



FINAL REPORT

Transferability & Forecasting of the Pedestrian Index of the Environment (PIE) for Modeling Applications

NITC-RR-1028 ■ March 2019

*NITC is a U.S. Department of Transportation
national university transportation center.*



TRANSFERABILITY & FORECASTING OF THE PEDESTRIAN INDEX ENVIRONMENT (PIE) FOR MODELING APPLICATIONS

Final Report

NITC-RR-1028

by

Kelly J. Clifton, PhD¹ (PI)

Jaime P. Orrego-Onate¹

Patrick A. Singleton, PhD²

Robert J. Schneider, PhD³

¹Portland State University

²Utah State University

³University of Wisconsin – Milwaukee

for

National Institute for Transportation and Communities (NITC)

P.O. Box 751

Portland, OR 97207



March 2018

Technical Report Documentation Page

1. Report No. NITC-RR-1028	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Transferability & Forecasting of the Pedestrian Index of the Environment (PIE) for Modeling Applications		5. Report Date March 2019	
		6. Performing Organization Code	
7. Author(s) Kelly J. Clifton; Jaime P. Orrego-Oñate; Patrick A. Singleton; Robert J. Schneider		8. Performing Organization Report No.	
9. Performing Organization Name and Address Portland State University, Department of Civil & Environmental Engineering PO Box 751 - CEE Portland, OR 97207-0751		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No. NITC-RR-1028	
12. Sponsoring Agency Name and Address National Institute for Transportation and Communities (NITC) P.O. Box 751 Portland, Oregon 97207		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplementary Notes			
<p>16. Abstract</p> <p>This project focuses on making our measures, models, and methods more transferable to other locations. Specifically, we re-evaluate, compare and test our pedestrian index of the environment (PIE) measure using data resources more commonly available to planning agencies across the country. Next, we test the results of PIE and its input data in models of pedestrian mode choice for stability of estimation results within a region (intra-regional) and between regions (inter-regional). This research is the next logical step in the MoPeD's enhancement and is critical to enabling its utility beyond the Portland region.</p> <p>The results of this project show that population density and pedestrian connectivity had the most consistent and strong relationship to walk mode choice across all of our regions, which echoed the long literature on this topic. However, the other components of the built environment included in PIE had more variability in their ability to explain walk mode choice. Employment density and its subset urban living infrastructure (ULI), intended to capture retail and service access, had less explanatory power and stability in the cities tested. Based upon these findings, we provide several guidelines for the construct of walkability indices, including variables and spatial scales.</p> <p>Our findings raise questions about the relationship between walking and the built environment within a region and thus, the intra-regional transferability of one walkability index is suspect. Estimation results suggest that there may be different responses to the built environment in lower-density vs. higher density regimes and that these relationships may be nonlinear. However, smaller sample sizes of travel survey data in high density areas in all of the US cities tested pose limitations to drawing more confident conclusions from these results.</p> <p>The inter-regional comparisons of PIE and walk mode share between Los Angeles and Portland showed promise for the use of the index in different regions. In these two regions, model results showed a similar walk mode share for the same values of PIE constructed at the block group level. This provides initial support that the PIEbg construct may be transferrable between metropolitan regions, in part, due to population density's prominent role in PIE.</p>			
17. Key Words Walking, pedestrians, travel demand forecasting, four-step models		18. Distribution Statement No restrictions. Copies available from NITC: www.nitc-utc.net	
19. Security Classification (of this report) Unclassified	20. Security Classification (of this page) Unclassified	21. No. of Pages 33	22. Price

ACKNOWLEDGMENTS

This project was funded by the National Institute for Transportation and Communities (NITC) and the City of Tigard, Oregon. The authors thank colleagues from Metro, Portland State University, University of Wisconsin, Milwaukee, Polytechnique Montréal, Technical University of Munich, and the Metropolitan Council, St. Paul, MN, for their insights and interest in this topic.

DISCLAIMER

The contents of this report reflect the views of the authors, who are solely responsible for the facts and the accuracy of the material and information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation University Transportation Centers Program and Portland State University in the interest of information exchange. The U.S. Government and Portland State University assumes no liability for the contents or use thereof. The contents do not necessarily reflect the official views of the U.S. Government and Portland State University. This report does not constitute a standard, specification, or regulation.

RECOMMENDED CITATION

Clifton, Kelly J; Orrego-Oñate, Jaime; Singleton, Patrick; and Schneider, Robert. *Transferability & Forecasting of the Pedestrian Index of the Environment (PIE) for Modeling Applications*. NITC-RR-1028. Portland, OR: Transportation Research and Education Center (TREC), 2019.

TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
1.0 INTRODUCTION	2
2.0 BACKGROUND	5
3.0 PEDESTRIAN INDEX OF THE ENVIRONMENT (PIE).....	6
4.0 DATA AND METHODS FOR NEW PIE (PIE _{BG})	9
5.0 TESTING THE TRANSFERABILITY OF PIE _{BG}	12
5.1 COMPARING PIE _{BG} AND PIE ₀	12
5.2 INTRAREGIONAL MEASURE TRANSFERABILITY	18
5.3 INTERREGIONAL MEASURE TRANSFERABILITY.....	22
6.0 DISCUSSION AND CONCLUSIONS	28
6.1 TRANSFERABILITY OF MODELS AND MEASURES.....	28
6.2 REPRESENTING THE BUILT ENVIRONMENT	30
6.3 FUTURE WORK.....	31
7.0 REFERENCES	32

LIST OF TABLES

Table 1 Binary Logit Models for PIE ₀	8
Table 2 PIE _{BG} estimation results: binary logit models of walk mode share	14
Table 3 Estimated weights for each built environment measure for PIE ₀ and PIE _{BG}	14
Table 4 Pooled vs. Urban/Suburban Walk Mode Share Models	21
Table 5 Trip information and built environment measures in each metropolitan area.....	23
Table 6 Unscaled PIE _{BG} coefficients estimated from Portland data.....	26

LIST OF FIGURES

Figure 1 Regional Map of the Built Environment Index PIE ₀	9
Figure 2 Comparison between block groups (bold line) and PAZ (grid) in downtown Portland, OR.....	11
Figure 3 Distribution of the scaled attributes of the built environment for PIE _{BG} in Portland, OR13	13
Figure 4 Frequency distribution of PIE _{BG} scores by block group	15
Figure 5 Land percentage by PIE _{BG} score	16
Figure 6 PIE _{BG} distribution in Portland	17
Figure 7 PIE ₀ versus PIE _{BG}	18
Figure 8 PIE _{BG} aggregated by deciles vs walking share	19
Figure 9 Predicted walk probabilities	22
Figure 10 Standardized coefficients for different built environment measures.....	24
Figure 11 Unscaled PIE _{BG} distribution in Los Angeles and Portland	27
Figure 12 The role of walkability gradients.....	29

EXECUTIVE SUMMARY

There have been important advances in non-motorized planning tools in recent years, including the development of the MoPeD pedestrian demand model (Clifton et al., 2013, 2015). This tool and others are increasingly requested by governments and agencies seeking to increase walking activity and create more walkable places. To date, the MoPeD tool has been piloted with success in the Portland region using data unique to Metro, the metropolitan planning organization. However, there is increasing interest from planning agencies in adapting the pedestrian modeling tools and their inputs for use in their own jurisdictions. Unfortunately, other regions often do not have uniform access to the same kinds of pedestrian environment data as Metro, particularly at such a fine-grained scale.

In this next phase of our pedestrian modeling work (see Clifton et al., 2013, 2015), this project focuses on making our measures, models, and methods more transferable to other locations. Specifically, we will re-evaluate, compare and test our pedestrian index of the environment (PIE) measure using data resources more commonly available to planning agencies across the country. Next, we test the results of PIE and its input data in models of pedestrian mode choice for stability of estimation results within a region (intraregional) and between regions (interregional). This research is the next logical step in the MoPeD's enhancement and is critical to enabling its utility beyond the Portland region.

In terms of index inputs, the results of this project show that population density and pedestrian connectivity had the most consistent and strong relationship to walk mode choice across all of our regions, which echoed the long literature on this topic. However, the other components of the built environment included in PIE had more variability in their ability to explain walk mode choice. Employment density and its subset urban living infrastructure (ULI), intended to capture retail and service access, had less explanatory power and stability in the cities tested. Based upon these findings, we provide several guidelines for the construct of walkability indices, including variables and spatial scales.

Our findings raise questions about the relationship between walking and the built environment within a region and thus, the intraregional transferability of one walkability index is suspect. Estimation results suggest that there may be different responses to the built environment in lower-density vs. higher density regimes and that these relationships may be nonlinear. However, smaller sample sizes of travel survey data in high density areas in all of the US cities tested pose limitations to drawing more confident conclusions from these results.

The interregional comparisons of PIE and walk mode share between Los Angeles and Portland showed promise for the use of the index in different regions. In these two regions, model results showed a similar walk mode share for the same values of PIE constructed at the block group level. This provides initial support that the PIEbg construct may be transferrable between metropolitan regions, in part, due to population density's prominent role in PIE.

1.0 INTRODUCTION

There is a growing demand for tools that can estimate pedestrian activity, forecast future walking levels, and be used for other transportation planning applications. Many metropolitan planning organizations (MPOs), in particular, seek tools to estimate pedestrian demand that can provide greater support for planning and policy-making activities, including: increased sensitivity to pedestrian facilities, land use changes, and the built environment; better accounting for mode shifts and resulting changes in greenhouse gas emissions; more accurate estimates of physical activity levels for health impact assessments and of pedestrian exposure for traffic safety analyses; etc. While not yet widespread, regional travel demand forecasting for pedestrians continues to develop (Kuzmyak et al. 2014; Singleton et al. 2018); however, different MPOs have used various ad-hoc techniques that depend on their unique decision support needs, data sources, and modeling capabilities, with little consistency. Other agencies lack the technical capacity and funding to undertake model improvements on their own. Facing this reality, travel model transferability is of increasing interest (Rossi and Bhat 2014). A framework for modeling pedestrian demand with consistent data inputs that can be more easily transferred between regions would be of great value to practitioners, especially those at agencies with limited resources.

Despite a long history of research documenting relationships between walking and environmental conditions—including effects of the built environment—models used in practice lag in their representation of the pedestrian environment. Pedestrian travel has been positively related to higher residential and employment densities, greater land use mix or diversity, more connected networks or higher intersection densities, greater accessibility to transit, and sidewalk and crossing conditions (e.g., Saelens and Handy 2008). Yet, most MPO regional models use only rough density measures to predict pedestrian demand (Singleton et al. 2018). While the lack of pedestrian-sensitive built environment data was a barrier in the past, today fine-grained archived spatial datasets are becoming more widely available, including point-, parcel-, or block-level measures of the pedestrian environment. This has led to the development of a number of measures of the pedestrian environment, including the Pedestrian Index of the Environment, (Singleton et al. 2014), the Pedestrian Environment Factor (Greenwald and Boarnet 2001), the Pedestrian Environment Index (Peiravian et al. 2014), and others.

The present-day issue is that the availability and formatting of these built environment data vary widely across and within regions, yielding challenges to the transferability of composite measures of the pedestrian environment (and models of pedestrian demand). A related issue involves finding the appropriate scale that balances tradeoffs between the use of small-scale and behaviorally-sensitive pedestrian environment data, the computational processing challenges that such data impose, and the general availability of those data now and in the future.

Pedestrian modeling has also not kept pace with nor taken full advantage of developments happening in pedestrian data collection. Historically, one reason why few MPO models represented pedestrian travel was that the regional household travel surveys used to estimate such models captured few walking trips, often by design (Clifton and Muhs 2012). Today, surveys have improved their capture of pedestrian activity, and data collection efforts are increasingly being supplemented by short- and long-term pedestrian counts (Schneider et al.

2005). Surprisingly, these count data are rarely used in the development of regional pedestrian demand forecasting models and analysis tools, despite being a potential source of data for external validation, if not calibration.

This project begins to address some of these research gaps and practical needs. It builds upon previous work developing a framework for Model of Pedestrian Demand (MoPeD) and a Pedestrian Index of the Environment (PIE) variable (Clifton et al. 2013, 2015) to make measures and models more transferrable and useful for forecasting.

PIE is a composite index that is constructed from a set of models that estimate the probability of walking trips as a function of a built environment attribute. The value of PIE is calculated from a set of built environment variables representing activity density, block density, sidewalk density, transit access, neighborhood-oriented businesses and other factors (Singleton et al. 2014). PIE was originally developed for Portland, OR using the 2011 Oregon Household Travel Survey (OHAS) and built environment data at a small grid scale of 80 by 80 meters.

MoPeD is a tool for predicting pedestrian demand using trip generation, mode choice, and destination choice models. These models predict walking trip probabilities for a variety of travel purposes using PIE and other socioeconomic and demographic information. MoPeD has been calibrated with OHAS data and applied to 80 by 80-meter grid cells in Portland, OR.

These tools and others are increasingly requested by those interested in increasing walking activity and creating more walkable places. While MoPeD and PIE have been piloted with success in the Portland region using data unique to Metro, the metropolitan planning organization, there is increasing interest from planning agencies within and outside of Portland and Oregon (e.g.: City of Tigard, OR; Metropolitan Council of the Twin Cities, MN; San Francisco Public Health Department, CA) in adapting these pedestrian modeling tools for use in their own jurisdictions. Local governments desire to apply these tools for a variety of planning and forecasting purposes, including but not limited to regional demand modeling. Unfortunately, other regions often do not have uniform access to the same kinds of pedestrian environment data as Metro, particularly at such a fine-grained scale. Important challenges remain in model development that must be overcome if these tools are to achieve widespread application. Among the most critical needs are the standardization and forecasting of model inputs, particularly measures of the built environment.

