This paper is published in *Transportation Research Part D: Transport and Environment* and is referenced as:

Wali, B., Frank, L. D., Chapman, J. E., & Fox, E. H. Developing policy thresholds for objectively measured environmental features to support active travel. *Transportation Research Part D: Transport and Environment*, *90*, 102678.

Developing Policy Thresholds for Objectively Measured Environmental Features to Support Active Travel

Behram Wali, Ph.D.
Lead Research Scientist
Urban Design 4 Health, Inc.
24 Jackie Circle East
Rochester, NY 14612
&
Postdoctoral Scholar
Senseable City Lab
Massachusetts Institute of Technology
bwali@ud4h.com; walibehram@yahoo.com

Lawrence D. Frank, Ph.D.
President
Urban Design 4 Health, Inc.
24 Jackie Circle East
Rochester, NY 14612
ldfrank@ud4h.com

Jim E. Chapman, MSCE
Managing Principal
Urban Design 4 Health, Inc.
24 Jackie Circle East
Rochester, NY 14612
jchapman@ud4h.com

Eric H. Fox, MScP
Senior Planner
Urban Design 4 Health, Inc.
24 Jackie Circle East
Rochester, NY 14612
efox@ud4h.com

Developing Policy Thresholds for Objectively Measured Environmental Features to Support Active Travel

Behram Wali, Lawrence D. Frank, Jim E. Chapman, & Eric H. Fox

ABSTRACT

A novel evidence-based methodology is presented for determining place-based thresholds of objectively measured built environment features' relationships with active travel. Using an innovative machine-learning based Generalized Additive Modeling framework, systematic heterogeneity fundamental to the development of well-justified and objective environmental thresholds is accounted for. The methodology is employed to model an individual's likelihood of transport walking as a function of environmental factors using California Household Travel Survey linked with comprehensive built environment data. The results reveal strong and complex non-linear dependencies of likelihood of transport walking on environmental features that cannot be quantified using standard threshold detection methods. Thresholds for key environmental features to enhance active travel vary significantly across different socioeconomic groups. Accounting for strong income-based differences in development of environmental benchmarks is emphasized. The thresholds can serve as a useful guiding tool for policymakers, planners, engineers, and public health officials to track existing environmental conditions and healthy behaviors.

Keywords: Walking for transport, objectively measured built environment, built-environment thresholds, Generalized Additive Models, thin plate regression splines, systematic heterogeneity.

1. INTRODUCTION

Physical activity is an integral lifestyle component for improving public health. Engagement in physical activity has declined substantially over the last few decades, having serious implications for population health. Approximately 3.2 million deaths each year globally are attributed to insufficient physical activity. Thus, promoting regular physical activity has emerged as a global public health priority (WHO 2010). The linkage between built environment and health outcomes through physical activity is now one of the most studied pathways in the literature (Ewing et al. 2015, Frank et al. 2019a), with studies emphasizing importance of socioecological models for providing evidence-based theoretical underpinnings for designing necessary interventions (Sallis et al. 2006, Ding et al. 2012). Although strong associations between built environment factors and physical activity exist (Ewing et al. 2015, Sarkar et al. 2015, Kang et al. 2017, Frank et al. 2019a), it remains a challenge to develop and modify built environments that encourage increased physical activity.

The idea of establishing thresholds is logically attractive from a policy perspective. Researchers and policy makers are often asked, “How dense or mixed do land uses need to be to get people to walk or take transit?” Past attempts to address these questions have suffered from analytical challenges addressed in part in the current study. There is little evidence regarding development of optimal thresholds/benchmarks of key variables related to walkability. Empirical evidence regarding optimal thresholds of key environmental variables is needed to promote health. While there are contextual factors unique to specific locations impacting generalizability, certain basic built environment features are required before walking is viable. Research is needed for transportation planners and policymakers to make an evidence-based case that retrofitting environments to promote active living is cost-effective (Kaczynski et al. 2009, Koohsari et al. 2013). That is, how much of specific types of changes to the built environment is required to realize shifts in travel patterns that will support healthier outcomes.

1.1. Research Objective

This study focuses on developing a clear, evidence-based methodology for creating thresholds for various environmental measures that support active travel and physical activity. The key research issue is to develop a methodology for understanding where the thresholds for key built environment benchmarks should be set for characterizing neighborhood level support for active travel. The thresholds can serve as a useful guiding tool for policymakers, planners, engineers and public health officials to track physical activity conditions of communities over time.

In this study, thresholds are established for four objectively measured built environment variables using an innovative machine-learning based Generalized Additive Modeling (GAM) framework with P-smoothing/cubic splines. This approach allows empirical identification of the potential thresholds from the data at hand rather than a priori imposing untested assumptions. The methodology is employed on a large survey for modeling the likelihood of transport walking. Data from the 2013 California Household Travel Survey are used to model likelihood of transport walking as a function of environmental factors and other controls. Objectively measured environmental data (as opposed to perceived environmental features) from the Robert Wood Johnson Foundation's National Environmental Database (NED) metrics are used, which objectively tracks how places within a given geographical area (e.g., census block group or tract) compare in terms of the level of support for active living and enhanced public health.

2. LITERATURE REVIEW

The potential to encourage active transportation by altering the built environment is undisputable (Cervero and Kockelman 1997, Ewing et al. 2015, Sarkar et al. 2015, Frank et al. 2019a). Both from the perspective of internal and external validity, strong correlations between subjective as well as objective measures of the built environment and travel behavior are well-documented (Ewing et al. 2015). However, from a threshold development standpoint, existing research on the built environment, land-use, and physical activity (such as transport walking) has mostly focused on "linear" relationships between specific attributes and physical activity (Brownson et al. 2001, Huston et al. 2003); e.g., greater street connectivity is associated with more walking. However, the correlations between built environment attributes and physical activity may be non-linear (Frank et al. 2004, Eom et al. 2015). For instance, the negative relationship between the distance to a park and walking may be negative in general but can be more pronounced when parks are located beyond a certain distance (such as greater than 400 meters) from a residents' home location. In other words, a deeper understanding of the optimal range of key environmental metrics (e.g., residential density) that support healthy outcomes can be helpful to practitioners. Distance decay is a core feature of the gravity model used widely in transportation planning and understood to be a fundamentally non-linear concept (Fotheringham 1981). However, the linear assumption also implies that one cannot infer a minimum threshold on an environmental attribute that must be exceeded before significant improvements in physical activity can be expected (trigger effects). It also precludes identification of a possible plateau or "ceiling effect", e.g., beyond a certain level of density, mix, or other features, only marginal changes in behavior and health benefits can be observed.

Notably, the issue of nonlinear dependencies of physical activity and healthy behaviors on environmental features is not just widely recognized in the literature (Kaczynski et al. 2009, Eom et al. 2015), but has been explored as far back as 1994 (Frank and Pivo 1994). However, these nonlinear dependencies are not yet adequately covered in a methodologically rigorous manner for development of thresholds[1] (except for few studies described that follow). Without

