *Health & Place* **32**, 1–7.

Cerin, E., Leslie, E., Toit, L., Owen, N., Frank, L., 2007. Destinations that matter: Associations with walking for transport. *Health & Place* **13**, 713–724.

Cerin, E., Saelens, B.E., Sallis, J.F., Frank, L.D., 2006. Neighbourhood Environment Walkability Scale: Validity and Development of a Short Form. *Medicine & Science in Sports & Exercise* **38**, 1682–1691.

Cervero, R., Kockelman, K., 1997. Travel demand and the 3Ds: density, diversity, and design. Transportation Research Part D: *Transport and Environment* **2**, 199–219.

Cervero, 1988. Land-Use Mixing and Suburban Mobility. *Transportation Quarterly* **42**, 429–446.

Clark, A., Scott, D., 2014. Understanding the Impact of the Modifiable Areal Unit Problem on the Relationship between Active Travel and the Built Environment. *Urban Studies* **51**, 284–299.

Dewulf, B., Neutens, T., Van Dyck, D., De Bourdeaudhuij, I., Van de Weghe, N., 2012. Correspondence between objective and perceived walking times to urban destinations: Influence of physical activity, neighbourhood walkability, and socio-demographics. *International Journal of Health Geographics* **30**, 35–37.

Doyle, S., Kelly-Schwartz, A., Schlossberg, M., Stockard, J., 2006. Active community environments and health: the relationship of walkable and safe communities to individual health. *Journal of the American Planning Association* **72**, 19–31.

Ewing, R., Meakins, G., Hamidi, S., Nelson, A.C., 2013. Relationship between urban sprawl and physical activity, obesity, and morbidity – Update and refinement. *Health & Place* **26**, 118–126.

Frank, L.D., 2000. Land Use and Transportation Interaction: Implications on Public Health and Quality of Life. *Journal of Planning Education and Research* **20**, 6–22.

Frank, L.D., Andresen, M.A., Schmid, T.L., 2004. Obesity relationships with community design, physical activity, and time spent in cars. *American Journal of Preventive Medicine* **27**, 87–96.

Frank, L.D., Pivo, G., 1994. Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. *Transportation Research Record* **1466**, 44–52.

Frank, L.D., Sallis, J.F., Conway, T.L., Chapman, J.E., Saelens, B.E., Bachman, W., 2006. Many pathways from land use to health: associations between neighbourhood walkability and active transportation, body mass index, and air quality. *Journal of the American Planning Association* **72**, 75–87.

Frank, L.D., Sallis, J.F., Saelens, B.E., Leary, L., Cain, K., Conway, T.L., Hess, P.M., 2009. The development of a walkability index: application to the Neighbourhood Quality of Life Study. *British Journal of Sports Medicine* **44**, 924–933.

Frank, L.D., Schmid, T.L., Sallis, J.F., Chapman, J., Saelens, B.E., 2005. Linking objectively measured physical activity with objectively measured urban form. *American Journal of Preventive Medicine* **28**, 117–125.

Frank, L.D., Saelens, B.E., Powell, K.E., Chapman, J.E., 2007. Stepping towards causation: Do built environments or neighbourhood and travel preferences explain physical activity, driving, and obesity*? Social Science & Medicine* **65**, 1898–1914.

Fotheringham, S., Rogerson, P., 2009. *The SAGE Handbook of Spatial Analysis*. London England: SAGE Publications Ltd.

Giles-Corti, B., Macaulay, G., Middleton, N., Boruff, B., Bull, F., Butterworth, I., Badland, H., Mavoa, S., Roberts, R., Christian, H., 2014. Developing a research and practice tool to measure walkability: a demonstration project. *Health Promotion Journal of Australia* **25**, 1–7.

Glazier, R.H., Creatore, M.I., Weyman, J.T., Fazli, G., Matheson, F.I., Gozdyra, P., Moineddin, R., Shriqui, V.K., Booth, G.L., 2014. Density, Destinations or Both? A Comparison of Measures of Walkability in Relation to Transportation Behaviors, Obesity and Diabetes in Toronto, Canada. *PLOS ONE* **9,** 1–9.