In this next phase of our pedestrian modeling work, we focus on testing the ability of our measures, models, and methods to be more transferable to other locations. Specifically, we capitalize on the increased availability of pedestrian environment data to create a new PIE (called PIE_{bg}) that utilizes more widespread data sources available at the census block group level. We re-evaluate, compare (with the original PIE, called PIE_0 in this report) and test our PIE_{bg} measure using data resources more commonly available to planning agencies across the country. Next, we examine this PIE_{bg} variable and its association with walking in Portland amid both urban and suburban contexts (intra-regionally). Additionally, we examine differences in the PIE_{bg} measure and its performance in regions outside of Portland (interregional transferability). These tasks balance data availability, scale, computational capacity, and behavioral realism. This updated suite of pedestrian modeling products has the potential to be more widely transferable and applicable to MPOs and other planning agencies beyond Portland. Finally, we test this new

measure across a large range of urban environments by estimating models of probability of walking using PIE_{bg} as a single explanatory variable. This analysis tests the issue of linearity in the parameters and questions the validity of using a single parameter to represent the behavioral response to walking environment for all locations.

Our overall research objective is to increase the availability of pedestrian demand tools for use by various planning agencies around the country. To do this, we are guided by the following research questions.

How can PIE be constructed using data that are widely available to planning agencies across the country, but remain sensitive to pedestrian scales and variations in environmental conditions? PIE is central to representing the built environment in the trip generation and destination choice models within MoPeD. In its current form, PIE relies on detailed, disaggregate, spatial data that are available in the Portland region; but, in other areas of the country, these data are not uniformly available. In order for the MoPeD model to be transferable to other locations, we redesign PIE to make use of data resources that are more commonly available around the US and elsewhere and compare results with our original PIE construct.

Should the measures of the built environment comprising PIE be weighted differently in different urban/suburban contexts? People walking in suburban locations may actually care about and respond to various environmental characteristics differently than people walking in urban locations. Suburban areas have been given relatively less attention in studies of pedestrian behavior, despite evidence of significant levels of walking observed there (Larco et al. 2013). To better understand walking behavior across a spectrum of contexts, we test for differences in the combinations of environmental features and thresholds used in PIE between urban and suburban locations for Portland. Additionally, we extend these tests to different cities to assess how PIE behaves and if we can find evidence that, the measurements are transferrable.

The remaining report is organized into five additional sections. The Section 2.0 provides an introduction to the pedestrian model. Then, a brief summary of the background of the project, a review of the main issues about walking and travel behavior, and a synthesis of the transferability measures are provided in Section 3.0. The next section, Section 4.0 is a description of the data inputs and outputs, and discusses the main assumptions in the analysis, and reviews the construction of the data sets including the built environment and travel data. Section 5.0 covers the research methods and the implications of our approach and presents results. It is separated in four subsections: the reconstruction of our new index (PIE_{bg}), the comparison with the older index (PIE_0), and two sections analyzing the intra- and intercity transferability of our new index. Finally, Section 6.0 is a discussion and conclusions of this project with suggestions for policy and future research.

2.0 BACKGROUND

The relationship between walking and the built environment has had a particular emphasis on the study of travel behavior. In 1997, Cervero and Kockelman proposed three attributes to characterize the built environment: diversity, design, and density (the D's). Analyzing different metro areas from the US, they found that relationship between these variables and travel behavior are modest to moderate. Although they did not find strong relationships for each variable, they did suggest that the synergy among the three could cause impacts that are more appreciable. Other studies, focusing explicitly on walking, have found a stronger association with the built environment. Two reviews (Handy et al. 2002; Frank 2000) found positive and significant relationships between density and mixed land use with walking. In a 2006 review, Sallis et al. found that studies have yielded surprisingly consistent results but that utilitarian walking (and cycling) is generally higher in the areas with mixed uses, good street connectivity, and higher population densities. In another 2008 review, Saelens and Handy (2008) found more consistency among these same relationships even after controlling for self-selection. However, they found evidence in some studies indicating that the random samples of travel observations taken from the general population do not ensure sufficient variation and that multicollinearity among built environment characteristics makes it hard to identify the unique contribution of a specific attribute.

None of the studies acknowledges any location bias in the selection of study participants. Specifically, there is a lack of variation in the overall urban structure and range of built environments tested in the literature. As noted by Saelens and Handy, the majority of studies are based on a small number of countries including the US. Overall residential densities in the US are the lowest in the world (Newman 2014; Huang et al. 2007). Additionally, the low level of compactness (Huang et al. 2007) adds the question of whether the density gradients have a role. Furthermore, having low residential densities on average means higher density areas may be scarce, thus, making them “hidden populations” or “hard-to-reach populations” for gathering travel behavior data from surveys. This means that the sampling process has to be very large to have a good accuracy level, or that places with a desirable density for walking may be nonexistent in many cases, especially in smaller size cities or towns. Thus, empirically-derived relationships between walking and the built environment using only lower-density observations may not apply to higher-density places or regions. A study of differences between major cities and rural small towns found that smaller towns did not have the nearby destinations to support walking, compared with city centers of major metropolitan areas that did have close destinations (Stewart et al. 2016).

Another significant approach to determine the degree to which built environment characteristics affects travel behavior has been the calculation of elasticities from a meta-analysis or meta-regressions (Ewing and Cervero 2010; Stevens 2017). This has been done for different cities and samples, with estimated elasticities of built environment variables ranging 0.0–0.3, suggesting an inelastic relationship with walking. Further discussions have focused on establishing the relative vs. absolute magnitude of this number (is it large or not?) but have neglected the potential complementarity among those variables characterizing the built environment. Besides, this approach avoids the discussion of multicollinearity in the characteristics of the built environment. The latest debates have centered on the usefulness of these methods (Clifton 2017;

Handy 2017; Ewing and Cervero 2017), as they have become common among different fields. The main concerns regarding these methods are the absence of longitudinal studies as well as the assumptions of linearity behind them, as there may be evidence that is not the case (Talen and Koschinsky 2014).

Although these issues may be responsible for mixed results in studies of the relationships between walking and the built environment, there is extensive evidence that these relationships exist. Nevertheless, there is no consensus around the level or magnitude of effects and suggests that findings may have limited applicability in other locations. This project tries to fill these gaps by building and testing new indicators of the built environment from data sources that are more available and from there test how transferable are the measures among cities.

Fox et al. (2014) identify a clear distinction between two types of model transferability: temporal transferability that refers to the validity of the models for long-term predictions, and spatial transferability that refers to the validity of models across different regions. For both types of transferability, there are methods to test model transferability (Ben-Akiva 1981; Lerman 1981) and approaches to updating the model parameters (Ben-Akiva and Bolduc 1987; Badoe and Miller 1995). However, the focus in the literature has been on identifying the statistical methods to transfer models and testing the accuracy of the prediction, rather than understanding the behavioral responses under different contexts.

For example, Nowrouzian and Srinivasan (2012) test spatial transferability of activity-based models in Florida and identify a gap in the study of transferability in this types of models. They found that transferability is limited and that further research must be done to identify the range of contexts in which the model can be applied. Karasmaa (2007) tests the spatial transferability of the simple four-step regional model in two different urban areas in Finland. The research shows that the models are largely transferable between these locations; however, doubts arise as to whether this evidence supports the idea of more universal model transferability or is limited to similar urban structures and cultural contexts. Another limitation in the transferability literature is that regional models focus more on motorized vehicles than walking and biking (see Singleton et al. 2018) and thus a specific examination of the transferability of relationships between the built environment and walking behavior is lacking. Thus, this project will try to fill this gap to test if there are systematic behavioral responses across and within different cities.

3.0 PEDESTRIAN INDEX OF THE ENVIRONMENT (PIE)

This study builds on our previous work by examining and improving upon the Pedestrian Index of the Environment (PIE), developed and described in the project reports (Clifton et al. 2013, 2015). In order to understand this study, it is important to give a brief overview of our original concept of PIE, referred to as PIE₀ in this report. PIE₀ is an index comprising six built environment measures, listed below, and was computed for each PAZ in the system (N= 1,465,252 in the four-county metro area of Portland, OR).

- comfortable cycling facilities; a proxy for low volume streets with traffic calming (local density of weighted bicycle network links in a one-mile buffer of each PAZ (1.6 km));
- block size (average block size in a quarter-mile (0.4 km) buffer of each PAZ);
- people per acre (population plus employment density in a quarter-mile (0.4 km) buffer of each PAZ);
- sidewalk density (total length of metro sidewalk inventory within a quarter-mile buffer of each PAZ);
- transit access (a density measure of transit stops weighted by the service frequency in a quarter-mile buffer); and
- urban living infrastructure (ULI) (the number of services, entertainment, and retail services in a quarter-mile buffer of the origin PAZ).

In the calculation of the index, each of the built environment measures is weighted based upon their relative importance to walking behavior (see Clifton et al., 2013 for more information). To calculate these weights, we used the standardized coefficients from a univariate model that predicts the likelihood of walking as a function of each of the six built environment measures for the PAZ of the trip origin.

For each built environment attribute j , a univariate (walk/don't walk), binary-logit mode choice model β_j is estimated to predict walking/other mode using as an explanatory variable the attribute score z_j . Each model calibration delivers a utility function:

$$z_j = \beta_j + \beta_{1j} z_j$$

After this, PIE_0 was constructed by weighting each built environment variable by the corresponding estimated coefficient β_{1j} and summing the values over a specific spatial unit.

$$PIE_0 = \sum_j \beta_{1j} z_j$$

The dataset used to provide the built environment characteristics used in PIE_0 is the Context Tool, developed by Portland Metro. This dataset calculates the built environment characteristics within 80-meter by 80-meter grid cells (called Pedestrian Analysis Zones, or PAZs) and associates each raw measure with discrete values between 1 and 5. These discrete values were assigned using natural breaks in the distribution of each measure. Additionally, the travel behavior data used in the estimation of the binary logit models for the weights used in PIE_0 were the Oregon Household Travel Survey (OHAS) from 2011. The sample consisted of 56,634 trips for a regular day, 36,463 trips of which correspond to the area where Portland Metro Context Tool had measured.

The coefficients for PIE_0 are shown in Table 1 and the values for PIE_0 for the Portland region are shown in Figure 1. The advantages of PIE_0 were that it captured fine-grained variations in the walking environment at the PAZ scale (80-meters by 80-meters). This improves upon the coarse scale used in most regional travel models, where Traffic Analysis Zones (TAZs) are the unit of analysis. Additionally, this index weighted the different built environment attributes by their association with walking behavior and improves our characterization of “walkability”. Thus, PIE_0 was statistically significant in our later analysis of travel behaviors (mode choice and trip generation). In addition to the contributions made by PIE_0 , MoPeD can be used as a stand-alone demand tool or in concert with Metro's regional travel model. Despite these improvements, the methodology still has some limitations. First, the data sources used for the model estimation used discrete values between 1 and 5 using natural breaks to transform the data rather than the direct values of the built environment variables. This presents problems in transferring the measures and the models to other places since each location will have its own distributions of these variables (and different corresponding natural breaks).

Table 1 Binary Logit Models for PIE_0

Context Variable (z_n)	Coefficient (β)	p-value	Model pseudo- R^2
Model 1			
Bicycle Access	0.494	0.00	0.057
Constant	-4.047	0.00	
Model 2			
Block Size	0.543	0.00	0.096
Constant	-3.729	0.00	
Model 3			
People per Acre	0.812	0.00	0.095
Constant	-4.304	0.00	
Model 4			
Sidewalk Density	0.500	0.00	0.083
Constant	-3.900	0.00	
Model 5			
Transit Access	0.621	0.00	0.083
Constant	-3.386	0.00	
Model 6			
ULI Density	0.549	0.00	0.073
Constant	-3.204	0.00	
Data used for all models			
Trips (n)	36,463		
Walk	3,560		
Not Walk	32,903		

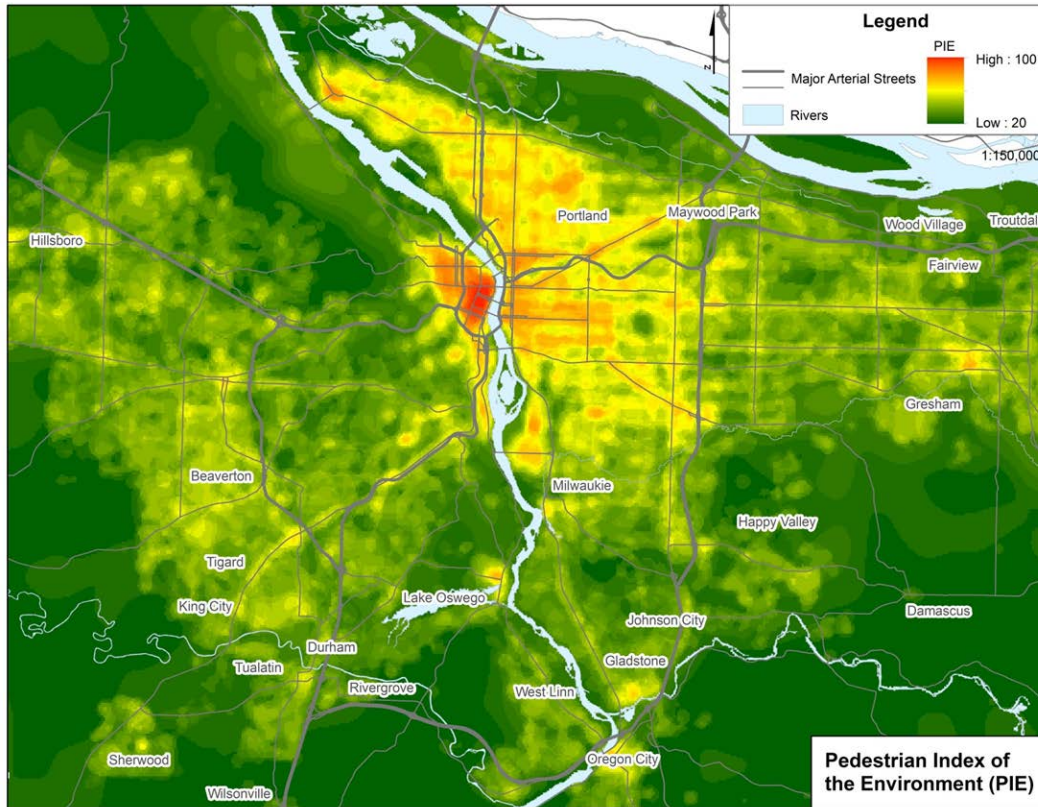


Figure 1 Regional Map of the Built Environment Index PIE_0

4.0 DATA AND METHODS FOR NEW PIE (PIE_{BG})

Our objective at large was to re-create pedestrian-specific measures of the built environment at a coarser scale (for the purpose of using more readily available data nation-wide), compare them with PIE_0 , and test whether their predictive capacity in models of walking mode choice in Portland, OR.