[1] As discussed earlier, for establishing reliable place-based thresholds, it is crucial to accommodate variations in individuals' responsiveness to level of changes in environmental attributes, i.e., systematic heterogeneity contours (Pinjari and Bhat 2006, Wali et al. 2018c). In particular, nonlinearity is an outgrowth of systematic heterogeneity due to variations in "observed" factors. In the literature, systematic response heterogeneity is typically accounted for by investigating differences in response sensitivities of environmental factors due to individual demographic and/or other relevant attributes, see (Frank et al. 2004, Ding et al. 2012, Van Dyck et al. 2012b). For instance, it is typically captured by interacting individual characteristics with environmental factors, i.e., the response of a respondent to residential density may be a function of an individual's gender or age. For the purpose of developing place-based

necessarily focusing on threshold development, a predominant approach to capturing systematic heterogeneity in the individuals' responsiveness to environmental attributes is the use of interactions, i.e., interacting environmental features with an individual's characteristics (Table 1). Evidence suggests that while the relationships of neighborhood environment with leisure walking differs among driving status, the relationships do not vary between driving status for transport walking (Ding et al. 2012). Likewise, relationships between urban form, travel patterns and obesity have been shown to differ by race (Frank et al. 2004). More recently, researchers have explicitly analyzed systematic heterogeneity arising from nonlinear responsiveness of individuals' physical activity to environmental factors. For instance, Van Dyck et al. (2012) examined the relationships between perceived neighborhood attributes and adults' self-reported sedentary behavior in a multi-country sample (Van Dyck et al. 2012a). They observed non-linear relationships between land use mix diversity measures (proximity of destinations and number of types of destinations) and (self-reported) weekly minutes of motorized transport (Van Dyck et al. 2012a). Similarly, nonlinear relationships of unknown form were quantified for perceived environmental attributes and adults' leisure-time physical activity (Van Dyck et al. 2013). However, the study did not discuss any apparent thresholds for (perceived) environmental factors. The key strength of both studies was the use of pooled data from culturally and environmentally diverse countries (Van Dyck et al. 2012a, Van Dyck et al. 2013). However, the studies focused on perceived neighborhood environmental attributes and which may differ from actual environmental features.

benchmarks/thresholds, we clarify that our emphasis is on another key element of systematic heterogeneity, i.e., systematic heterogeneity arising from nonlinear responsiveness of an individuals' physical activity to environmental factors. More detailed discussion on the importance of systematic heterogeneity due to nonlinear responsiveness can be found in (Pinjari and Bhat 2006, Wali et al. 2018b, Wali et al. 2018c, Wali et al. 2020).

Recently, the nonlinear associations between objectively measured built environment and commute mode choice and vehicle use (driving distances) are also explored. Using gradient boosting decision trees, strong nonlinear dependencies between built environment (around home and work locations), commute mode choice, and driving distances were quantified (Ding et al. 2018a, Ding et al. 2018b). Importantly, built environment variables collectively contributed to 65% of the predicting power for car mode choice (Ding et al. 2018b). Most in line with the present study, a recent paper focused on finding the desirable values (thresholds) of objectively measured D-variables to increase active transportation[2] (Park et al. 2020). For 28 metropolitan regions of the U.S., cross-sectional travel behavior and built environment data were linked to estimate GAMs. By characterizing the complex nonlinear dependencies, careful recommendations were provided regarding desirable ranges of key built environment (BE) metrics (see Table 1). The study demonstrated the value of GAMs in approximating complicated non-linear dependencies that could not be objectively approximated using simpler polynomial models (Park et al. 2020). However, the study did not focus on likely differences in travel behavior arising due to socioeconomic status of individuals. The present study contributes by developing and applying a rigorous methodology to establish place-based thresholds using comprehensive travel behavior and objectively measured built environment data. In doing so, potential differential income effects are also explicitly considered and analyzed.

[2] While zealously acknowledging and appreciating the valuable contribution of Park et al. (2020), we wish to note that the present study was formally conceived in mid-2017 as part of a research project funded by the Robert Wood Johnson Foundation and completed independently of the work reported in Park et al. (2020). The authors learned about the work of Park et al. (2020) through a personal interaction with Dr. Keunhyun Park following a lectern session during the 2020 Transportation Research Board Annual Meeting in which findings from the present study were presented.

Table 1: Relevant Studies About Relationships of Neighborhood Environment with Physical Activity

Measures/Explanatory Factors	Key Outcome	Focus	Statistical Method	Approach to capture systematic heterogeneity*	Key Results
Neighborhood environment: street connectivity, walking-bicycling infrastructure, traffic safety, pedestrian safety structures, & overall microscale sum score **(Ding et al. 2012).**	Objectively measured physical activity (transportation & leisure) among adults	Relationship between neighborhood environment and physical activity across driving status	Mixed linear regression, Mixed generalized linear regression	Interactions between environmental factors & individual characteristics	1) Effect of neighborhood environment on leisure walking differed between driving status. 2) Effect of neighborhood environment on transport walking did not differ between driving status.
Urban Form: connectivity, net residential density, land-use mix **(Frank et al. 2004).**	Self-reported travel patterns (walking and time in a car), BMI, & obesity	Relationship between urban form and health differ across ethnicity and gender	Logistic regression	Same as above	Relationships between urban form, travel patterns and obesity differed across whites and blacks
Perceived neighborhood environment: residential density, land use mix diversity, land use mix access, street connectivity, pedestrian facilities, etc. **(Van Dyck et al. 2012a).**	Adults' sedentary behavior	Strength, shape, and direction of relationship between perceived environmental factors & sedentary behavior	Generalized additive models	Explicitly considered nonlinear responsiveness of individuals' physical activity to environmental factors.	Strong non-linear dose-response relationships of land use mix diversity measures (proximity of destinations and number of types of destinations) and (self-reported) weekly minutes of motorized transport
Perceived neighborhood environment: residential density, land use mix diversity, land use mix access, street connectivity, pedestrian facilities, etc. **(Van Dyck et al. 2013)**	Adults' leisure-time physical activity	Strength, shape, and direction of relationship between perceived environmental factors & leisure-time physical activity	Generalized additive models	Same as above	1) Curvilinear main effects were observed for walking and cycling facilities, residential density, and aesthetics 2) Gender was a significant moderator of the relationships of leisure walking with aesthetics and crime safety
Objective neighborhood environment: distance to city center, distance to local center, population density, employment density, transit zone **(Ding et al. 2018a).**	Driving distance	Examining non-linear effects of the built environment on driving distances in Oslo	Gradient boosting decision trees	Considered nonlinear relationships through machine-learning classifier	1) Non-linear relationships between built environment (BE) and driving distances 2) On weekdays, the collective influence of BE is larger than that of demographics.

Notes: BMI is Body Mass Index; (*) See footnote #1 for a description of approaches to capture systematic response heterogeneity and its relevance to creating place-based thresholds/benchmarks; CBD is Central Business District.

Table 2: Relevant Studies About Relationships of Neighborhood Environment with Physical Activity (Continued)

Measures/Explanatory Factors	Key Outcome	Focus	Statistical Method	Approach to capture systematic heterogeneity*	Key Results
Objective built environment at home and work locations: population density, employment density, land use mix, bus stop density, distance from CBD, etc. **(Ding et al. 2018b).**	Commute mode choice (car, transit, walk, bicycle)	Non-linear effects of the built environment around home and work locations on mode choice	Gradient boosting decision trees	Considered nonlinear relationships through machine-learning classifier	1) Built environment variables collectively contribute to 65% of the predicting power for car mode choice 2) Distance from workplace to CBD has a much larger impact on car mode choice than distance from residence to CBD.
Objective neighborhood environment: activity density, job-population balance, intersection density, % of 4-way intersections, count of transit stops, etc. **(Park et al. 2020).**	Travel mode choices and vehicle use	Finding desirable values of D-variables to reduce vehicle use and increase walking and transit use	Generalized additive models	Same as above	1) Strong non-linear relationships between D-variables & travel behavior 2) Recommendations to increase walking/transit use: activity density (pop + emp/sq.mi.) (20K - 25K), job-population balance (minimum 0.2-0.5), intersection density (minimum 150-300), transit stop density (minimum 25 for small center to minimum 150 for large centers)

Notes: (*) See footnote #1 for a description of approaches to capture systematic response heterogeneity and its relevance to creating place-based thresholds/benchmarks; CBD is Central Business District.