Gordon, Peter; Richardson, Harry W, 1997. Are compact cities a desirable planning goal? *American Planning Association. Journal of the American Planning Association* **63**, 95–106.

Grasser, G., Van Dyck, D., Titze, S., Stronegger, W., 2013. Objectively measured walkability and active transport and weight-related outcomes in adults: a systematic review. *International Journal of Public Health* **58**, 615–625.

Hajna, S., Ross, N.A., Brazeau, A.-S., Bélisle, P., Joseph, L., Dasgupta, K., 2015. Associations between neighbourhood walkability and daily steps in adults: a systematic review and meta-analysis. *BMC Public Health* **15**, 1–8.

Handy, S.L., Boarnet, M.G., Ewing, R., Killingsworth, R.E., 2002. How the built environment affects physical activity: views from urban planning. *American journal of preventive medicine* **23**, 64–73.

Handy, S., Sallis, J.F., Weber, D., Maibach, E., Hollander, M., 2008. Is Support for Traditionally Designed Communities Growing? Evidence From Two National Surveys. *Journal of the American Planning Association* **74**, 209–221.

Harvey, A., 1990. Time use studies for leisure analysis. *Social Indicators Research* **23**, 309–336.

Hoehner, C.M., Brennan Ramirez, L.K., Elliott, M.B., Handy, S.L., Brownson, R.C., 2005. Perceived and objective environmental measures and physical activity among urban adults. *American Journal of Preventive Medicine* **28**, 105–116.

Jacobs, J., 1961. *The Death and Life of Great American Cities*. New York: Random House

Jelinski, D.E., Wu, J., 1996. The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology* **11**, 129–140.

Kang, B., Moudon, A.V., Hurvitz, P.M., Reichley, L., Saelens, B.E., 2013. Walking Objectively Measured: Classifying Accelerometer Data with GPS and Travel Diaries. *Medicine & Science in Sports & Exercise* **45**, 1419–1428.

Lee, C., Moudon, A.V., 2006. The 3Ds+R: Quantifying land use and urban form correlates of walking. *Transportation Research Part D: Transport and Environment* **11**, 204–215.

Leslie, E., Coffee, N., Frank, L., Owen, N., Bauman, A., Hugo, G., 2007. Walkability of local communities: Using geographic information systems to objectively assess relevant environmental attributes. *Health & Place* **13**, 111–122.

Leslie, E., Saelens, B., Frank, L., Owen, N., Bauman, A., Coffee, N., Hugo, G., 2005. Residents' perceptions of walkability attributes in objectively different neighbourhoods: a pilot study. *Health & Place* **11**, 227–236.

Li, W., Joh, K., Lee, C., Kim, J.-H., Park, H., Woo, A., 2015. Assessing Benefits of Neighbourhood Walkability to Single-Family Property Values A Spatial Hedonic Study in Austin, Texas. *Journal of Planning Education and Research* **35**, 471–488.

Lund, Hollie. 2003. Testing the Claims of New Urbanism. *Journal of The American Planning Association* **69**, 414–449.

Lynch, Kevin. 1960. *The Image of the City*. MIT Press: Cambridge, Mass.

Katz, Peter. 1993. *The New Urbanism: Toward an Architecture of Community*. McGraw-Hill Publishers: Columbus, Ohio.

Kwan, Mei-Po. 2013. Beyond space (as we knew it): Toward temporally integrated geographies of segregation, health, and accessibility. *Annals of the Association of American Geographers* **103**, 1078–186.

Kerr, J., Rosenberg, D., Sallis, J.F., Saelens, B.E., Frank, L.D., Conway, T.L., 2006. Active Commuting to School: Associations with Environment and Parental Concerns. *Medicine & Science in Sports & Exercise* **38**, 787–793.

Marshall, J.D., Brauer, M., Frank, L.D., 2009. Healthy Neighbourhoods: Walkability and Air Pollution. *Environmental Health Perspectives* **117**, 1752–1759.

Mayne, D.J., Morgan, G.G., Willmore, A., Rose, N., Jalaludin, B., Bambrick, H., Bauman, A., 2013. An objective index of walkability for research and planning in the Sydney Metropolitan Region of New South Wales, Australia: an ecological study. *International Journal of Health Geographics* **12**, 61–84.