The extension to this methodology was to estimate a new index that we called PIE_{bg} , named for the spatial scale of its construct: The Census block group geographic unit. There is a total of 1,324 block groups with an average area of 11.4 sq. meters (SD = 111.1 sq. meters).

Figure 2 shows the difference between both units of analysis. The block group scale is similar to the traffic analysis zones (TAZ) commonly used in transportation planning methods. PIE_{bg} was calibrated using an analogous methodology as PIE_0 . One of our objectives was to test the loss of power of our models in changing the scale and the data source, so we framed the PIE_{bg} score to be similarly constructed as PIE_0 .

In PIE_0 the built environment data were scaled using natural breaks and assigned discrete values between 1 and 5. However, in PIE_{bg} , we decided that this method lose too much information and

the induce error could even be higher at the larger scale of the block group, particularly when using a block group scale and thus, continuous values were assigned. For that reason, we scaled the data to be between 1 and 5 using continuous breaks. This is done by using the following formula:

$$z_{new} = \frac{5 - 1}{\max(Z) - \min(Z)} * (z_{old} - \min(Z)) + 1,$$

where z is the value that wants to be transform and the max and min values correspond to the maximum and minimum value of the built environment data set for each city.

After this, we used the same methodology of PIE_0 where the weights of each attribute of the built environment are the coefficients of univariate logit models of the choice to walk estimated using household travel survey data for each region. The new indicator PIE_{bg} is again transformed by multiplying with a constant to have a minimum value of 20 and a maximum value of 100 as with PIE_0 .

$$constant = \frac{20}{\sum \beta_{1j} * \min z_j}$$

One of the objectives of this project was to use widely-available data in order to make our measures of PIE transferrable. The US Environmental Protection Agency (EPA) created a unique dataset called the Smart Location Database (SLD) (Ramsey and Bell 2014; US EPA 2014) that includes information from the Census, the street design from Navteq’s NAVSTREETS street data, and transit data from each local agency, in General Transit Feed Specification (GTFS), that include all the time schedule and transit routes. This data source was selected because it is a resource that offers a large number of built environment variables, measured using equivalent methods, and it covers the entire country at a relatively small spatial scale. The Census block groups are clusters of blocks with a population between 600 and 3,000 people. Although pedestrian behavior is associated with a broad array of built environment characteristics, the following were selected based upon their availability in the SLD database (and thus their availability in all of our US study locations), their strong association with walking, and similarity to that used in PIE_0 .

- People per acre: It is the addition of population and employment density. Population density is the gross population density of people per acre using the population of the decennial 2010 Census and the unprotected area defined in the Census Block Group. Employment density is calculated using the Census Longitudinal Employer-Household Dynamics (LEHD) data and dividing it by the unprotected area defined in the Census Block Group in acres.
- Urban living infrastructure (ULI): Using the same data from employment density we filtered to include only commercial and entertainment employment. However, unlike the original ULI measure, which is a count of these establishments within a quarter-mile buffer, this ULI measure is a density measure. While they are both effectively describing density, the coefficients from these two different ULI measures cannot be directly compared because of their different construction.

- Transit accessibility: Aggregate frequency of all transit in a buffer of 0.25 miles of the block group in the peak hour using GTFS data from the local transit agencies.
- Total road network density: The total miles of roadway per square mile. Highways that are divided are counted as two separate roadways.

These variables in this new index differ from those used in the original PIE_0 in two ways. First, our original measures of connectivity, block size, was substituted with total road network density. The reason for this is partially convenience: it is not available in the SLD dataset. The other alternative, intersection densities, were poorly defined in the SLD database making it difficult to interpret. Additionally, Berrigan et al. (2010) showed that intersection density and network density have almost a 90% of correlation, making them very similar. Therefore, we used total road network density to assess connectivity, which also gives the best fit in the model.

Second, we did not include measures of sidewalk density and comfortable cycling facilities in our new construct. This omission was justified as both measures are not widely available. Further, the comfortable cycling facilities variable itself has a weak theoretical association with pedestrian travel. It was originally incorporated in PIE_0 a proxy measure for the degree to which streets accommodated multimodal travel. Furthermore, the primary motivation for this exercise is to test how transferable are the relationships in PIE are across different locations and not necessarily to estimate a model to use for forecasting.

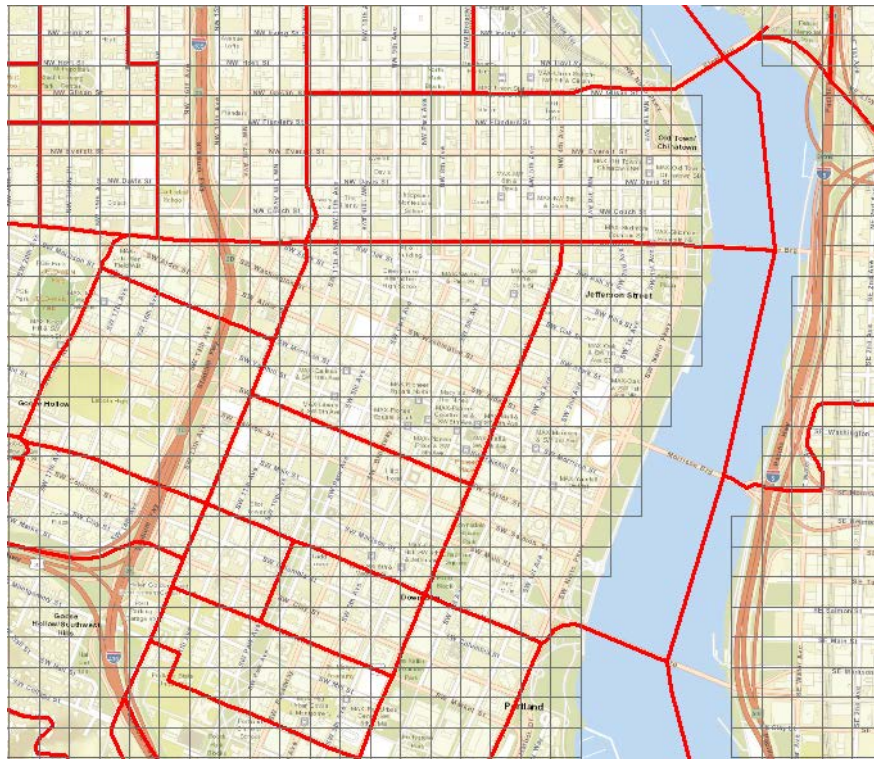


Figure 2 Comparison between block groups (bold line) and PAZ (grid) in downtown Portland, OR

The next step in the process was to construct the built environment measures from the SLD data. The dataset, as explained previously, includes the residential density and employment density, so those measures were directly incorporated. To create ULI, we used commercial and retail employment density measures that also was a direct calculation. The total road network and the transit aggregated frequency were defined directly in the database so no calculation was needed.

In synthesis, PIE_{bg} is estimated with the same method of aggregating the univariate coefficients of the four variables described above. This process is as follows:

- a. Rescale the built environment variables from the SLD values to be a continuous measure between 1 and 5, with 1 representing the group of lowest values and 5 the highest.
- b. Do a spatial join of the Oregon Household Survey trips origins data with the SLD. In this way, we manage to get a data set of each trip with the associated data of the block group at the trip origin.
- c. Estimate individual univariate binary logit models of walking mode choice as a function of each of the built environment variables.
- d. Take the coefficients of each model and generate a PIE_{bg} value for the block group level with the weights. The coefficients are multiplied by a constant so that PIE_{bg} has a minimum value of 20 and a maximum value of a 100. For this task, we replicate the same method used in PIE_0 .

The final PIE_{bg} score is calculated using the following formula.

$$PIE_{bg} = 4.6 * \textit{people per acre} + 4.8 * ULI + 4.7 * \textit{transit accessibility} + 6.1 * \textit{street network density}$$

5.0 TESTING THE TRANSFERABILITY OF PIE_{BG}

This is section we present the results of our testing, comparison, and analysis of both walkability measures PIE_0 and PIE_{bg} constructed at different scales and with different data inputs. First, we compare our both walkability measures. Then, we test the relationship between PIE_{bg} and walking within the Portland region to investigate whether one walkability measure is suitable for all urban contexts within one region. Or said differently, we are testing whether there is a different behavior response to PIE in urban vs. suburban areas. Finally, we introduce a more extensive analysis that includes different cities in the US. Our aim here is to test the inter-regional transferability by examining how these measures perform in different metropolitan areas of the US.

5.1 Comparing PIE_{bg} and PIE_0

To compare the new measures of the walkability PIE_{bg} constructed at the block group scale with nationally available data to the original measure PIE_0 at a finer scale with region-specific data, we needed to compute PIE_{bg} for the Portland region. The distributions of various data used to

construct PIE_{bg} in Portland are shown in Figure 3. ULI, shown in orange, has a highly positive skew, meaning that there are very few block groups with scores above 2 and most of the block groups have low ULI score with a score of 1. The people density measure, shown in green, is less skewed, but also shows a large concentration of low values. This means that high values of both of these measures are scarce in the total composition of the region. Transit access, shown in pink, is also skewed to be high in very few places. The total road network density shown in purple is the only measure that it distributed more evenly across all of the values. Overall, this distribution makes sense for a region such as Portland, where the total land area is dominated by suburban landscapes and large park reserves with only a few highly urbanized nodes. This supports the notion that the landscape of North American cities may not have enough locations with sufficient densities to support non-automobile travel.

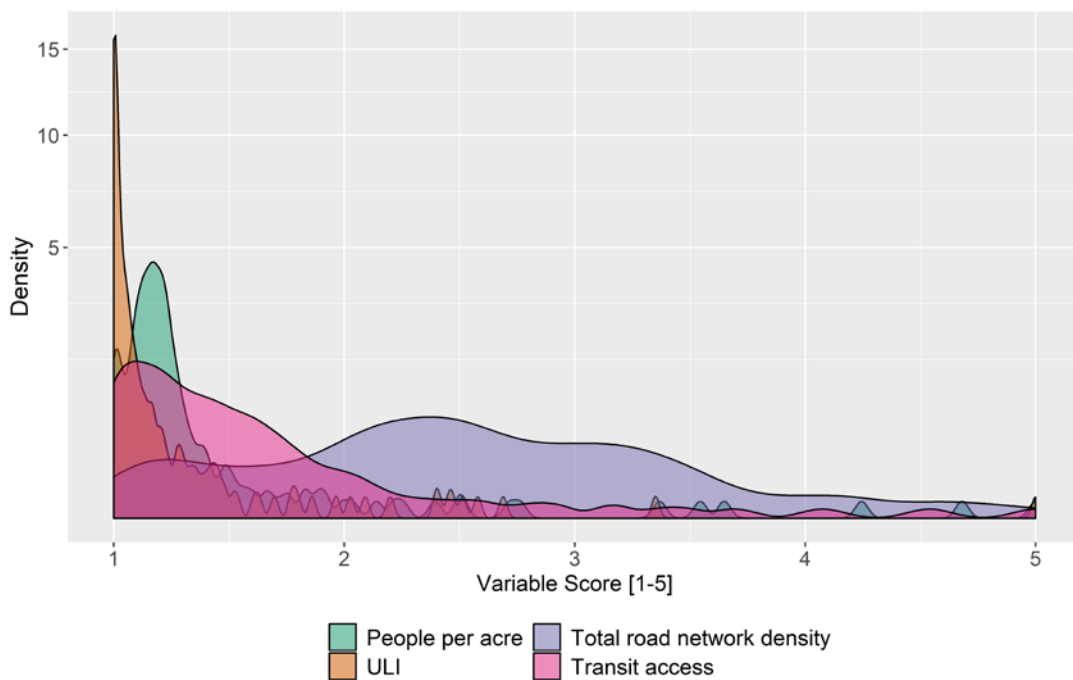


Figure 3 Distribution of the scaled attributes of the built environment for PIE_{bg} in Portland, OR

Using these data inputs, PIE_{bg} was computed using the process outlined in the previous section. The resulting coefficients from the univariate binary logit models of walk mode choice for each built environment measure are shown in Table 2. The estimated coefficients range between 0.5-0.7 with the total road network density showing the highest coefficient value. These coefficients are used to weight each individual measure in the construction of PIE, after scaling to give PIE a range of values between 20 and 100.

Table 2 PIE_{bg} estimation results: binary logit models of walk mode share

Built Environment variable (z_n)	Coefficient (β)	p-value	Model pseudo-R ²
Model 1			0.03
People (employment and residential) per acre [1-5]	0.52	0.00	
Constant	-3.01	0.00	
Model 2			0.02
ULI [1-5]	0.54	0.00	
Constant	-2.94	0.00	
Model 3			0.05
Road network density [1-5]	0.69	0.00	
Constant	-4.12	0.00	
Model 4			0.03
Transit accessibility [1-5]	0.50	0.00	
Constant	-3.10	0.00	
Data used for all models			
Trips (n)	41,316		
Walk	3,946		
Not Walk	37,370		

The resulting weights for PIE_{bg} are shown in Table 3 E below and compared to the commensurate measures from PIE₀. Here, we can see that the person density has a similar weight of 4.6; however, the other variables were apparently more sensitive to the change in scale and resulted in higher weights in PIE_{bg} than in PIE₀.