3. METHODS

3.1. Conceptual Framework

The key focus of the present study is to define a methodology that can guide the development of place-based thresholds of key objective environmental variables for characterizing active transportation at the neighborhood level. Fundamentally, the issue of developing place-based thresholds for key environmental variables translates to the categorization of continuous variables in physical activity prediction. The number of classes (thresholds) into which a continuous built environment variable can be categorized, depends mainly on the relationship between the predictor and outcome. The need for adequate characterization of the relationship must be carefully considered. While the application of linear models can help identify thresholds, these models significantly limit the ability to identify thresholds that reflect the 'true' relationship between a predictor and an outcome. Alternatively, this assumption of linearity precludes identification of possible *'trigger'* and *'ceiling'* effects. While non-linear models (non-linear in variables and/or non-linear in parameters) offer a solution, several important caveats exist. First, it is seldom obvious how to specify the non-linear functional form that best fits the data. Second, the use of piecewise regression as an alternative method also requires a priori identification of cut-points in the data, and it is always subjective where to split a piecewise regression. To address these issues, an innovative machine-learning based Generalized Additive Modeling (GAM) framework with P-smoothing/cubic splines is employed to empirically identify the thresholds from the data at hand rather than imposing untested assumptions a priori. Addressing the issues stated above, GAM offers a data-driven customized non-linear method for identifying thresholds for key environmental factors. Importantly, the identification of thresholds in GAM is not independent of physical activity; rather, it is informed by the relationships between key benchmarks and the physical activity outcome[3]. As a result, it facilitates the identification of thresholds within the benchmark variables that can distinguish between poorer

[3] One alternative to generalized additive modeling could be polynomial-type non-linear regressions. Conceptually, compared to generalized additive models (GAMs), polynomial regression methods, while relatively simpler to understand, comes at a cost of substantial subjectivity. Polynomial models require the analyst to make prior assumptions about the functional forms for specific variables. As noted in the relevant literature (Wali et al. 2018b, Ding et al. 2018a, Park et al. 2020), correct identification of appropriate non-linear functional form that best fits the data is almost always not obvious. Further, due to the underlying complex and context-specific dependencies of travel behavior on environmental features, no definitive answers exist in the literature regarding appropriate functional forms. To this end, by blending machine learning and inferential methods, GAMs offer an intuitive approach to characterize complex dependence structures in an objective and defensible way without imposing any unrealistic and inappropriate assumptions on the data a priori (Van Dyck et al. 2012b, Van Dyck et al. 2012a, Van Dyck et al. 2013, Hu et al. 2018, Wali et al. 2018b, Park et al. 2020).

and better physical activity outcomes at a specific geographic scale (such as census block group, tract, county and state). A visual description of the concept follows in Figure 1 below.

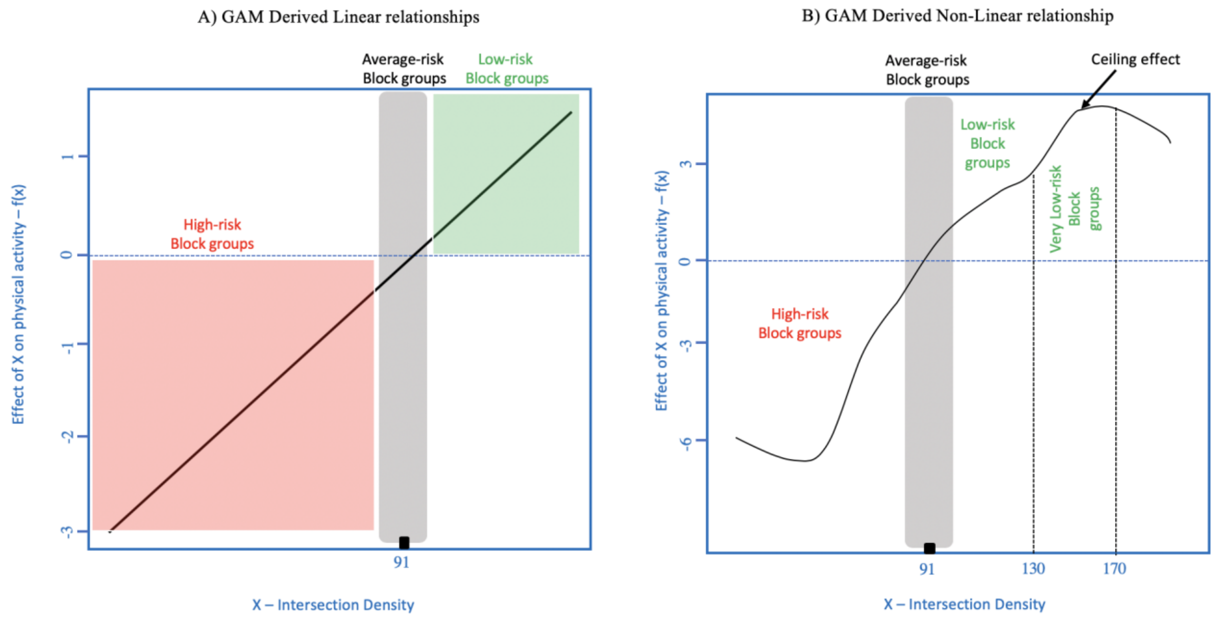


Figure 1: Graphical Illustration of Two Hypothetical Relationships between Built-Environment and Physical Activity

(Notes: GAM is Generalized Additive Model; Raw values of predictor (hypothetical intersection density) are plotted on X-axis; Smoothed values of hypothetical predictor effects obtained from GAM plotted on Y-axis).

As shown for a hypothetical linear relationship (Figure 1A – left panel), an increase in intersection density correlates with a linear increase in the likelihood of physical activity (such as walking). As a starting point, an average-risk category can be identified based on where the slope line intersects the "zero-effect" or "50-50 chance" reference point. In this case, a hypothetical intersection density value of 91 intersections per sq. kilometer defines the average-risk census block groups, whereas an increase beyond 91 defines low-risk census block groups (i.e., block groups where residents have a greater than 50% likelihood of participation in transport walking). As shown, a threshold of 91 can be derived only if the relationship between intersection density and physical activity is truly linear. If the 'true' relationship is non-linear (Figure 1B), application of GAM can help extract valuable information directly relevant to the project scope.

First, referring to Figure 1B, an average-risk category can be defined in a similar fashion based on where the non-linear slope line intersects the zero-effect reference point which indicates a 50-50 chance (in this case 91). In terms of negative dependencies, the likelihood of participating in physical activity is below 50% (albeit increasing) up until an intersection density of 91. These can be regarded as "high-risk" block groups. An increase beyond this point results in an above 50% chance of participation in physical activity (transport walking). In particular, based on the non-linear contour derived from hypothetical GAM analysis, a block group with an intersection density value greater than 91 (and less than 130) can be defined as "low-risk block group", whereas an intersection density value greater than 130 but smaller than 170 can be categorized as "very low-risk" block groups, given the sharp increase in the physical activity participation slope between 130 and 170. Lastly, a ceiling effect can also be observed for block groups with intersection density greater than 150 - in that only marginal improvements in physical activity are observed. Using this richer set of information allows for the development of context-specific thresholds for key benchmarks.

3.1.1. A Case for Objectively Measured Environmental Features

The usefulness of objectively-measured environmental data is well-demonstrated in the transport, planning, and health literature (Saelens et al. 2003, Owen et al. 2004, Sallis et al. 2004, Frank et al. 2005, Leslie et al. 2007). A supportive objective built environment is necessary for enabling greater physical activity (Ma and Dill 2015). In the present study, given the key focus on examining optimal values of built environment, the use of objective measures of built environment is fundamental to answering the key question of how much and what types of infrastructure is required to support healthy outcomes. Determining thresholds or optimal values for individuals' perceptions of the environment may be less useful from a policy standpoint since retrofitting an existing environment based on individuals' perception could be challenging especially when perceptions do not reflect actual environments in the context under discussion (see below). Also, engineers and planners work with real on-ground physical features and cannot typically control for individuals' perceptions. Perception-based environmental data are often used as substitutes when objectively measured environmental data are not available (Ma and Dill 2015). When both are available (unlike the case in the present study), perception-based data should be included in the spirit of controlling for important psychological factors that may help explain individuals' behaviors (Saelens et al. 2003). Social cognitive theory and theory of planned behavior highlights an important distinction between the built environment as it is objectively evaluated and the built environment as perceived by users (especially vulnerable road users) (Bandura 1986). Considering perception-based data as a substitute for objective data, when the latter is available, can lead to different conclusions because it remains to be determined how an individual's perception regarding the built environment reflects reality. In fact, when the few times both objective and perceived data are considered together (Kirtland et al. 2003, McGinn et al. 2007, Ma and Dill 2015), a poor agreement between the two measures of the built environment is found (Kirtland et al. 2003, McGinn et al. 2007). Given this lack of agreement between the two, little or no change in the point estimates of objective BE variables was found when combining both objective and subjective measures into one model (McGinn et al. 2007). Also, the associations of self-reported, objective environmental data, and physical activity are quite different and often contradictory (McGinn et al. 2007, Ma and Dill 2015). Thus, when possible, both subjective and objective measures should be considered (Saelens et al. 2003, Ma and Dill 2015). As part of future work, with the availability of necessary data, subjective data on environmental features should also be considered to control for individuals' psychological factors as optimal values for key objective environmental features are examined (e.g., more fine-grained measures of subjective and objective measures collected through recent studies could be used (Frank et al. 2019b, Wali et al. 2021)).