McCormack, GR., Spence JC., Berry, T., Doyle-Baker, P., 2009. Does perceived behavioral control mediate the association between perceptions of neighbourhood walkability and moderate- and vigorous-intensity leisure-time physical activity? *Journal of physical activity & health* **6**, 657–66.

Millward, H., Spinney, J., Scott, D., 2013. Active-transport walking behavior: destinations, durations, distances. *Journal of Transport Geography* **28**, 101–110.

Mowat, D.L., 2015. Healthy Canada by Design: Translating science into action and prevention. *Canadian Journal of Public Health* **106**, 3–4.

Openshaw, S. and Taylor, P. 1979. A million or so correlation coefficients: three experiments on the modifiable areal unit problem. *Statistical Applications in the Spatial Sciences* **11**, 127-144.

Owen, N., Cerin, E., Leslie, E., duToit, L., Coffee, N., Frank, L.D., Bauman, A.E., Hugo, G., Saelens, B.E., Sallis, J.F., 2007. Neighbourhood Walkability and the Walking Behavior of Australian Adults. *American Journal of Preventive Medicine* **33**, 387–395.

Peiravian, F., Derrible, S., Ijaz, F., 2014. Development and application of the Pedestrian Environment Index (PEI). *Journal of Transport Geography* **39**, 73–84.

Pucher, J., Renne, J.L., 2003. Socioeconomics of urban travel: evidence from the 2001 NHTS. *Transportation Quarterly* **57**, 49–77.

Rodríguez, D.A., Joo, J., 2004. The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D: Transport and Environment* **9**, 151–173.

Saelens, B.E., De Bourdeaudhuij, I., Sallis, J.F., 2003. Environmental correlates of physical activity in a sample of Belgian adults. *American Journal of Health Promotion* **18,** 83–92.

Saelens, B.E., Handy, S.L., 2008. Built Environment Correlates of Walking: A Review. *Medicine & Science in Sports & Exercise* **40**, S550–S566.

Saelens BE, Sallis JF, Black JB, and Chen D., 2003. Neighbourhood-based differences in physical activity: an environment scale evaluation. *American Journal of Public Health* **93**, 1552– 8.

Sallis, J.F., Saelens, B.E., Frank, L.D., Conway, T.L., Slymen, D.J., Cain, K.L., Chapman, J.E., Kerr, J., 2009. Neighbourhood built environment and income: Examining multiple health outcomes. *Social Science & Medicine* **68**, 1285–1293.

Spinney, J.E.L., Millward, H., 2011. Weather impacts on leisure activities in Halifax, Nova Scotia. *International Journal of Biometeorology* **55**, 133–145.

Spinney, J.E., Millward, H., Scott, D., others, 2012. Walking for transport versus recreation: a comparison of participants, timing, and locations. *Journal of Physical Activity and Health* **9**, 153–162.

Terrain Group (WSP Canada Inc.), 2009. *Spring Garden Streetscape: Final Report*. Halifax Regional Municipality

Thompson, B., 1995. Stepwise Regression and Stepwise Discriminant Analysis Need Not Apply here: A Guidelines Editorial. *Educational and Psychological Measurement* **55**, 525–534.

TURP, 2008. *Halifax Regional Space Time Activity Research (STAR) Survey: Data Products*. Saint Mary's University, Time Use Research Program, Halifax.

Ulmer, J.M,Chapman, J.E., Kershaw, S.E., Campbell, M., Frank, L.D., 2014. Application of an evidence-based tool to evaluate health impacts of changes to the built environment. *Canadian Journal of Public Health* **106**, eS26–eS32.

Van Holle, V., De Bourdeaudhuij, I., Deforche, B., Van Cauwenberg, J., Van Dyck, D., 2015. Assessment of physical activity in older Belgian adults: validity and reliability of an adapted interview version of the long International Physical Activity Questionnaire (IPAQ-L). *BMC Public Health* **15**, 1–14.

Villanueva, K., Knuiman, M., Nathan, A., Giles-Corti, B., Christian, H., Foster, S., Bull, F., 2014. The impact of neighbourhood walkability on walking: Does it differ across adult life stage and does neighbourhood buffer size matter? *Health & Place* **25**, 43–46.