Table 3 Estimated weights for each built environment measure for PIE₀ and PIE_{bg}

Component	Range of values	Weights for PIE ₀	Weights for PIE _{bg}
People per acre	1 to 5	4.6	4.6
ULI	1 to 5	3.1	4.8
Road network density	1 to 5	-	6.1
Block size	1 to 5	3.1	-
Transit access	1 to 5	3.5	4.7
Sidewalk completeness	1 to 5	2.8	-
Bicycle access	1 to 5	2.8	-

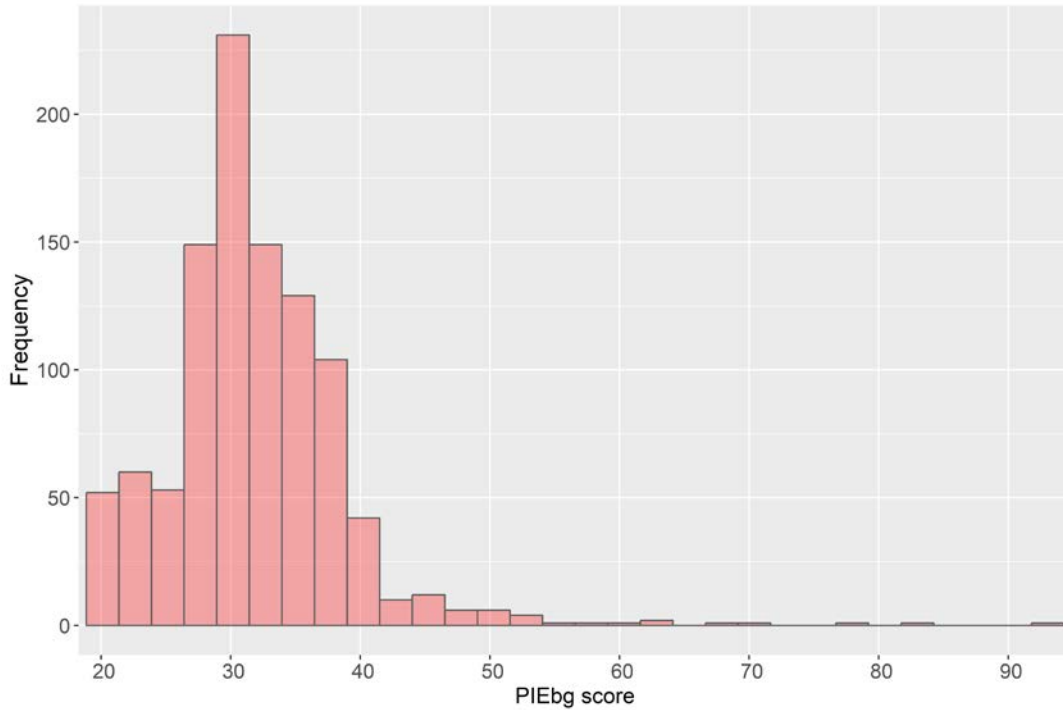


Figure 4 shows the histogram of the resulting values of PIE_{bg} for each block group computed using these data and weights across the Portland region. The most frequent scores occur in a range between 26 and 38. We see a limited number of locations that have high scores of PIE_{bg} , which is characteristic of many North American cities.

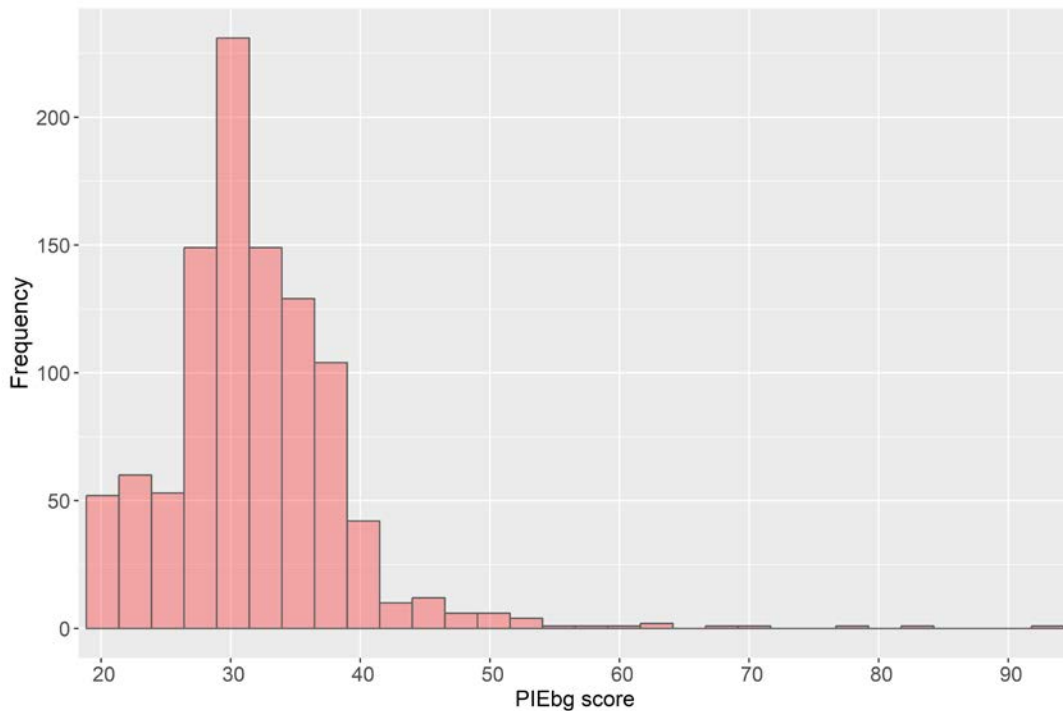


Figure 4 Frequency distribution of PIE_{bg} scores by block group

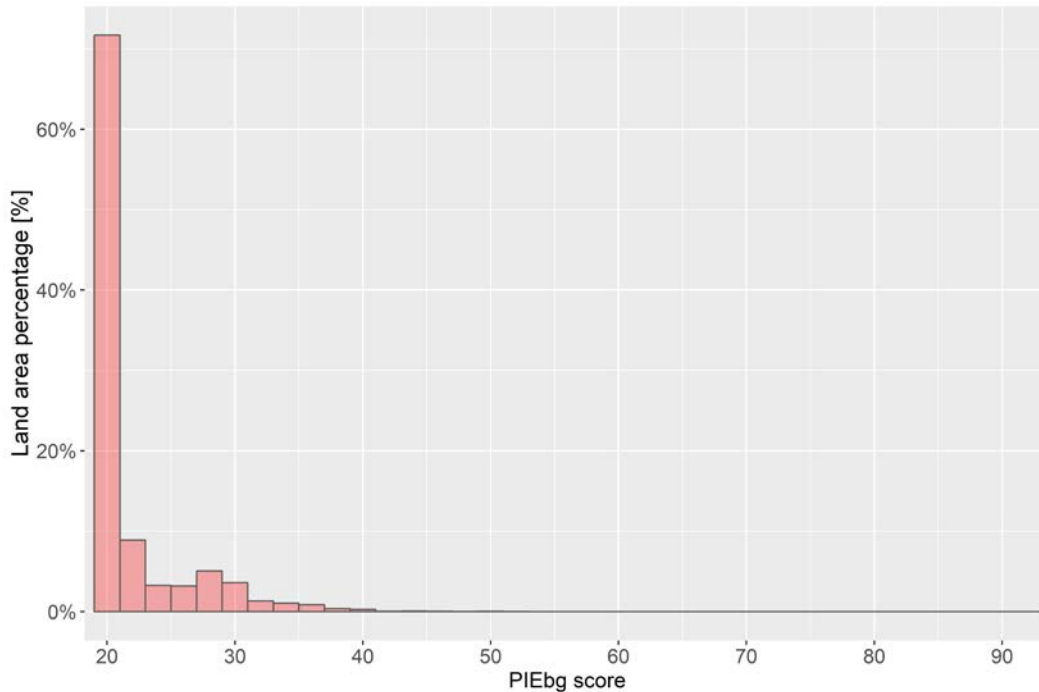


Figure 5 Land percentage by PIE_{bg} score

Figure 6 shows this distribution spatially. The highest scores of PIE_{bg} (shown in red) are all concentrated in the Central City of Portland. In contrast, the vast majority of the region has scores that are under 30 (dark green). When analyzing the large patterns of the metro area, around the 90% of the land area has slightly less than half of the total population, a large share of this area is mainly rural/exurban. Furthermore, the areas with the highest walkability have scores that ranges between 40 and 100 represent less than the 1% of the area and only 5% of the population (~6.5% of the households sampled) and generates almost 14% of the total trips of the region. Visually from Figure 1 and Figure 6, we can see that despite the differences in the data sources, scale, and weighting scheme, the two walkability measures track similarly across the region. We will investigate these differences statistically and in models of walking below.

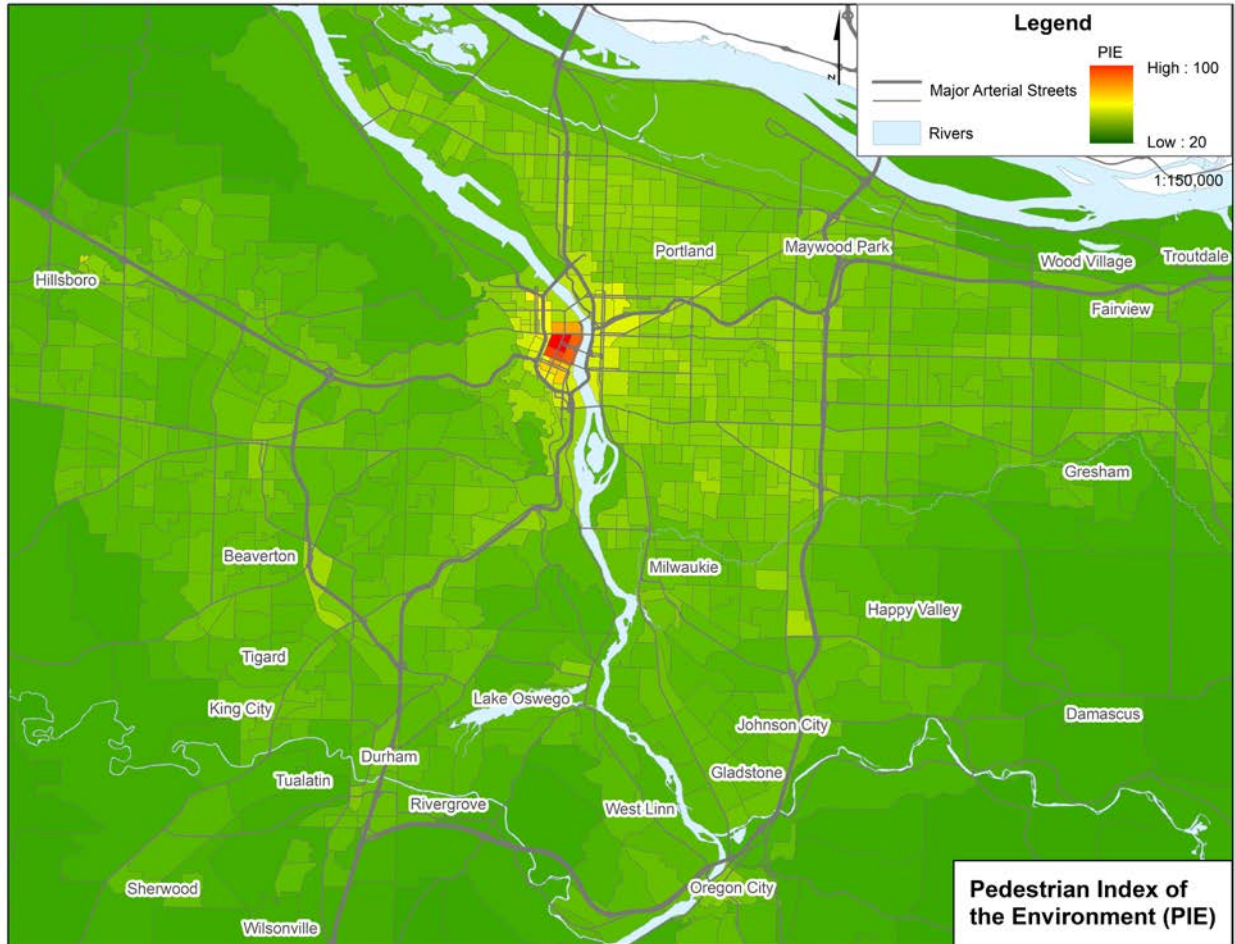


Figure 6 PIE_{bg} distribution in Portland

The two versions of PIE, PIE_0 , and PIE_{bg} are compared for each block group in the study area. PIE_0 is calculated at the pedestrian analysis zone, was averaged over the block group and plotted against the value of PIE_{bg} . Both indices were calculated using dependent variables between 1 and 5 and were weighted to have values between 20 and 100. Figure 7 shows the comparison of the two constructs in Portland, OR. As mentioned earlier, PIE_{bg} scores are generally lower in value than PIE_0 scores. This is due to differences in our approaches to scaling the originally built environment data (discrete values with natural breaks for PIE_0 ; continuous rescaled data for PIE_{bg}). However, both measures are coherent as the differences only reflect a change in the scale of the score, but both rise together.

An interesting observation of this graph the breakpoint in the PIE_{bg} score of 40 or PIE_0 of 70. This is what Figure 6 would be the central city of Portland plus some outer neighborhoods or cities commercial areas. The notion is that in more walking areas the difference between both PIE become smaller, almost reaching the same value in the highest values. This would mean, that in higher values the continuous rescaling of PIE_{bg} is more similar to the natural breaks approach of PIE_0 . Consequently, this could be interpreted as that the higher values of the characterization of the built environment are more clustered, meaning that the measures used by both versions of PIE become similar.