3.2. Data

This study used data from two key publicly available sources: (1) the California Household Travel Survey (CHTS) and (2) the National Environmental Database (NED) (developed by Urban Design 4 Health (UD4H) for the Robert Wood Johnson Foundation). In this study, the travel diary data are used to extract information on individual (person-level) participation in transport walking. Components of the CHTS data set used in this analysis include: (1) geocoded household, trip, and activity location shapefile, (2) household characteristics table, (3) individual characteristics table, (4) individual trip table, and (5) individual activity table. Using the most recent 2013 CHTS housed in the National Renewable Energy Laboratory (NREL) Transportation Secure Data Center (TSDC), a total of 42,454 households in California is considered. After data cleaning (missing data, outliers, incomplete survey, etc.), a total of 62,778 adult participants (belonging to 33,644 households) is retained. Given that physical activity thresholds are highly likely to be context-specific, the threshold development and analysis presented next is conducted only for adults (18-65 years of age) in metropolitan areas resulting in a sample of 44,253 respondents belonging to 24,842 households. The metropolitan areas were designated according to a 2010 U.S. Department of Agriculture (USDA) rural to urban categorization system for census tracts (USDA 2010).

Data on key built environment/environmental factors at census block group (CBG) level are extracted from the UD4H-developed NED and matched with geo-coded CHTS travel behavior data. Funded by the Robert Wood Johnson Foundation (RWJF), the NED compiles a unique set of nationally consistent and comprehensive standardized set of built, natural, and social environmental metrics from multiple sources to assess how local environments impact health outcomes. These metrics are environmental predictors of physical activity, obesity, travel behavior and health outcomes in one or more population subgroups to support a *Culture of Health*. Five core built-environment factors used in the creation of National Walkability Index, work which was also funded by RWJF, are considered in this study:

1. Residential Density (density)
2. Intersection Density (connectivity)
3. 8-Tier Employment Mix (diversity)
4. Access to transit stops (transit)
5. Vehicle Miles Travelled (automobile use)

3.3. Environmental Measures Descriptions

Residential density is measured as total housing units per acre of unprotected land (non-conservation protected land, such as parks and natural areas) and relies on base data from the U.S. Census Bureau (Frank et al. 2003). Street connectivity is measured as pedestrian-weighted intersections with automobile-orientated intersections (intersections at high speed, high traffic volume and limited access roadway links removed) per square miles of land and is derived from the U.S. EPA Smart Location Database. The 8-tier employment entropy is a measure of the diversity of job types based on eight employment classifications (Cervero and Kockelman 1997): 1) retail, 2) office, 3) industrial, 4) service sector, 5) entertainment, 6) education, 7) healthcare and 8) public administration. Employment base data is based on the Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) furnished by the U.S. Census Bureau. This variable considers the relative similarity of the number of jobs by the static eight employment types such that:

$$8\text{-}tier\ employment\ entropy = -E/(\ln(N)) \qquad \text{Eq. 1}$$

Where: E contains each of the employment classification types and N is the number of employment types with jobs > 0. Lastly, access to public transit is measured by the count of transit stops from all rail and bus modes equally weighted based on General Transit Feed Specification (GTFS) sources released by transit agencies nationwide. An important measure of travel behavior related to automobile usage is annual household Vehicle Miles Traveled (VMT). This measure was acquired from the U.S. DOT and U.S. Department of Housing & Urban Development (U.S. HUD) based on a series of sources including national driving records from the National Household Travel Survey (NHTS). The acquired VMT variable was modeled using ordinary least squares (OLS) regression to provide nationwide estimates of automobile usage. VMT is included in the National Walkability Index (components of which are used in this study) since it is the most understood measure of car dependence (McIntosh et al. 2014). People who drive more walk less. In this study, VMT gives an overall picture of average household behavior in a CBG. VMT as a measure of car dependence is inversely related to walking and picks up the observed CBG-level car-dependence behavior that could be due to the underlying hindrance to active travel in terms of less supportive built environment. Also, VMT is widely viewed as the strongest single correlate of environmental degradation and resource consumption in the transport sector (Cervero and Murakami 2010). Along this line, inclusion of VMT as a control helps capture the key exposure pathways between built environment and active travel (Frank et al. 2019a) – which is a key but often ignored pathway in the literature (Frank et al. 2019a). For example, with higher VMT (greater car dependence), the amount of local pollution (e.g., particulate matter) and global pollution (e.g., greenhouse gas emissions) increases – ultimately increasing the exposure related externalities associated with active travel. Thus, controlling for

VMT in the analysis helps provide a fuller and broader picture of the determinants of walkability.

3.4. Statistical Methodology for Identifying Thresholds

The response outcome in this study is a binary indicator of participation in walking (0/1) (over the one-week period preceding respondents' survey day) at the person-day level. Owing to the discrete nature of the response outcome, the Generalized Additive Binary Logit Model (GABLM) approach was selected. GABLM is an extension of a standard Binary Logit (BL) model where the modeling of the mean utility function relaxes the assumption of linearity. A more detailed overview of the methodological aspects of GABLMs have been published elsewhere (Wali et al. 2018b). In a standard binary logit model, the key quantity of interest is the mean value of the response outcome given the values of the independent variables. In developing the standard logistic regression equation, the logarithm of the odds represents a logit transformation, where the logit is a function of explanatory factors such that:

$$Y_i = logit(P_i) = LN\left(\frac{P_i}{1 - P_i}\right) = \beta_o + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \cdots . + \beta_K X_{K,i} \qquad \text{Eq. 2}$$

Where: β_o is the model constant term and the $\beta_1 \ldots . \beta_K$ are the unknown parameters to be estimated corresponding with the independent variables ($X_K, k = 1, \ldots \ldots, K$ the set of explanatory variables). The main limitation of Eq. 2 is the assumption of linearity. That is, all explanatory factors are assumed to linearly influence the response outcome (i.e., probability that a person chooses transport walking). To account for complex nonlinearities without imposing assumptions a priori, GABLMs are introduced. Following Wood (2017), GABLMs exhibit smooth functions (such as kernel smoothers, lowess, smoothing splines or regression splines) for explanatory factors rather than single parameter estimates as in the case of BL models. If spline terms are included for all explanatory factors, the model specification adjusts to be fully non-parametric (Wood 2017, Wali et al. 2018b). On the other hand, if the mean functions of some covariates are assumed to be linear (and in which case a single β is estimated) while the mean functions of other covariates are represented as splines, the model specification turns to be semi-parametric. In general, the model has the following structure:

$$g(u) = \alpha_o + \sum_{j=1}^{K} \beta_j X_j + \sum_{i=1}^{P} f_i(Z_i) \qquad \text{Eq. 3}$$