Wagner, C.R., Mueller, D.S., 2011. Comparison of bottom-track to global positioning system referenced discharges measured using an acoustic Doppler current profiler. *Journal of Hydrology* **401**, 250–258.

Whittingham, M.J., Stephens, P.A., Bradbury, R.B., Freckleton, R.P., 2006. Why do we still use stepwise modelling in ecology and behaviour?. Stepwise modelling in ecology and behaviour. *Journal of Animal Ecology* **75**, 1182–1189.

Williams, D.M., Matthews, C.E., Rutt, C., Napolitano, M.A., Marcus, B.H., 2008. Interventions to Increase Walking Behavior. *Medicine & Science in Sports & Exercise* **40**, S567–S573.

Wilson, J., Spinney, J., Millward, H., Scott, D., Hayden, A., Tyedmers, P., 2013. Blame the exurbs, not the suburbs: Exploring the distribution of greenhouse gas emissions within a city region. *Energy Policy* **62**, 1329–1335.

Wood, L., Frank, L.D., Giles-Corti, B., 2010. Sense of community and its relationship with walking and neighbourhood design. *Social Science & Medicine* **70**, 1381–1390.

Appendix A

Dependent Variable: Walked Distance by Population (W/P)									
	(Enter Method)	**Unstandardized Coefficients B**	**Standardized Coefficients Beta**	**Sig.**	**t**	**R^2**	**Adjusted R^2**	**F**	**Sig.**
Model 1						0.2	0.171	6.9	0.000
	Constant	4.470							
IV:	Intersection Density	1.914	0.321	0.031	2.201				
	Entropy	1.917	0.321	0.005	2.875				
	Dwelling Density	-1.121	-0.188	0.190	-1.321				
Model 2						0.359	0.328	11.49	0.000
	Constant	4.470							
IV:	Intersection Density	0.144	0.024	0.870	0.164				
	Entropy	1.649	0.277	0.008	2.734				
	Dwelling Density	-2.473	-0.415	0.003	-3.012				
	Retail Lot Coverage Ratio	3.855	0.647	0.000	4.512				
Model 3						0.523	0.48	12.351	0.000
	Constant	13.145							
IV:	Intersection Density	-0.440	-0.074	0.597	-0.53				
	Entropy	1.626	0.273	0.005	2.876				
	Dwelling Density	-2.909	-0.488	0.001	-3.405				
	Retail Lot Coverage Ratio	3.018	0.506	0.001	3.406				
	% Middle Aged	-0.384	-0.482	0.002	-3.182				
	% Older Adult	0.147	0.159	0.076	1.798				
	Average Income	0.085	0.312	0.001	3.408				
	% Young Adults (Excluded)	-	-	-	-				
Model 4						0.62	0.558	10.045	0.000
	Constant	1.887							
IV:	Intersection Density	-0.869	-0.146	0.353	-0.934				
	Entropy	0.379	0.064	0.606	0.518				
	Dwelling Density	-1.615	-0.271	0.080	-1.778				
	Retail Lot Coverage Ratio	2.807	0.471	0.002	3.214				
	% Middle Aged	-0.172	-0.216	0.175	-1.371				
	% Older Adult	0.148	0.160	0.062	1.892				
	Average Income	0.097	0.359	0.000	4.103				
	% Residential Land	-0.011	-0.031	0.775	0.286				
	% Commercial Land	0.007	0.011	0.913	0.11				
	% Institutional Land Use	0.246	0.345	0.002	3.26				
	% Office Land use	0.780	0.321	0.003	3.065				
	% Industrial Land	-0.033	-0.063	0.508	-0.665				
	% Park Land (Excluded)	-	-	-	-				
	% Young Adults (Excluded)	-	-	-	-				
Model 5						0.341	0.309	10.601	0.000
	Constant	11.635							
IV:	WI-4*	0.278	0.154	0.297	1.051				
	% Middle Aged	-0.346	-0.435	0.006	-2.841				
	% Older Adult	0.201	0.218	0.022	2.341				
	Average Income	0.071	0.262	0.011	2.59				
	% Young Adults (Excluded)	-	-	-	-				
Model 6						0.559	0.507	10.823	0.000
	Constant	3.133							
IV:	WI-4*	0.225	0.125	0.470	0.726				
	% Middle Aged	-0.149	-0.187	0.194	-1.311				
	% Older Adult	0.143	0.155	0.061	1.898				
	Average Income	0.102	0.375	0.000	4.147				
	% Residential Land Use	-0.070	-0.201	0.042	-2.064				
	% Commercial Land	0.007	0.010	0.914	0.109				
	% Institutional Land Use	0.243	0.341	0.001	3.356				
	% Office Land use	0.834	0.343	0.001	3.588				
	% Industrial Land	-0.040	-0.076	0.434	-0.786				
	% Park Land (Excluded)	-	-	-	-				
	% Young Adults (Excluded)	-	-	-	-				
	*WI-4: four-variable walkability index								