In sum, PIE_{bg} is coincident with an aggregated PIE_0 with some caveats. PIE_{bg} is less sensitive to changes in the built environment for suburban contexts and more sensitive in more urban areas than PIE_0 . These differences are likely due to the reliance on natural breaks in PIE_0 and continuous breaks in PIE_{bg} than the use of different data sources or the change in scale. The next section will explore how PIE_{bg} as a measure of walkability is associated with walk mode share at the block group level across the region.

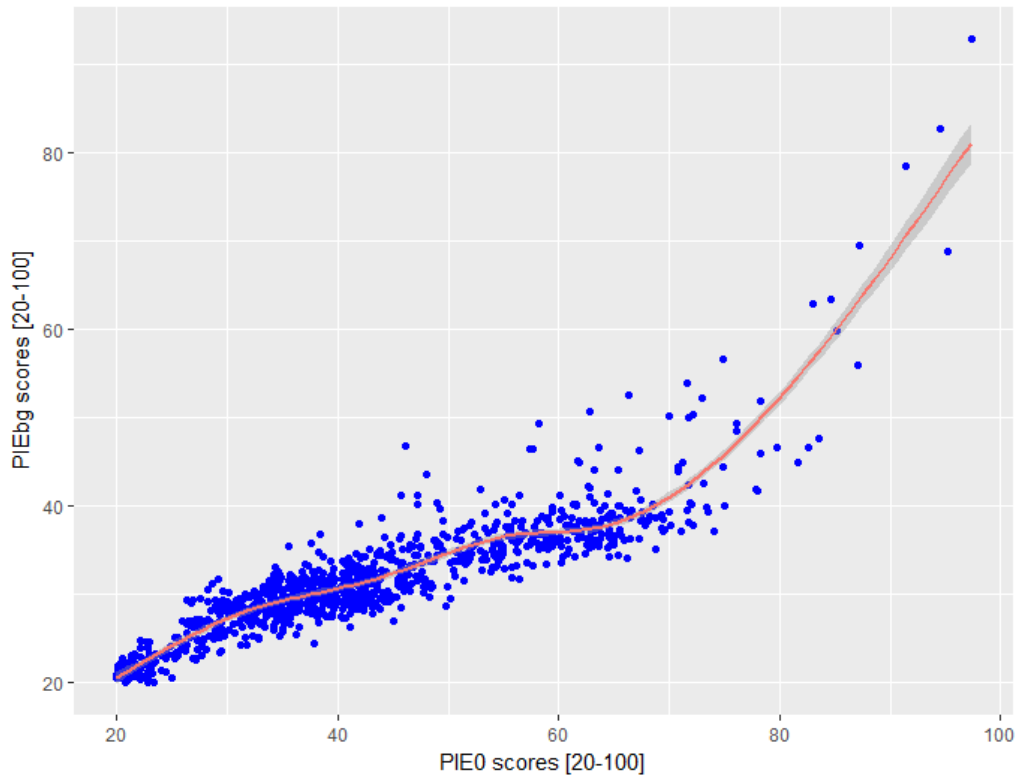


Figure 7 PIE_0 versus PIE_{bg}

5.2 Intraregional measure transferability

We now turn our attention to examining the question of how walk mode share varies within a region and the relationship with our measure of walkability, PIE_{bg} . Here we question if the relationship between walking and the built environment changes over the spectrum of urban spaces and if one walkability measure (or one model of walking behavior estimated from a pooled sample from the region) is adequate. Or contrarily, do we need to segment models (or walkability measures) by the urban regime in order to better reflect this varying relationship over the region.

To do this, we used the weighted expansion of Oregon Household Travel Survey, which inflates the sample of trips to represent the total population of trips, to compute the total amount of trips and walk trips in each block group (walk mode share). Figure 8 shows the resulting plot of the walk modal share by deciles of PIE_{bg} . The graph shows two trends (or three if we include the first rise from decile 1 to 2). For the bottom, 60% of PIE_{bg} areas have a walking share that varies

between 6-8%. The second tendency in the graph indicates an increase of walking share with the PIE_{bg} scores at the seventh decile or greater. These two main differences in the slopes suggest that they might be two different behavioral responses to the built environment.

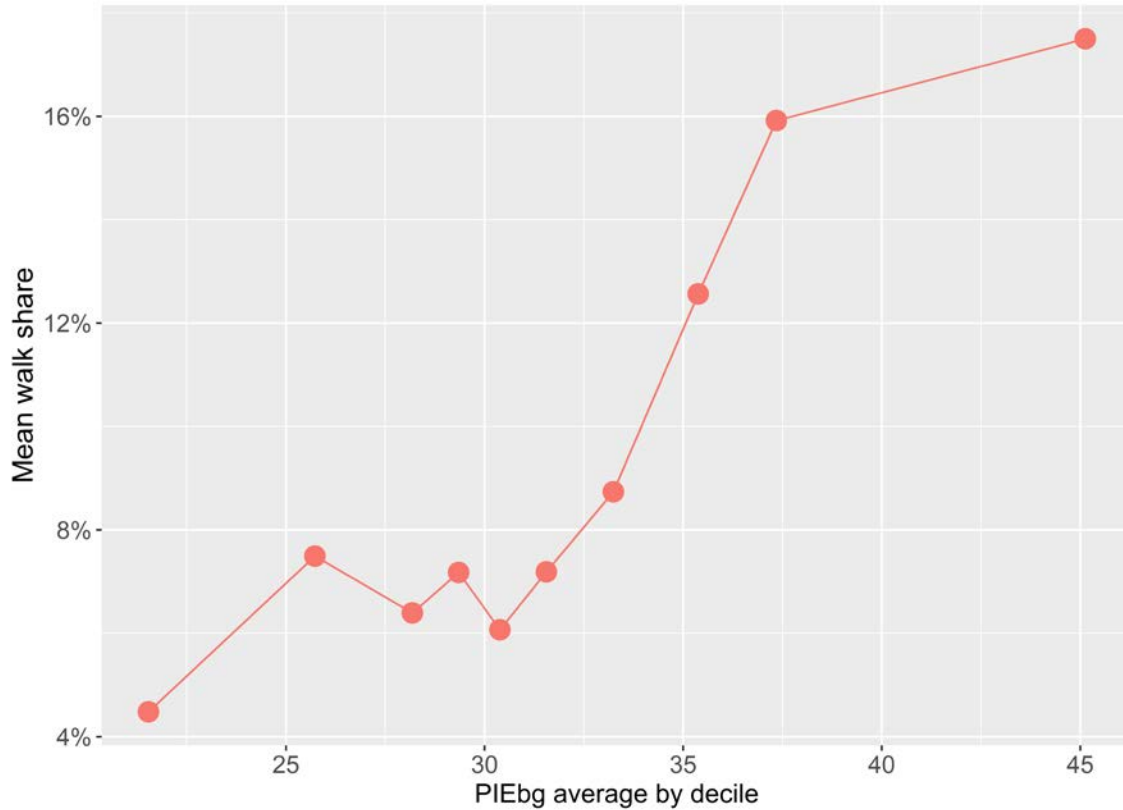


Figure 8 PIE_{bg} aggregated by deciles vs walking share

These two regimes could be identified as the separation between a suburban environment with low people density and an urban environment with a higher people density. This suggests that the relationships between walking and the built environment should not be represented by just one formal functional and that this function may not be linear. More detailed information about the walking behavior and more data points across the spectrum of PIE_{bg} are needed before we can discern and quantify relationship. In the next section, we are going to test the different regime hypothesis by contrasting two separated models versus a single one.

These results suggest that it may not be appropriate to model the relationship between walk mode share and the built environment as a simple linear relationship or estimating a single model parameter for this relationship across all locations. Our findings suggest that there may be a threshold that separates two different regimes of walking behavior relative to the built environment. To examine this more closely, we test the results of a) a singular model with one parameter estimate for the built environment across all urban environments, with b) a model estimated at two intervals, with one parameter for areas that are more urban, another representing suburban environments. The model is a univariate binary logit model (the same of section 4.0)

that predicts the probability of walking as a function of a single independent variable, an index of the built environment, which is represented as PIE_{bg} .

The model estimation results are shown in Table 4. Model 1 and 2 are separate models for urban and suburban areas, respectively, where urban areas are defined as those areas with a PIE_{bg} score of 40 or greater and suburban areas are those with a score below 40. Model 3 is a pooled model across all values of PIE_{bg} . We used a PIE_{bg} score of 40 because this was the approximate breakpoint in the relation between PIE_0 and PIE_{bg} in Figure 7 and the natural break in the distribution of PIE shown in Figure 1. However, one potential topic for future research is determining the exact breakpoint(s) of PIE in the defining these different regimes. For the purposes of this research, we will use 40 as the breakpoint and test the relationship between two regimes only due to sample size limitations and the maximum values of PIE in Portland, we could not test more with statistical accuracy.

The three models had very low explanatory power, which is expected from behavioral models in general and from a univariate model such as this. However, we are only estimating these models for comparison. From simple inspection of the results, it is noticeable that the coefficients are different between the three models. We tested the validity of separating the models using the likelihood ratios test. The test rejects the pooled model as valid ($p < 0.000$), meaning that the correct specification should be two separated models. Therefore, this evidence supports our hypothesis that the built environment does not have the same effect on the choice to walk in a city. However, this difference may be due to a spatial sorting of people according to the differences in the attitudes and inherent preferences of the people in those respective areas. However, we are lacking information about those aspects and cannot test them in this study.

A visualization of these model results is shown in Figure 9. This figure shows the predicted walking mode share plotted against the PIE_{bg} score. The graph shows a different prediction for the probability of walking for the different set of models. From the two separated models (urban and suburban), we can see that the effect of the built environment is larger in the suburban context and less steep in the urban environment, but still considerable. It is important to notice that almost all Portland has a low PIE score, as only 0.6% of the block groups have over 60 in the score, and 0.04% of the total land area (the very central city). Remember the area in the analysis is the Metropolitan Statistical Area that is based in counties and includes plenty of rural lands. Therefore, the larger portion of Portland as shown in Figure 4 is around a PIE_{bg} score of 30 that is consistent with the 8% overall walking share in the region. The pooled model in the vicinity of the score 30 shows a similar walking probability, however, there is an overestimation of probability trips in the rural/suburban area and underestimation in the transition between urban and suburban, and in the more urban areas.

Despite the limitations, the exercise done here lends empirical evidence to support that concern that the relationship between walking and the built environment may be different (non-linear or linear in different regimes with different coefficients) across different environments within the same region. Thus, one model estimate may not be sufficient to capture this relationship and thus not transferable to all areas within a region.

Table 4 Pooled vs. Urban/Suburban Walk Mode Share Models

Context variable (x_n)	Coefficient (β)	p-value	Model pseudo-R ²
Model 1 Urban			
PIEbg (40-100)	0.02	0.00	0.03
Constant	-2.83	0.00	
Log-Likelihood	-2993		
Chi-squared	166		
Model 2 Suburban			
PIEbg (20-40)	0.10	0.00	0.03
Constant	-5.73	0.00	
Log-Likelihood	-9364		
Chi-squared	521		
Model 3 All urban types			
PIEbg (20-100)	0.036	0.00	0.02
Constant	-3.56	0.00	
Log-Likelihood	-12489		
Chi-squared	1060		
Log-Likelihood ratio test segmentation of the model		0.00	
Data used for all models			
Trips (n) – all urban types	41,316		
Walk	3,946		
Not Walk	37,370		
Trips (n) – urban	6,350		
Walk	1,199		
Not Walk	5,151		
Trips (n) – suburban	34,966		
Walk	2,747		
Not Walk	32,219		

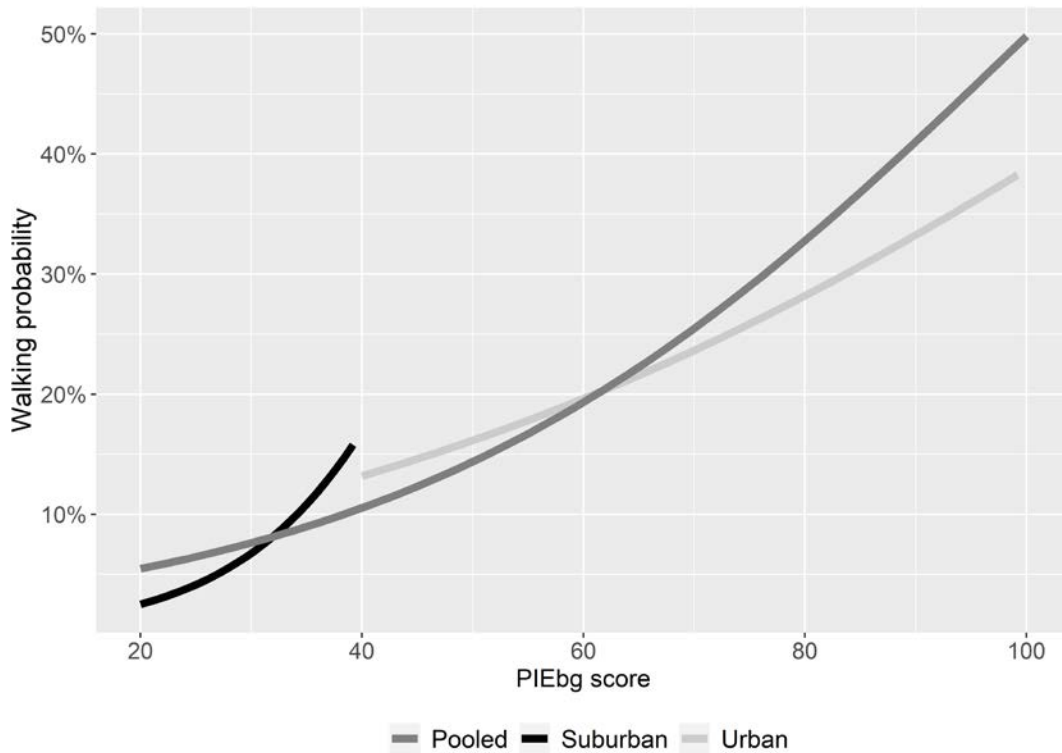


Figure 9 Predicted walk probabilities

5.3 Interregional measure transferability

In this section, we test the differences in various estimations of PIE_{bg} for different cities. As the objective was to assess the transferability of our walkability measure, we needed to associate it with travel behavior data from each city in our study. To test relationships within and across regions, we assembled four US travel surveys: those covering the state of California (Kunzman and Daigler 2013) as well as the metropolitan areas of Minneapolis–St. Paul (Cambridge Systematics 2014), Portland (Oregon Modeling Steering Committee 2012), and Seattle (Cambridge Systematics 2006). The California travel survey allowed for the inclusion of the Metropolitan Statistical Areas of Los Angeles, San Diego, and San Francisco-Oakland (without San Jose), bringing the total of US regions to six.