Where: $u = E(Y)$ for Y (from Equation 2), g is the link function (such as logit link), and f_i are some smoothing functions corresponding to explanatory factors Z_i for each explanatory factor whose mean functions are assumed to be non-linear, $i = 1, \dots., P$. Whereas β_j is a set of parameter estimates for explanatory variables which are assumed to be linear (as in a standard logit model). GABLMs provide greater flexibility than BLs, as they relax the strong (and often unrealistic) hypothesis of linear dependence between the correlates and the expected value of the response outcome. Following Wood (2017), there are different types of spline smoothing bases (such as cubic regression splines, cyclic cubic regression splines, P-splines, thin plate splines, and thin plate regression splines) that can be used to estimate f_i. In this study, the authors use a technique known as thin plate regression splines that can be used for producing univariate knot free bases, for details see (Wood 2017, Wali et al. 2018b). For characteristics of thin plate regression splines and estimation of smoothing parameters, see (Wali et al. 2018b). Theoretically, if any smooth functions were considered/allowed in model fitting then the maximum likelihood estimation of GAMs would result in complex over-fitting estimates of the smooth functions (Wood 2017). For this reason, using penalized likelihood maximization, the model (negative log) likelihood is updated by the inclusion of a penalty for each smooth function associated with a variable, penalizing its ‘wiggliness’ (Wood 2017). To control the trade-off between penalizing wiggliness vs badness of fit (in summary to avoid overfitting), each penalty is multiplied by an associated smoothing parameter (Wood 2017). The selection of optimal smoothing parameter is conducted using Un-Biased Risk Estimator (UBRE) criterion. For more details about the regularization scheme used to avoid overfitting, see Wood (2017).

3.4.1. Modeling Strategy

To estimate dose-response relationships for key environmental correlates, we estimate two sets of models (Table 2). Referred to as Strategy 1, the first set of models estimated the dose-response relationships of single objectively measured environmental factors (as discussed in section 3.3) with participation in transport walking, adjusting for sociodemographic, household size and tenure, income, age, locality indicators and employment/education status. The modeling specification for the first set of models is:

$$g(u) = \alpha_o + \sum_{j=1}^{K} \beta_j X_j + f_i(Z_i) \qquad \text{Eq. 4}$$

Where: Z_i represents each of the five environmental covariates whose mean functions are assumed to be non-linear (as described above). That is, while controlling for all other factors (as listed in Table 3 below), five GABLMs are estimated each with a single environmental variable at a time. The non-parametric functions $f_i(Z_i)$ provide a more complete approximation of the effects of the environmental variables on outcome than would have been achieved by $\beta(Z_i)$, because the former does not assume that the effect of an environmental variable on outcomes is necessarily fixed or constant. The second set of models (referred to as strategy 2) estimated dose-response relationships of multiple objectively measured environmental factors with participation in transport walking, adjusting for all other sociodemographic, locality indicators, household size and tenure, and other controls listed in Table 3. Considering multiple environmental measures is important since urban form factors tend to covary across space and a natural synergy for environmental features could exist (greater land use mix is usually found in areas with high residential density and better connectivity – all of which are associated with higher active transportation) (Frank and Engelke 2001, Sevtsuk et al. 2016). Methodologically, considering only a single environmental feature in a model specification (as is done under Strategy 1) can lead to omitted variable bias problem, where the estimated correlation between the single environmental feature and active transportation could change if other environmental features are simultaneously considered in the same model specification (Frank and Engelke 2001, Van Dyck et al. 2013, Wali et al. 2018c). Mathematically, the following specification was estimated:

$$g(u) = \alpha_o + \sum_{j=1}^{K} \beta_j X_j + f_i(Residential\ density) + f_i(Intersection\ density) + f_i(Employment\ mix) + f_i(Access\ to\ transit\ stops) + f_i(Vehicles\ miles\ travelled) \quad \text{Eq. 5}$$

The key motivation behind testing strategies 1 and 2 was to examine if considering multiple environmental variables in a same model specification affects the contours of relationships observed in model with single environmental variables at a time (Eq. 4). Finally, to account for likely differences arising due to socioeconomic status of the participants, the models were estimated using the aggregate data (i.e., adult participants in metro regions) as well as stratified data by income-level, i.e., low-income metro adults, medium-income metro adults, high-income metro adults[4].

[4] Compared to a full segmentation approach (i.e., estimating separate models for low, medium and high-income groups), a more parsimonious approach could be to conduct a partial segmentation analysis. In particular, the income indicator can be interacted with the non-parametric smooth functions associated with built environment features. We present the results obtained from the partial segmentation approach later in section 4.

Table 2: Modeling Scheme

Modeling Strategy	Predictors	Aggregate Data	Low-income Metro Adults	Medium-income Metro Adults	High-income Metro Adults
Strategy 1	Employment Mix + All controls	×	×	×	×
	Residential Density + All controls	×	×	×	×
	Number of Transit Stops + All controls	×	×	×	×
	Vehicle Miles Travelled + All controls	×	×	×	×
	Intersection Density + All controls	×	×	×	×
Strategy 2	**All Five Walkability Variables + Controls:** Employment Mix + Residential Density + Number of Transit Stops + VMT + Intersection Density + All controls	×	×	×	×

3.5. Threshold Selection

Recognizing the need for adequate characterization of the relationship between environmental factors and participation in transport walking, the number of categories as well as the location of the cut points will depend on the graphical contours of relationships obtained from GABLMs with thin-plate regression splines. Based on these models (Eq. 4 and Eq. 5 above), we develop visualizations showing the relationship between the raw values of Z (environmental factors) plotted on x-axis and the probability of walking on y-axis (obtained from smoothed values in $f(Z)$)[5]. For a specific environmental variable such as residential density, the resulting graph obtained from GABLMs provides a point-by-point estimate of the effect of each of the continuous levels of residential density on the outcome. A positive slope in the GABLM graph

[5] An alternative and equivalent way would be to plot the raw values of Z (environmental factors) on x-axis and the "smoothed" values obtained from $f(Z)$ on y-axis. However, as probability-based contours are relatively easier to interpret, we prefer to work on the scale of probability for threshold selection. If visualizations are created on the basis of "smoothed" values obtained from $f(Z)$, the individual contour plots will be on the scale of the linear predictor (as is formulated through the utility function of a logit model), i.e., a scale that ranges between $-\infty$ and $+\infty$. Following guidelines by Wood (2006), $f(Z)$ in our case is the centered mean function where the centering coefficient is the constant term, α_o. In summary, $f(Z) = 0$ refers to the average value of the explanatory factor and in the context under discussion represents a 50-50 chance of an adult in metropolitan region participating in transport walking. To shift to the probability scale, we use the inverse transformation of the GABLM link function (in this case logit link) to map from the linear predictor scale to the probability scale (0-1).

suggests an increasingly stronger effect of an environmental factor on an outcome. Contrarily, a negative slope indicates a decreasing or weaker effect. In particular, the portion of the contour above the 50-50 chance line indicates the range of an explanatory variable which leads to greater than 50% chance of an individual participating in transport walking. Thus, the thresholds for a particular environmental variable are based on where the non-linear contour line intersects the 50-50 chance line (more details to follow). In most standard applications of discrete choice models, a default and typical threshold of 0.5 or 50% is used to classify a prediction as positive outcome (1 vs. 0) (Hensher et al. 2005, Greene and Hensher 2010, StataCorp 2013). That is, if the predicted probability is greater than 0.5 (greater than a 50-50 chance or “more likely than not”), then that prediction is classified as a positive outcome (participation in transport walking). However, the definition of minimum probability defining a threshold can vary across users – which can bring in a qualitative aspect to the interpretation of quantitative results. To this end, we provide applicable probability ranges in presenting the thresholds in the results section so the reader/user can select appropriate threshold based on their preferences.

As discussed in section 3.1, threshold and ceiling effects (if any) can be identified as well, providing a fuller understanding of the effects of a particular environmental covariate. For example, if there is a minimum threshold on the residential density that must be exceeded before significant improvements in likelihood of transport walking can be observed, it is expected that there will be a flat (or negative) slope in the GABLM graph, followed by a sharp positive slope likely indicating an upsurge in effect. Likewise, if there is a threshold indicating overdose effects such that improvements in, for example, residential density beyond this threshold are correlated with only marginal gains (or reduction) in physical activity, a flat or negative slope in the GABLM plot will then be observed in that particular range of environmental variable (such as residential density). Finally, as models in strategy 2 can better account for the effects of (likely) correlated environmental variables (empirical evidence below), we base the selection of thresholds on the results obtained under Strategy 2. All the statistical analysis is conducted in R software using mgcv package for model estimations (Wood 2017).