Dependent Variable: Walked Distance by Road (W/R)									
	(Enter Method)	**Unstandardized Coefficients B**	**Standardized Coefficients Beta**	**Sig.**	**t**	**R^2**	**Adjusted R^2**	**F**	**Sig.**
Model 7						0.307	0.282	12.262	0.000
	Constant	0.947							
IV:	Intersection Density	0.312	0.216	0.159	0.116				
	Entropy	0.389	0.269	2.583	0.012				
	Dwelling Density	0.271	0.187	1.411	0.162				
Model 8						0.422	0.394	14.993	0.000
	Constant	0.947							
IV:	Intersection Density	-0.054	-0.037	-0.266	00.791				
	Entropy	0.334	0.231	2.401	0.019				
	Dwelling Density	-0.009	-0.006	-0.047	0.963				
	Retail Lot Coverage Ratio	0.798	0.551	4.046	0.000				
Model 9						0.553	0.513	13.942	0.000
	Constant	2.97							
IV:	Intersection Density	-0.172	-0.119	-0.881	0.381				
	Entropy	0.336	0.232	2.527	0.014				
	Dwelling Density	-0.116	-0.080	-0.579	0.564				
	Retail Lot Coverage Ratio	0.609	0.420	2.920	0.005				
	% Middle Aged	-0.086	-0.445	-3.033	0.003				
	% Older Adults	0.027	0.120	1.406	0.164				
	Average Income	0.019	0.290	3.271	0.002				
	% Young Adults (Excluded)	-	-	-	-				
Model 10						0.664	0.609	12.167	0.000
	Constant	0.124							
IV:	Intersection Density	-0.227	-0.157	-1.069	0.289				
	Entropy	0.031	0.022	0.186	0.853				
	Dwelling Density	0.218	0.150	1.049	0.297				
	Retail Lot Coverage Ratio	0.565	0.390	2.831	0.006				
	% Middle Aged	-0.030	-0.157	-1.063	0.291				
	% Older Adult	0.029	0.130	1.630	0.107				
	Average Income	0.023	0.350	4.260	0.000				
	% Residential Land	-0.004	-0.043	-0.425	0.672				
	% Commercial Land	0.009	0.055	0.597	0.553				
	% Institutional Land Use	0.060	0.347	3.488	0.001				
	% Office Land use	0.186	0.314	3.194	0.002				
	% Industrial Land	-0.017	-0.133	-1.489	0.141				
	% Park Land (Excluded)	-	-	-	-				
	% Young Adults (Excluded)	-	-	-	-				
Model 11						0.496	0.471	20.164	0.000
	Constant	3.581							
IV:	WI-4*	0.433	0.299	2.326	0.022				
	% Middle Aged	-0.095	-0.49	-3.662	0.000				
	% Older Adult	0.026	0.115	1.419	0.160				
	Average Income	0.016	0.242	2.734	0.008				
	% Young Adults (Excluded)	-	-	-	-				
Model 12						0.639	0.597	15.144	0.000
	Constant	1.939							
IV:	WI-4*	0.526	0.363	2.335	0.022				
	% Middle Aged	-0.052	-2.690	-2.085	0.040				
	% Older Adult	0.016	0.072	0.979	0.330				
	Average Income	0.022	0.340	4.157	0.000				
	% Residential Land Use	-0.013	-0.153	-1.738	0.086				
	% Commercial Land	0.002	0.014	0.160	0.873				
	% Institutional Land Use	0.050	0.287	3.117	0.003				
	% Office Land use	0.145	0.245	2.838	0.006				
	% Industrial Land Use	-0.023	-0.178	-2.032	0.046				
	% Park Land (Excluded)	-	-	-	-				
	% Young Adults (Excluded)	-	-	-	-				