Following the general procedures outlined in Section 4.0, we estimate PIE_{bg} for five additional cities (Los Angeles, Minneapolis, San Diego, San Francisco, and Seattle) using household travel surveys and built environment data from each region. Then we compare the coefficient estimates across these regions to test if they are comparable and thus, indicate that PIE estimated from one region could indeed be applicable in another. Table 5 shows details from each metropolitan area travel survey and built environment attributes statistics.

Table 5 Trip information and built environment measures in each metropolitan area

	LOS ANGELES	MINNEAPOLIS – ST. PAUL	PORTLAND	SAN DIEGO	SAN FRANCISCO	SEATTLE
Year of survey	2010-12	2010	2011	2010-12	2010-12	2006
Total trips	32.8M	11.1M	6.5M	7.8M	13.4M	12.6M
Population	12.8M	3.3M	2.2M	3.1M	4.3M	3.4M
Walking mode share (%)	11.0%	6.2%	8.6%	8.5%	15.9%	7.9%
No. of block groups	8248	2314	1555	1795	2903	2483
Population density range [ppl/acre]	[0,300]	[0,180]	[0,100]	[0,93]	[0,319]	[0,173]
Population density mean [ppl/acre] (std. deviation)	20 (16)	8 (9)	8 (7)	14 (11)	21 (24)	9 (10)
Employment density range [ppl/acre]	[0,579]	[0,354]	[0,251]	[0,330]	[0,900]	[0,1120]
Employment density mean [ppl/acre] (std. deviation)	14 (5)	16 (4)	14 (4)	12 (4)	36 (8)	29 (5)
ULI range [ppl/acre]	[0,66]	[0,127]	[0,56]	[0,73]	[0,795]	[0,115]
ULI mean [ppl/acre] (std. deviation)	1 (3)	1 (4)	1 (3)	1 (3)	3 (24)	1 (4)
Road network density range [miles/sq. miles]	[0,69]	[0,55]	[0,54]	[0,62]	[0,65]	[0,86]
Road network density mean [miles/sq. miles] (std. deviation)	21 (7)	17 (8)	18 (10)	19 (8)	22 (9)	17 (9)
Transit accessibility range [aggregated frequency]	[0,4401]	[0,5038]	[0,2351]	[0,526]	[0,2028]	[0,1839]
Transit accessibility mean [aggregated frequency] (std. deviation)	71 (140)	128 (230)	151 (211)	35 (45)	86 (144)	49 (97)

Some basic analysis shows that the city that has the most walking is San Francisco (15.9%). The larger city in population is Los Angeles with 12.8 million people. The population density ranges are higher in Los Angeles and San Francisco while the employment density is higher in Seattle and San Francisco.

Following the assembly and standardization of the datasets, we made a few exceptions to the methodology. We did not scale each of the built environment data to have values between 1 and 5. Instead, we scaled the data to have an average of 0 and a standard deviation of 1. This allowed us to have standardized coefficients that let us directly compare the magnitude of the coefficients between cities. In doing this, the estimated standardized coefficients are only meant to be used for comparison across regions and not for application. Finally, we did not sum the employment and population density variables as we wanted to test the independent effect of each measure, however, we re-combine them as people density later in this section in an application to Los Angeles.

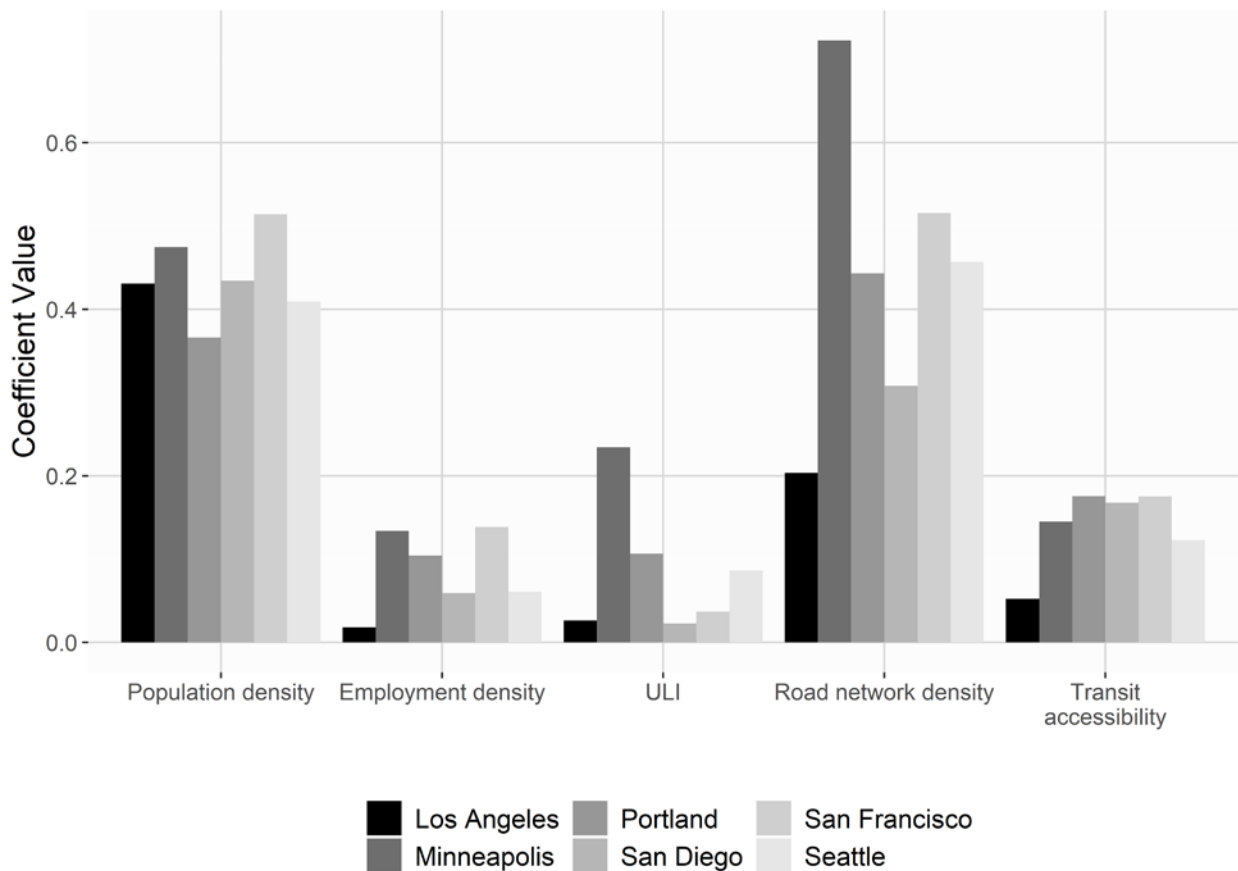


Figure 10 Standardized coefficients for different built environment measures

The results of these estimations for the different cities are shown in Figure 10. All the estimates were statistically significant ($p < 0.001$). Of the various built environment measures tested, the standardized coefficients for population density appear to have the most consistent values across the cities in the study, with values ranging from 0.37 in Portland to 0.51 in San Francisco. For the Los Angeles and San Diego regions, population density was the most important built environment characteristic in explaining walk mode choice. Given its stability and its relatively

high standardized coefficient, population density may be the most important built environment measure associated with the odds of walking across all of the cities, which is consistent with the literature.

In contrast, road network density has much more variability based upon the statistics in Table 5 that is reflected in the values of the standardized coefficients in Figure 10, which range from a low of 0.20 in Los Angeles to a high of 0.72 in Minneapolis. This was the most important measure in explaining walking for several cities: Minneapolis, Portland, San Francisco, and Seattle. This range of values is perplexing but may be a function of the computation of the attribute itself. Road network density includes all facility types, including freeways, and for that reason, it may have confounding results. For future work, we suggest using intersection density as the measure of pedestrian connectivity.

The standardized coefficients for employment density and retail and entertainment employment density are also variable across the cities. The values for employment density range from a low of 0.02 in Los Angeles to a high of 0.14 in San Francisco. Overall, the coefficient values are low compared to the other built environment attributes, and probably the spatial distribution of employment, something that we have not covered here, might explain its influence on walking behavior. It may also be that employment density is an imperfect proxy for access to destinations and a local accessibility measure or a mixed use measure may be a more stable measure across locations.

Narrowing the scope to retail, service, and entertainment establishments, as measured by ULI did not result in a decrease in the variation across places. The standardized coefficients varied from a low of 0.02 for San Diego to a high of 0.23 for Minneapolis. The suggestion to improve the measure of access to destinations holds.

These results raise a lot of questions about the cause of these results, including how these variables are distributed across the different regions, where are the locations where walking trips are observed in the sample, and perhaps variations in the trip purposes. But the results suggest that a walkability measure estimated in one location may only be cautiously applied in other regions, if at all.

To explore this issue further, we want to test the predictive ability of PIE_{bg} estimated in one region and applied in a model of walk mode share in another. In order to make the broadest comparison, we chose to utilize PIE_{bg} estimated from Portland in a univariate model of walk mode share in Los Angeles - the largest city in our study. We make additional modifications to the process of estimating PIE_{bg} using Portland data. This time, the data for each built environment measure (x_n) used in the calibration are not scaled at all because our aim is not to compare the coefficients but rather apply PIE in a model of walking. Because Portland and Los Angeles have different ranges and distributions of these variables, as shown in Table 5, we needed to ensure that the data were unscaled in the application across different locations. Second, population and employment density are re-combined into people per acre. The results are shown in Table 6.

Table 6 Unscaled PIE_{bg} coefficients estimated from Portland data

Built Environment variable (x_n) ¹	Coefficient (β)	p-value	Model pseudo- R^2
Model 1			0.03
People per acre	0.007	0.00	
Constant	-2.486	0.00	
Model 2			0.02
ULI	0.039	0.00	
Constant	-2.402	0.00	
Model 3			0.05
Road network density	0.052	0.00	
Constant	-3.463	0.00	
Model 4			0.04
Transit access	0.001	0.00	
Constant	-2.598	0.00	
Data used for all models			
Trips (n)	41,316		
Walk trips	3,946		
Not walk trips	37,370		

Using the coefficients for each of these built environments as weights, we computed PIE_{bg} for each block group in the Los Angeles region by using the following formula:

$$PIE_{bg} = 0.007 * \text{people per acre} + 0.039 * ULI + 0.052 * \text{road network density} + 0.001 * \text{transit accessibility}$$

Figure 11 shows both the distribution in the city of unscaled PIE_{bg} . Both areas have a clear central area that it is more walkable. In both cases, this area is small compared to the region. Even though the urban structure of Los Angeles seems very different from that of Portland, it seems from a visual inspection of the map that the PIE_{bg} specification of Portland applied to Los Angeles make some intuitive sense. However, as seen in Table 5 Trip information and built environment measures in each metropolitan area Table 5 many of the built environment attributes are much larger in Los Angeles than Portland. For example, the mean population density in each block group in Los Angeles is 20 people per acre, while in Portland is only 8. This means that the model estimation in Portland can omit and effect of what larger values of the built environment attributes may cause, causing that the PIE_{bg} scores in Los Angeles only reach higher values in very few places. Additionally, as shown also in Table 5 is that Los Angeles have a lower average transit service than Portland and also as Figure 10 shows the coefficient for transit is smallest in Los Angeles than in the other cities, for that reason, the effect of transit might also be creating some noise in the scores.