4. RESULTS

4.1. Descriptive Statistics

Table 3 presents the descriptive statistics of participation in transport walking (response outcome), key walkability related environmental variables, and additional controls including household size and structure, sociodemographic, education, and family composition (Table 2). Based on their distributions, key summary statistics, and extraction from a well-organized, integrated, and validated database, the underlying data are of reasonable quality. As shown in Table 3, around 16% of the respondents participated in transport walking over the one-week period preceding their survey day. Significant variations are observed in terms of how walkable the sampled block groups are across California. Similar patterns of large variations are observed for number of transit stops and intersection density (count per square kilometer) (Table 3). In terms of diversity, the average 8-tier employment entropy is 0.508, albeit with significant variations across all the CBGs in California (note employment entropy is bounded between 0 and 1 with ‘one’ indicating the greatest diversity). The data seem reasonable based on average household size and different income levels (see Table 3). Around 21% of the respondents were younger than 35 years of age, compared to 60% of the respondents being 45 years or older (Table 3).

Table 3: Descriptive Statistics of Key Variables

Category	Variable Name	N	Mean	SD	COV	Min	Max
Outcome: Physical Activity	Walking (0/1)	44253	0.159	0.366	2.300	0	1
	Biking (0/1)	44253	0.026	0.159	6.117	0	1
	Auto-use (0/1)	44253	0.828	0.377	0.455	0	1
	Transit-use (0/1)	44253	0.056	0.230	4.105	0	1
	Duration of walking (minutes)	7034	36.496	34.953	0.958	1	304.727
	Duration of minutes (log-minutes)	7034	3.195	0.973	0.305	0	5.719
Key Walkability Related Variables	Employment mix	44253	0.508	0.274	0.540	0	0.972
	Residential density	44253	5.045	7.007	1.389	0	196.740
	Number of transit stops	44253	2.244	4.153	1.851	0	100
	Vehicle Miles Travelled (VMT)	44253	24925.790	4568.426	0.183	0	37701.720
	Intersection density	44253	81.836	53.738	0.657	0	251.010
Household Size & Tenure	Population	44253	2055.647	1421.440	0.691	0	39017
	Household size	44253	3.165	1.454	0.459	1	8
	Housing in tenure	44206	13.892	10.679	0.769	1	81
Income, Age, & Locality Indicators	Low-income ($0-$50,000 USD)	44253	0.293	0.455	1.554	0	1
	Medium-income ($50,000 - $100,000 USD)	44253	0.322	0.467	1.450	0	1
	High-income (≥$100,000 USD)	44253	0.385	0.487	1.265	0	1
	Age: 18 - 24 years old	44253	0.094	0.292	3.100	0	1
	Age: 25 - 34 years old	44253	0.119	0.324	2.722	0	1
	Age: 35 - 44 years old	44253	0.185	0.388	2.099	0	1
	Age: 45 - 54 years old	44253	0.278	0.448	1.610	0	1
	Age: 55 - 64 years old	44253	0.323	0.468	1.446	0	1
	Adult person (18 - 65 years old)	44253	1	0	0	1	1
	Metropolitan area (1/0)	44253	1	0	0	1	1
Controls	Employed (1/0)	44253	0.724	0.447	0.617	0	1
	College degree (1/0)	44253	0.467	0.499	1.068	0	1
	Rented home (1/0)	44253	0.232	0.422	1.817	0	1
	Male (1/0)	44253	0.479	0.500	1.044	0	1
	No kids (1/0)	44253	0.596	0.491	0.823	0	1
	Kids, single parent (1/0)	44253	0.014	0.119	8.255	0	1
	Kids, multi parent (1/0)	44253	0.389	0.488	1.253	0	1
	Has disability (1/0)	44253	0.060	0.237	3.968	0	1
	Student count in household	44253	0.978	1.157	1.183	0	8
	License count in household	44253	2.197	0.943	0.429	0	8
	Bike count in household*	44213	1.855	1.828	0.986	0	15
	Number of walk trips*	42916	4.795	5.444	1.135	0	50
	Number of bike trips*	30728	0.836	2.252	2.695	0	50
	Number of transit trips*	43737	1.166	0.372	0.319	1	2

Notes: N is sample size; SD is standard deviation; COV is coefficient of variation; The binary physical activity outcomes are based on respondents' self-reported participation in physical activity over the one-week period preceding their survey day; (*) indicates variables that have some missing data; Descriptive statistics for walk duration are shown for adults who participated in transport walking (i.e., N = 7034 individuals).

Finally, the boxplots in Figure 2 provide preliminary insights about the distributions of key CBG-level walkability indicators by participation in transport walking. All five key walkability indicators are not normally distributed with heavy right tails (Figure 2). Overall, respondents who participated in transport walking on-average are residing in CBGs with higher (more walking-supportive environments) median values of objectively measured walkability indicators (see the black solid horizontal lines in Figure 2 indicating median). For example, the CBGs of respondents who did participate in walking tended to be more diverse, dense, and connected as shown by higher median values of employment mix, residential density, and intersection density, respectively (Figure 2).

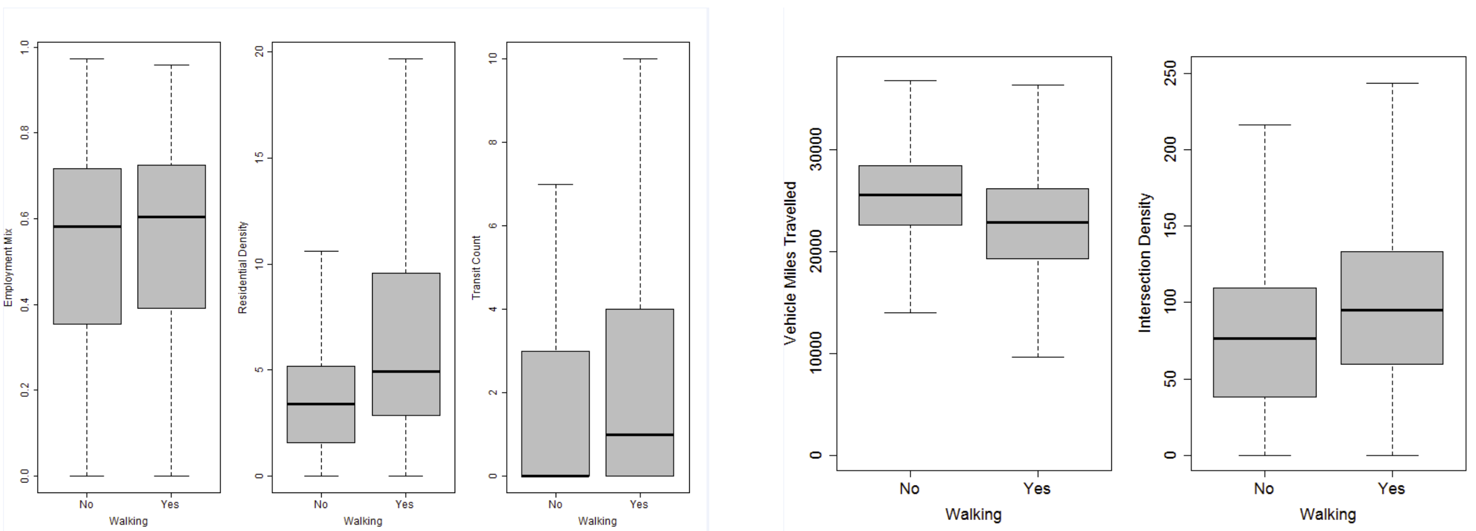


Figure 2: Distribution of Key CBG-Level Walkability Indicators by Participation in Transport Walking
(Note: Outliers for Residential Density and Count of Transit Stops are not shown).