*WI-4: four-variable walkability index

Dependent Variable: Walked Distance by Developed Area (W/DA)									
	(Enter Method)	Unstandardized Coefficients B	Standardized Coefficients Beta	Sig.	t	R^2	Adjusted R^2	F	Sig.
Model 13						0.428	0.407	20.702	0.000
	Constant	0.010							
IV:	Intersection Density	0.008	0.413	3.352	0.001				
	Entropy	0.004	0.204	2.160	0.034				
	Dwelling Density	0.003	0.150	1.248	2.160				
Model 14						0.559	0.538	26.005	0.000
	Constant	0.01							
IV:	Intersection Density	0.003	0.144	1.180	0.241				
	Entropy	0.003	0.164	1.949	0.055				
	Dwelling Density	-0.001	-0.056	-0.488	0.627				
	Retail Lot Coverage Ratio	0.011	0.587	4.940	0.000				
Model 15						0.598	0.562	16.759	0.000
	Constant	0.03							
IV:	Intersection Density	0.00	0.132	1.038	0.302				
	Entropy	0.00	0.181	2.084	0.040				
	Dwelling Density	0.00	-0.133	-1.008	0.317				
	Retail Lot Coverage Ratio	0.01	0.498	3.653	0.000				
	% Middle Age	0.00	-0.279	-2.006	0.048				
	% Older Adults	-9.19	-0.032	-0.398	0.692				
	Average Income	0.000	0.180	2.138	0.036				
	% Young Adults (Excluded)	-	-	-	-				
Model 16						0.729	0.685	16.616	0.000
	Constant	-0.009							
IV:	Intersection Density	0.001	0.079	0.599	0.551				
	Entropy	-0.002	-0.084	-0.810	0.421				
	Dwelling Density	0.003	0.161	1.253	0.214				
	Retail Lot Coverage Ratio	0.009	0.463	3.752	0.000				
	% Middle Aged	4.470	0.018	0.137	0.892				
	% Older Adult	5.480	0.019	0.269	0.789				
	Average Income	0.000	0.253	3.433	0.001				
	% Residential Land	-5.161	-0.048	-0.534	0.595				
	% Commercial Land	0.000	0.138	1.671	0.099				
	% Institutional Land Use	0.001	0.254	2.850	0.006				
	% Office Land use	0.003	0.395	4.470	0.000				
	% Industrial Land	0.000	-0.149	-1.862	0.067				
	% Park Land (Excluded)	-	-	-	-				
	% Young Adults (Excluded)	-	-	-	-				
Model 17						0.540	0.518	24.082	0.000
	Constant	0.032							
IV:	WI-4	0.003	0.547	4.459	0.000				
	% Middle Age	-0.001	-0.295	-2.306	0.024				
	% Older Adult	-4.209	-0.015	-0.190	0.850				
	Average Income	0.000	0.176	2.082	0.040				
	% Young Adults (Excluded)	-	-	-	-				
Model 18						0.684	0.647	18.476	0.000
	Constant	0.016							
IV:	WI-4	0.003	0.613	4.214	0.000				
	% Middle Aged	0.000	-0.117	-0.966	0.337				
	% Older Adult	-9.385	-0.033	-0.477	0.635				
	Average Income	0.000	0.272	3.556	0.001				
	% Residential Land Use	0.000	-0.157	-1.911	0.060				
	% Commercial Land	7.171	0.034	0.431	0.667				
	% Institutional Land Use	0.000	0.151	1.758	0.083				
	% Office Land use	0.002	-0.185	-2.256	0.027				
	% Industrial Land Use	0.000	0.301	3.723	0.000				
	% Park Land (Excluded)	-	-	-	-				
	% Young Adults (Excluded)	-	-	-	-				

* WI-4: four-variable walkability index