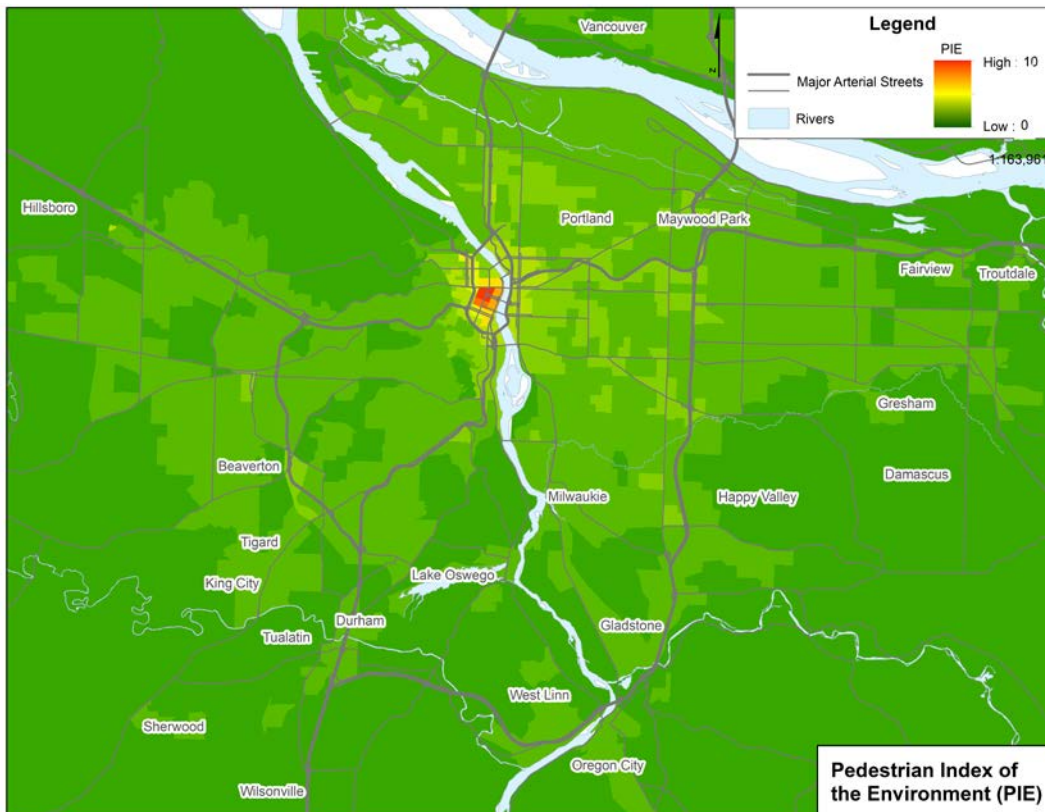
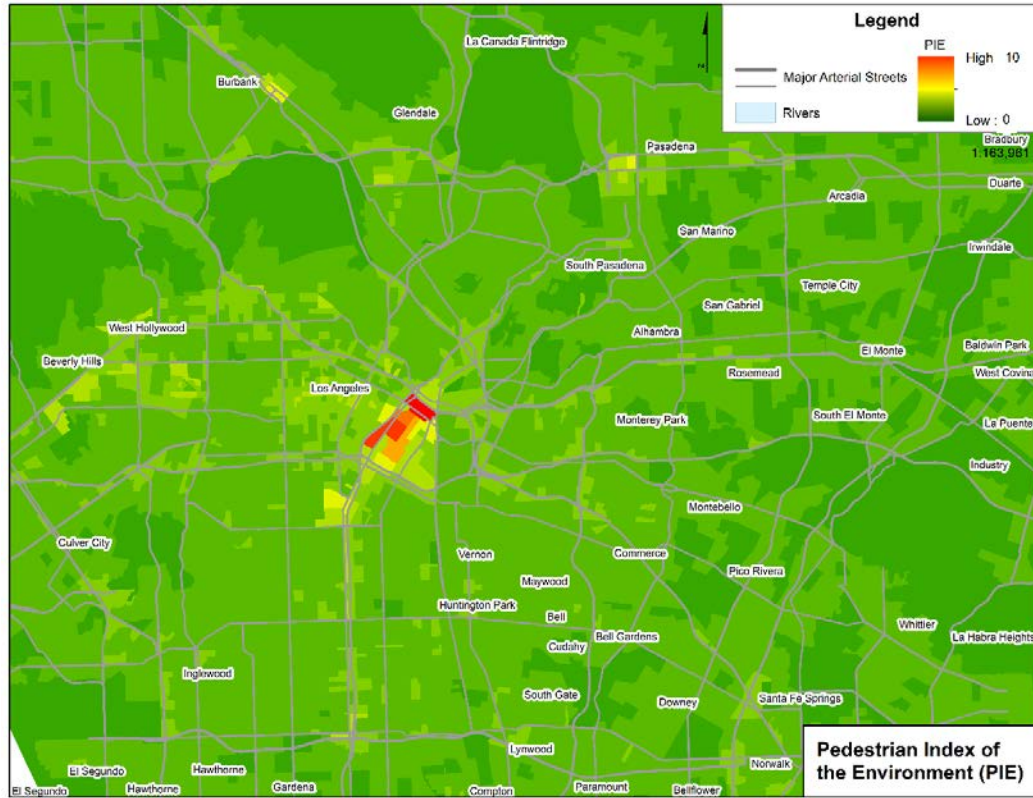


Figure 11 Unscaled PIE_{bg} distribution in Los Angeles and Portland

6.0 DISCUSSION AND CONCLUSIONS

The analysis of aggregate and disaggregate pedestrian trips presented here examined the potential transferability of relationships between walking and the built environment within and across different regions. To summarize some of these findings:

- The probability of walking has a clear relationship with PIE_{bg}
- An examination of PIE_{bg} and walk mode share suggests that there may be different relationships between walking behaviors and the built environment across the spectrum of environments within a region. Our intraregional tests for transferability of models and measures revealed two regimes where these relationships may differ: urban versus suburban contexts. However, more research is needed to better understand these variations within a region and the conditions defining these regimes.
- Population density has the strongest and most consistent relationship with walk mode share across the six US metropolitan regions tested in our study.
- Interregional tests of transferability of PIE_{bg} revealed similar walk mode share results in Los Angeles and Portland. This provides initial support that the PIE_{bg} construct may be transferrable between metropolitan regions.

More detailed discussion of these conclusions is below.

6.1 Transferability of models and measures

The findings were mixed in terms of the degree to which PIE could be applicable to areas beyond its estimation region. In our examination of intraregional transferability, there appeared to be different relationships between walking and the built environment in different parts of the region, divided simply into lower density (suburban) and higher density (urban) areas. This suggests that perhaps different constructs may be needed to represent different walking regimes within a region. Yet, the characteristics (thresholds) defining each walking regime are not the same in Portland and Los Angeles. Further, our findings also suggest that the relationship between PIE_{bg} and walk mode shares may not be linear. This notion suggests the existence of different regimes of response from people to the built environment that should be better represented by our predictions tools.

One reason that defining these regimes and the functional form of the relationship between walking and the built environment was difficult is because of the relatively smaller sample sizes in the urban areas resulted in noise in the estimates. Thus, larger sample sizes from higher density areas are needed to better understand these relationships. Further, there are potential correlations with the socio-economic characteristics of travelers and the characteristics of the locations of travel, which point to the need to control for this in walkability measures (discussed in the next subsection).

Based upon our analysis, travel behavior surveys from the US cities we studied tend to have greater proportion of sample from lower density areas associated with American suburban form and smaller sample sizes in areas with higher walkability and overall walking activity, which represent a smaller proportion of both land mass and population of those cities. This may not be the case in international contexts, where more of the urbanized regions are of a higher density than the US. Creating pooled samples with comparable and consistent data from cities around the world is becoming a more realistic endeavor with standardization of travel surveys and the availability of built environment data. This would be an asset to understanding the relationships between travel behavior and the built environment, broadly, and identifying the various regimes of walking, specifically.

From our analysis, the distribution of walkable places and their relative locations across the region appear to be important. For example, Los Angeles region has more variation in its urban structure and is more polycentric than Portland. This raises the question of the role of the larger urban spatial structure in supporting walking activity. By urban structure we mean the number and distribution of centers or sub-centers across a region, the density gradients, the clustering or contiguity of walkable places, and the spatial extent of the urbanized area (total land area). These are characteristics that are not captured well in local and regional accessibility measures and perhaps represent a meso-scale description of the built environment.

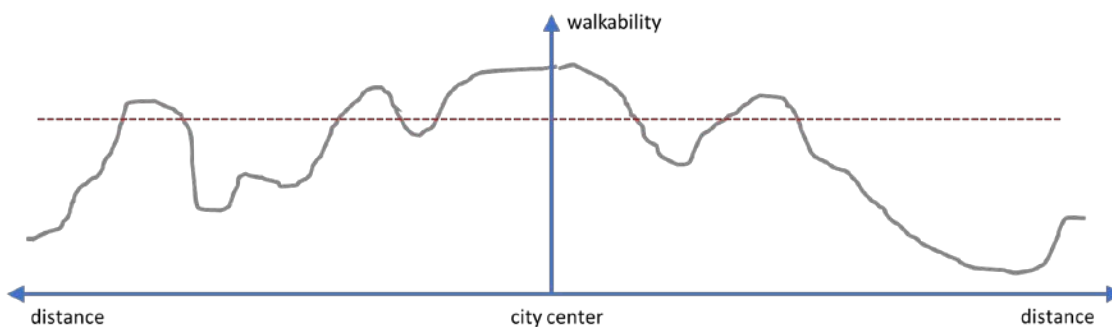


Figure 12 The role of walkability gradients

Figure 12 shows an example of walkability across a fictitious cross section of an urban region, from the city center outward. While all areas across the red dashed line have the same objective measure of walkability, there is significant variation in the walkability gradients as well the total space that has that level of walkability or more. One question to investigate is would those locations proximate to the city center, where walkability is relatively high and sustained across the space, have the same walk mode share or absolute levels of walking as those farther from the center with the same objective measure of walkability but different gradients and total area of walkable space. Or put another way, would we expect pockets of walkability to perform similarly as places where walkability is sustained over a larger area? At the moment, our methods of analysis do not consider this and this may be an area of fruitful future work that may help to explain the differences we see between regions.

6.2 Representing the built environment

The ways that we have been constructing PIE have limited its transferability and have problematic issues that limit its usefulness. This new construct, PIE_{bg} has some important differences to our original PIE_0 , in part because of our need to have comparable measure across cities. When PIE_0 was developed in 2011, the data available in Portland were unique and at a scale not widely available (PAZs - 80m x 80m grid cells). This presented the opportunity to examine walking at a scale more consistent with the behavior of interest and with zones of uniform size. However, these data were originally continuous and were then reduced to categorical on a scale of 1 to 5 based upon natural breaks in their distribution over the Portland region in the base year. This resulted in the loss of variation as well as reducing its applicability to other regions or even future Portland conditions.

In this research, we attempted to address this limitation in the creation of PIE_{bg} . However, this new built environment construct also has limitations. In using the SLD data at the block group level, we then had comparable data across regions but the fine-grained and consistently-sized spatial resolution of the PAZ grid cells was lost. Further, we realize that in estimating the various coefficients (or weights) for PIE with built environment variables while not controlling for socio-economic characteristics may bias those estimates. Also, there may be econometric issues with our constructs in that we are using the coefficients from a walk mode choice model as weights in our final construction of PIE and then in turn, using that very construct in other models of mode choice and trip distribution in MoPeD. These limitations are problematic.

Now at the end of this research process, we have some important recommendations about future attempts to create a walkability measure using various built environment characteristics. The literature has revealed a myriad of associations between the built environment and walking behavior. Many of these characteristics are highly correlated with one another and then, are problematic when including them as separate independent variables in a model. Thus, this is one of the rationales for creating an aggregate index or measure that reflects the walking environment. Given the variations in the availability of data, the difficulties in interpretation, and the correlations between these phenomena, we recommend parsimony in the selection of built environment characteristics to represent in such a measure.

We recommend using continuous data to represent the built environment rather than reducing the information to a categorical measure. The reasons for this are maintaining the variation in the measure as well as the difficulty in representing environments that exceed the conditions present in the estimation year.

Of those measures tested in our study, population density consistently explained more of the variation in walk mode share, followed by measures representing pedestrian connectivity. These two built environment characteristics are important to include and not as highly correlated with one another as we originally believed. Several measures of connectivity were tested over the course of this project: average block size, road network density, and average streets per node. There are various concerns and limitations with each. Employment density was highly variable, including the ULI measure, that attempts to capture access to local destinations.

Contemporaneous with this study, new guidebook on measuring non-motorized connectivity was completed (FHWA, 2018) with the input and participation of this study's author, and offers suggestions that may resolve these issues. Based upon this guidebook, we recommend including a pedestrian connectivity measure that captures access to destinations into future walkability indices. This compensates for the low performance of employment density and incorporates network connectivity in one measure.

Transit accessibility and service is a difficult measure to incorporate consistently. From a theoretical perspective, the links between transit and access and egress mode of travel is logical. However, in this work, we are only considering trips where walking is the single mode of transport. Thus, the rationale for including transit accessibility is less clear. One could argue that in areas with good transit, more trips would be made on foot by persons traveling to the area via transit. However, this may be a proxy for (or confounded with) the built environment conditions in these locations. In development of a walkability measure, it may be best to omit transit accessibility from the direct representation of walkability and instead include as an additional measure for better interpretation of results.

When data permit, using uniform grid cells as the spatial unit of analysis is preferable to block groups, which suffer from a size differential that is correlated with population density. Others have examined the effects of the size of these grid cells on model performance and have found tradeoffs between size, model run times, and accuracy (Zhang et al. 2018). However, the loss of accuracy from doubling the grid size from 80m grids to 160m grids is nominal compared to the advantages for model runs and data availability.

In terms of the econometric issues with PIE, future representations of walkability using PIE index will be computed as the sum of the portion of utility function of a walking mode choice model that deals with the built environment. By minimizing the number of built environment variables included in this estimation, correlation between built environment variables does not exceed the threshold for concern. Thus, they can be included directly without the use of an index. Here socio-economic and other important characteristics will be controlled for to include inflation of the coefficients for the built environment variables. This development and presentation of these results are outside the scope of this report but are based upon lessons learned during the process of this research.

6.3 Future work

Several limitations inhibit stronger conclusions from this analysis. First, there are a limited number of cities included in this exploratory analysis. A more comprehensive study needs a larger number of cities with different activity density distributions (particularly on the high end) and varying urban spatial structure, including those from other countries. This expansion of the work could help to define regimes, and how these measures interact at different thresholds. Additional analysis may also include more advanced spatial analysis techniques.

In addition to considering the recommendations for the construct of PIE and other walkability measures, future work may want to include more coarse measures of the built environment to better capture urban structure, such as amount and distribution of density, total land area, density

gradients and continuity, and mixed-use, as they may play a role in understanding traveler response to the built environment, including pedestrian behaviors.