4.2. Modeling Results

Before conducting advanced statistical analysis, simple correlations and binary logit models were estimated to explore the strength and direction of relationships between key built environment features and the likelihood of adults participating in transport walking in metropolitan regions. The analysis was performed for aggregate data as well as stratified data by income level. The modeling results under strategy 1 and 2 are presented below.

4.2.1. *Strategy 1*

GABLMs were estimated under Strategy 1, i.e., the models estimated the dose-response relationships of single objectively measured environmental factors (as discussed in section 3.4) with participation in transport walking, adjusted for sociodemographic, household size and tenure, income, age, locality indicators, and employment/education status. Taking one built-environment feature as an input at a time, a total of five models were estimated given the five built-environment features considered in this study. To examine if the nature of relationships (and underlying thresholds) varies across the socioeconomic status of the participants, separate models were estimated for low-income, medium-income, and high-income adults. In summary, a total of 20 GABLMs were estimated. Referring to Table 4, thin plate regression splines are estimated for the built-environment features. Whereas all other controls such as sociodemographic, household size/tenure, age, locality, and employment/education status (results not shown in the Table 4) are parametric, i.e., standard β's are estimated. Regarding the spline terms, note that the estimated univariate smooth functions such as $f_i(Employment\ Mix)$ cannot be represented using single β value and are are usually represented in terms of degrees of freedom. A statistically significant degree of freedom greater than 1 suggests strong non-linear pattern in the dependency of transport walking on a specific correlate. For details, see (Wali et al. 2018b).

Table 4: Estimation Results of Strategy 1 Models (including single built-environment variable at a time in a model specification)

Variables	Aggregate Data (N = 44253)		Low-income Metro Adults (N = 12963)		Medium-income Metro Adults (N = 14269)		High-income Metro Adults (N = 17021)	
	Effective DF	Chi-square (p-value)	Effective DF	Chi-square (p-value)	Effective DF	Chi-square (p-value)	Effective DF	Chi-square (p-value)
Model 1:								
Employment Mix	2.09	30.80 (0.000)	6.38	14.04 (0.060)	1.83	20.77 (0.000)	1.07	8.44 (0.005)
Log-likelihood at convergence	-18002		-5827.14		-5180.59		-6868.66	
Akaike Information Criterion (AIC)	36044.2		11699.05		10396.85		13771.49	
Model 2:								
Residential Density	5.26	992.55 (0.000)	4.24	196.58 (0.000)	8.66	389.19 (0.000)	3.97	463.36 (0.000)
Log-likelihood at convergence	-17516.8		-5733.37		-4987.18		-6640.00	
Akaike Information Criterion (AIC)	35080.2		11507.23		10023.69		13319.94	
Model 3:								
Number of Transit Stops	5.76	306.25 (0.000)	5.28	75.82 (0.000)	3.57	115.99 (0.000)	3.22	127.90 (0.000)
Log-likelihood at convergence	-17868.7		-5796.63		-5134.96		-6810.68	
Akaike Information Criterion (AIC)	35785		11635.85		10309.08		13659.80	
Model 4:								
Vehicles Miles Travelled (in 1000s)	8.26	1013.1 (0.000)	6.77	232.99 (0.000)	7.52	375.86 (0.000)	7.28	451.62 (0.000)
Log-likelihood at convergence	-17508.3		-5715.72		-5003.47		-6645.21	
Akaike Information Criterion (AIC)	35069.1		11476.98		10053.99		13336.99	
Model 5:								
Intersection Density	6.92	497.3 (0.000)	2.36	85.17 (0.000)	3.54	175.97 (0.000)	7.58	279.75 (0.000)
Log-likelihood at convergence	-17767.8		-5791.00		-5105.76		-6732.98	
Akaike Information Criterion (AIC)	35585.3		11618.74		10250.62		13513.14	

Notes: DF is Degree of Freedom; The model estimated using aggregate data control for income levels. All the models control for age, employment, family type, race, gender, household tenure and renter status, household size, number of students and number of driving license holders in a household (results not shown for brevity); Low-income is 0 – 50,000 USD; Medium income is 50,000 – 100,000 USD; High income is greater than or equal to 100,000 USD.

Several important insights can be obtained from Table 4. First, for all the five built-environment features in models estimated using aggregate data, the effective degrees of freedom (DF) are statistically significantly greater than 1 at a 95% confidence level (see Effective DF for Model 1 through Model 5 in Table 4). This suggests that the dependencies between built-environment features and likelihood of transport walking are highly non-linear. For instance, as opposed to a DF equaling one (which would suggest a fully linear relationship), the effective DF for employment mix is 2.09 whereas effective DF for VMT is as much as 8.26. Next, the nature of non-linear relationship varies significantly across different income groups. For example, considering employment mix, the DF statistically significantly varies between 1.07 (for high-income adults) and 6.38 (for low-income adults) (see estimates in Table 4). This suggests that not only is the relationship between employment mix and likelihood of transport walking non-linear, the extent and shape of non-linearity varies across different income groups. Similar insights can be obtained for other built-environment variables as shown in Table 4.

4.2.2. Strategy 2

Next, as discussed in section 3.4.1, models were estimated under strategy 2 where multiple built environment factors were considered in the same model specification. The key motivation behind this was to examine if considering multiple environmental variables in a same model specification affects the contours of relationships observed in model with single environmental variables at a time (as shown in Table 4). The results of models with multiple built environment variables estimated using aggregate data, as well as stratified data (by income), are shown in Table 5. For the models with multiple built environment covariates, visualizations are developed showing the relationships between the environmental factors and likelihood of walking (discussed later). Likewise, thresholds effects corresponding to different probability ranges of transport walking are determined based on the estimation results. Overall, after accounting for multiple environmental variables in the same model specification, none of the individual environmental variables lose statistical significance (see the effective DF and associated p-values in Table 4 and 5). However, in terms of nonlinearities (and which are directly relevant to determination of thresholds), the extent as well as shape of nonlinearities associated with environmental variables differ between strategy 1 and 2 (see the effective DF estimates for the five built-environment features in Table 4 and 5). This highlights the importance of addressing the covariance issue discussed earlier. For instance, a nonlinear relationship between employment mix and walk probability (effective DF of 2.09) was observed in the aggregate data model when only employment mix and control variables were considered as independent variables (Table 4). However, when multiple built-environment features are considered, the nonlinearity associated with employment mix diminished (effective DF of 1.08 – Table 5). Likewise, for the aggregate data models, the "wiggliness" or "spline complexity" for each of the other four environmental variables significantly reduced under strategy 2 (see the lower DF estimates for environmental variables in Table 5 compared to those in Table 4). Similar patterns can be observed in income-based segmented models. As discussed earlier, one reason behind the change in nonlinear contours between strategy 1 and 2 could be the important issue of omitted variable bias (Wali et al. 2018c), and conceptually the potential synergies among different environmental features. That is, when only one environmental variable is considered at a time (such as employment mix), the 'effects' of other correlated environmental variables (such as transit accessibility) could be manifested through the observed associations between employment mix and walk likelihood – eventually leading to seemingly nonlinear associations for employment mix.

Overall, these findings emphasize that it is important to consider multiple (interrelated) built environment variables in the same model specification to quantify the non-linearities associated with individual variables. Since the key focus of this study is on environmental variables, the non-parametric functional form analysis is mainly conducted for walkability related measures. In addition, most of the control variables in this study are discrete (binary) in nature and thus non-parametric functions cannot be ideally estimated. Furthermore, despite the parametric β coefficients for control variables, the use of categorical variables helps capture the potential nonlinear associations (such as for income, age, and family structure related variables – see Table 5). As part of future work, non-parametric functional form analysis can also be conducted for continuous control variables such as population, student count, etc.