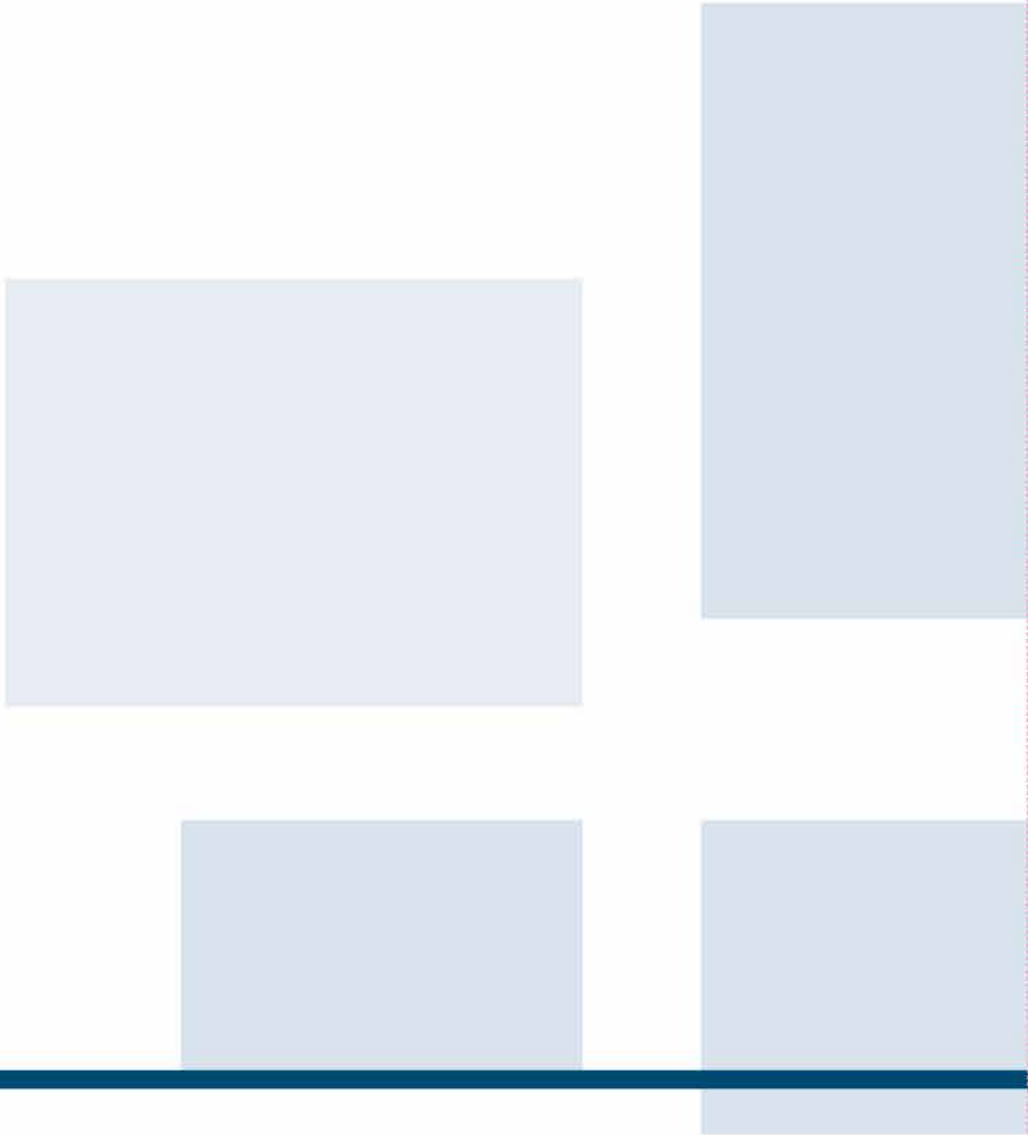
The consideration of non-linearities in the built environment-walking relationship is one contribution of this work that deserves more exploration. Although more analysis is needed to understand the specific functional form that these non-linearities take, our analysis provides evidence that continuing to assume linear relationship across the urban spectrum is questionable.

7.0 REFERENCES

- Badoe, Daniel A., and Eric J. Miller. 1995. "Analysis of Temporal Transferability of Disaggregate Work Trip Mode Choice Models." *Transportation Research Record*, no. 1493. <https://trid.trb.org/view/452588>.
- Ben-Akiva, Moshe. 1981. "Issues in Transferring and Updating Travel Behavior Models." In *New Horizons in Travel-Behaviour Research*, edited by Peter R. Stopher, A.H. Meyburg, and W. Brög, 665–86. Lexington: Lexington Books. <https://trid.trb.org/view/1153691>.
- Ben-Akiva, Moshe, and Denis Bolduc. 1987. "Approaches to Model Transferability and Updating: The Combined Transfer Estimator." *Transportation Research Record*, no. 1139. <https://trid.trb.org/view/289161>.
- Berrigan, David, Linda W Pickle, and Jennifer Dill. 2010. "Associations between Street Connectivity and Active Transportation." *International Journal of Health Geographics* 9 (1): 20. <https://doi.org/10.1186/1476-072X-9-20>.
- Cervero, Robert, and Kara Kockelman. 1997. "Travel Demand and the 3Ds: Density, Diversity, and Design." *Transportation Research Part D: Transport and Environment* 2 (3): 199–219. [https://doi.org/10.1016/S1361-9209\(97\)00009-6](https://doi.org/10.1016/S1361-9209(97)00009-6).
- Clifton, Kelly J. 2017. "Getting From Here to There: Comment on 'Does Compact Development Make People Drive Less?'" *Journal of the American Planning Association* 83 (2): 148–51. <https://doi.org/10.1080/01944363.2017.1290494>.
- Clifton, Kelly J., and Christopher Muhs. 2012. "Capturing and Representing Multimodal Trips in Travel Surveys." *Transportation Research Record: Journal of the Transportation Research Board* 2285 (October): 74–83. <https://doi.org/10.3141/2285-09>.
- Clifton, Kelly, Patrick Singleton, Christopher Muhs, and Robert Schneider. 2015. "Development of a Pedestrian Demand Estimation Tool." *Civil and Environmental Engineering Faculty Publications and Presentations*, September. <https://doi.org/10.15760/trec.124>.
- Clifton, Kelly, Patrick Singleton, Christopher Muhs, Robert Schneider, and Peter Lagerwey. 2013. "Improving the Representation of the Pedestrian Environment in Travel Demand Models, Phase I." *Civil and Environmental Engineering Faculty Publications and Presentations*, September. <https://doi.org/10.15760/trec.120>.
- Ewing, Reid, and Robert Cervero. 2010. "Travel and the Built Environment." *Journal of the American Planning Association* 76 (3): 265–94. <https://doi.org/10.1080/01944361003766766>.
- . 2017. "'Does Compact Development Make People Drive Less?' The Answer Is Yes." *Journal of the American Planning Association* 83 (1): 19–25. <https://doi.org/10.1080/01944363.2016.1245112>.

- Fox, James, Andrew Daly, Stephane Hess, and Eric Miller. 2014. "Temporal Transferability of Models of Mode-Destination Choice for the Greater Toronto and Hamilton Area." *Journal of Transport and Land Use* 7 (2): 41. <https://doi.org/10.5198/jtlu.v7i2.701>.
- Frank, Lawrence D. 2000. "Land Use and Transportation Interaction: Implications on Public Health and Quality of Life." *Journal of Planning Education and Research* 20 (1): 6–22. <https://doi.org/10.1177/073945600128992564>.
- Federal Highway Administration (FHWA). 2018. "FHWA Guidebook for Estimating Multimodal Network Connectivity." US Department of Transportation, Washington, DC. Available online: https://www.fhwa.dot.gov/environment/bicycle_pedestrian/publications/multimodal_connectivity/fhwahep18032.pdf
- Greenwald, Michael, and Marlon Boarnet. 2001. "Built Environment as Determinant of Walking Behavior: Analyzing Nonwork Pedestrian Travel in Portland, Oregon." *Transportation Research Record: Journal of the Transportation Research Board* 1780 (January): 33–41. <https://doi.org/10.3141/1780-05>.
- Handy, Susan. 2017. "Thoughts on the Meaning of Mark Stevens's Meta-Analysis." *Journal of the American Planning Association* 83 (1): 26–28. <https://doi.org/10.1080/01944363.2016.1246379>.
- Handy, Susan L., Marlon G. Boarnet, Reid Ewing, and Richard E. Killingsworth. 2002. "How the Built Environment Affects Physical Activity: Views from Urban Planning." *American Journal of Preventive Medicine* 23 (2): 64–73. [https://doi.org/10.1016/S0749-3797\(02\)00475-0](https://doi.org/10.1016/S0749-3797(02)00475-0).
- Huang, Jingnan, X. X. Lu, and Jefferey M. Sellers. 2007. "A Global Comparative Analysis of Urban Form: Applying Spatial Metrics and Remote Sensing." *Landscape and Urban Planning* 82 (4): 184–97. <https://doi.org/10.1016/j.landurbplan.2007.02.010>.
- Karasmaa, Nina. 2007. "Evaluation of Transfer Methods for Spatial Travel Demand Models." *Transportation Research Part A: Policy and Practice* 41 (5): 411–27. <https://doi.org/10.1016/j.tra.2006.09.009>.
- Kuzmyak, J. Richard, Jerry Walters, Mark Bradley, and Kara M. Kockelman. 2014. "Estimating Bicycling and Walking for Planning and Project Development: A Guidebook." *NCHRP Report*, no. 770. <https://trid.trb.org/view/1320568>.
- Larco, Nico, Jean Stockard, Bethany Steiner, and Amanda West. 2013. "Trips to Strips: Walking and Site Design in Suburban Multifamily Housing." *Journal of Urban Design* 18 (2): 281–303. <https://doi.org/10.1080/13574809.2013.772886>.
- Lerman, R. 1981. "A Comment on Interspatial, Intraspatial, and Temporal Transferability." In *New Horizons in Travel-Behaviour Research*, edited by Peter R. Stopher, A.H. Meyburg, and W. Brög, 628–32. Lexington: Lexington Books.
- Newman, Peter. 2014. "Density, the Sustainability Multiplier: Some Myths and Truths with Application to Perth, Australia." *Sustainability* 6 (9): 6467–87. <https://doi.org/10.3390/su6096467>.
- Nowrouzian, Roosbeh, and Sivaramakrishnan Srinivasan. 2012. "Empirical Analysis of Spatial Transferability of Tour-Generation Models." *Transportation Research Record: Journal of the Transportation Research Board* 2302 (December): 14–22. <https://doi.org/10.3141/2302-02>.

- Peiravian, Farideddin, Sybil Derrible, and Farukh Ijaz. 2014. "Development and Application of the Pedestrian Environment Index (PEI)." *Journal of Transport Geography* 39 (July): 73–84. <https://doi.org/10.1016/j.jtrangeo.2014.06.020>.
- Ramsey, Kevin, and Alexander Bell. 2014. "The Smart Location Database: A Nationwide Data Resource Characterizing the Built Environment and Destination Accessibility at the Neighborhood Scale." *Cityscape* 16 (2): 145.
- Rossi, Thomas F., and Chandra R. Bhat. 2014. "Guide for Travel Model Transfer," October. <https://trid.trb.org/view/1353320>.
- Saelens, Brian E., and Susan L. Handy. 2008. "Built Environment Correlates of Walking: A Review." *Medicine and Science in Sports and Exercise* 40 (7 Suppl): S550–66. <https://doi.org/10.1249/MSS.0b013e31817c67a4>.
- Sallis, James F., Robert B. Cervero, William Ascher, Karla A. Henderson, M. Katherine Kraft, and Jacqueline Kerr. 2006. "An Ecological Approach to Creating Active Living Communities." *Annual Review of Public Health* 27 (1): 297–322. <https://doi.org/10.1146/annurev.publhealth.27.021405.102100>.
- Schneider, Robert, Robert Patten, and Jennifer Toole. 2005. "Case Study Analysis of Pedestrian and Bicycle Data Collection in U.S. Communities." *Transportation Research Record: Journal of the Transportation Research Board* 1939 (January): 77–90. <https://doi.org/10.3141/1939-10>.
- Singleton, Patrick A., Robert J. Schneider, Christopher Muhs, and Kelly J. Clifton. 2014. "The Pedestrian Index of the Environment: Representing the Walking Environment in Planning Applications." In . <https://trid.trb.org/view/1289281>.
- Singleton, Patrick A., Joseph C. Totten, Jaime P. Orrego-Oñate, Robert J. Schneider, and Kelly J. Clifton. 2018. "Making Strides: State-of-the-Practice of Pedestrian Forecasting in Regional Travel Models." *Transportation Research Record: Journal of the Transportation Research Board*.
- Stevens, Mark R. 2017. "Does Compact Development Make People Drive Less?" *Journal of the American Planning Association* 83 (1): 7–18. <https://doi.org/10.1080/01944363.2016.1240044>.
- Stewart, Orion T., Anne Vernez Moudon, Brian E. Saelens, Chanam Lee, Bumjoon Kang, and Mark P. Doescher. 2016. "Comparing Associations Between the Built Environment and Walking in Rural Small Towns and a Large Metropolitan Area." *Environment and Behavior* 48 (1): 13–36. <https://doi.org/10.1177/0013916515612253>.
- Talen, Emily, and Julia Koschinsky. 2014. "Compact, Walkable, Diverse Neighborhoods: Assessing Effects on Residents." *Housing Policy Debate* 24 (4): 717–50. <https://doi.org/10.1080/10511482.2014.900102>.
- US EPA, OA. 2014. "Smart Location Database Technical Documentation and User Guide." Data and Tools. US EPA. February 27, 2014. <https://www.epa.gov/smartgrowth/smart-location-database-technical-documentation-and-user-guide>.
- Zhang, Qin; Clifton, Kelly J.; and Moeckel, Rolf, 2018. "Investigation of an Appropriate Spatial Resolution for Large-scaled Pedestrian Travel Demand Models." presented at Mobil.TUM 2018 Urban Mobility: Shaping the Future Together, Munich, Germany, June 13-14, 2018.



NITC

NATIONAL INSTITUTE for
TRANSPORTATION and COMMUNITIES
Transportation Research and Education Center
Portland State University
1900 S.W. Fourth Ave., Suite 175
Portland, OR 97201