4.2.3. Additional Analysis – polynomial models and partial-segmentation based GAM analysis

To see how simpler and more subjective polynomial-type models perform compared to GAMs, we estimated a polynomial logit regression for the aggregate data using all the built environment and control variables (results can be requested from the authors upon request). To determine the order of polynomial terms for each of the built environment variables, polynomial terms were added to a linear specification until further addition of higher-order polynomial terms did not reduce information criteria measures (AIC and BIC) (Bozdogan 2000, Wagenmakers and Farrell 2004). Once the orders of polynomial terms for built environment variables were determined, a polynomial model was estimated using all the built environment (with polynomial terms) and control variables. Order three (3) polynomial terms were found as best fit for employment mix and intersection density whereas order five (5) polynomial terms provided the best fit for residential density, number of transit stops, and vehicles miles travelled. The comparison suggested better performance of more parsimonious and objectively driven GAMs compared to less parsimonious polynomial model. The AIC for the polynomial model (for aggregate data) was 34872.44, which is greater than (indicating poorer fit) the AIC for the best-fit GAM model (AIC of GAM model 34867.13). A difference of 5 between two AICs for two competing models provides strong support in favor of the model with the lowest AIC (GAM model in this case). Compared to the DF for a specific variable quantified by GAM (Table 5), greater number of polynomial terms (indicating overestimation/overfitting) were observed for employment mix, number of transit stops, and to some extent residential density, whereas the best-fit polynomial terms were smaller for vehicle miles traveled and intersection density than their DF counterparts in GAM (indicating underestimation/underfit). In line with the conceptual motivation discussed earlier (section 3.1), these findings collectively provide strong evidence in favor of GAM approach used in this study.

To examine the feasibility and performance of partial segmentation-based GAM analysis (as discussed in footnote 4), we also conducted analysis based on interacting the income indicator with the non-parametric functions for BE features (partial segmentation) in a single model. In particular, the smooth functions for each of the 5 BE features are interacted with income indicator – providing income-specific spline effects in one model. The results obtained from partial segmentation approach are provided in Table A1 and Figure A1 in Appendix A. Compared to the GAM based on aggregate data (no income specific spline terms – see Table 5), the AIC of GAM based on partial segmentation (income-BE splines interactions) presented in Table A1 is lower by a magnitude of 138 (indicating substantially better fit for the model based on partial segmentation). However, the combined AIC of three separate income-based GAMs (full segmentation approach in Table 5) is even further lower than the AIC of GAM based on partial segmentation by a magnitude of 195 (see the comparison summary statistics in Table A1).

This provides substantial statistical evidence in favor of full segmentation approach used in the study. As a rule of thumb, a difference of greater than 10 in the AICs of two competing models provides essentially no support in favor of the model with greater AIC (in this case GAM model with partial income-BE based segmentation). Also, between the two approaches (full segmentation vs. partial segmentation), the magnitudes of Effective DF terms (capturing non-linearities) do not differ significantly (except for VMT spline effects for medium-income group) (see the Effective DF for partial segmentation model in Table A1 and those in Table 5 for full segmentation approach). Given the similarities in the spline terms between the two approaches, the marked improvement in model goodness of fit for full segmentation approach can be attributed to the fact that the full segmentation approach accounts for the potential differences in the coefficients for control variables across the three income groups, whereas, the partial segmentation approach restrict them across the three income groups (see the significant differences in β estimates for control variables across the three income groups – Table 5). This also provides substantial support in favor of full segmentation model. Admittedly, the conclusion that full segmentation approach is warranted in this study may be data dependent and a more parsimonious partial segmentation approach may be better in other contexts.

Note that the McFadden Pseudo R^2 values in this study varies between 0.08 and 0.15 for the aggregate and income-based stratified GAMs. Travel behavior models often have low Pseudo R^2 values (below 0.2) given the intrinsic and substantial variations in individuals' travel patterns (Cervero 2002, Guo and He 2020). Unlike, traditional R^2 for linear models, the values for McFadden Pseudo R^2 for discrete outcome models are often low with values below 0.2 not uncommon. Several other discrete outcome studies in the field of transport achieved similar Pseudo R^2 values (Hensher and Sullivan 2003, Buehler 2012, Aziz et al. 2015, Li et al. 2020).

Table 5: Estimation Results of Strategy 2 Models (including multiple built-environment variables in a same model specification)

Variables	Aggregate Data (N = 44,253)		Low-income Metro Adults (N = 12,963)		Medium-income Metro Adults (N = 14,269)		High-income Metro Adults (N = 17,021)	
	Effective DF	**Chi-square (p-value)**	**Effective DF**	**Chi-square (p-value)**	**Effective DF**	**Chi-square (p-value)**	**Effective DF**	**Chi-square (p-value)**
Walkability-Related Spline Effects								
Employment Mix	1.08	10.71 (0.001)	1.01	1.44 (0.229)	1.79	10.68 (0.000)	1.02	3.21 (0.073)
Residential Density	4.61	79.81 (0.000)	1.02	16.34 (0.000)	6.67	59.08 (0.000)	4.35	35.17 (0.000)
Number of Transit Stops	2.66	93.60 (0.000)	1.74	17.83 (0.000)	5.38	53.28 (0.000)	3.31	40.06 (0.000)
Vehicles Miles Travelled (in 1000s)	7.7	28.88 (0.000)	3.02	29.63 (0.000)	6.37	8.49 (0.343)	6.8	14.79 (0.060)
Intersection Density	4.33	45.22 (0.000)	2.52	10.20 (0.019)	3.58	18.40 (0.001)	6.68	34.18 (0.000)
Controls	**β**	***z-score***	**Β**	***z-score***	**β**	***z-score***	**β**	***z-score***
Intercept	-0.52	-5.53	-0.21	-1.56	-0.89	-4.58	-1.27	-5.63
Medium-income (50,000 - 100,000 USD)	-0.17	-4.25	---	---	---	---	---	---
High-income (≥ 100,000 USD)	0.08	1.85	---	---	---	---	---	---
Age: 25 - 34 years old	-0.37	-5.97	-0.23	-2.43	-0.51	-4.23	-0.42	-3.46
Age: 35 - 44 years old	-0.46	-7.82	-0.27	-3.05	-0.57	-4.85	-0.52	-4.58
Age: 45 - 54 years old	-0.43	-7.62	-0.34	-3.90	-0.61	-5.52	-0.45	-4.26
Age: 55 - 64 years old	-0.48	-8.51	-0.46	-5.14	-0.44	-4.12	-0.58	-5.46
Population (in 1000s)	-0.12	-8.68	-0.09	-3.78	-0.14	-5.23	-0.11	-5.42
Rented home (1/0)	0.30	8.30	0.37	6.79	0.26	3.86	0.03	0.36
Employed (1/0)	-0.34	-10.59	-0.47	-9.31	-0.23	-3.69	-0.23	-3.89
College degree (1/0)	0.43	13.09	0.20	3.20	0.51	8.96	0.64	10.54
Kids, multi parent (1/0)	-0.25	-5.20	-0.57	-6.87	-0.03	-0.36	0.01	0.09
Kids, single parent (1/0)	-0.28	-2.59	-0.58	-3.96	-0.18	-0.79	0.02	0.06
Household size	0.18	9.20	0.21	7.55	0.10	2.12	0.02	0.46
No disability (1/0)	0.16	2.81	0.29	3.86	0.04	0.31	0.34	1.98
Student count in household	-0.05	-2.43	-0.02	-0.57	-0.06	-1.39	-0.02	-0.44
License count in household	-0.56	-25.36	-0.77	-24.11	-0.35	-6.89	-0.24	-4.99
Male (1/0)	-0.13	-4.50	-0.29	-6.02	-0.08	-1.46	-0.01	-0.32
Summary Statistics								
Log-likelihood at constant, LL_{NULL}	-19,379		-6,605		-5,491		-7,123	
Log-likelihood at convergence, $LL_{CONVERGENCE}$	-17,395.23		-5,698.46		-4,933.14		-6,581.75	
AIC	34,867.13		11,447.50		9,945.88		13,239.82	
McFadden Pseudo R-Squared	0.11		0.15		0.10		0.08	
UBRE Score	-0.21		-0.12		-0.30		-0.22	

Notes: UBRE is Un-biased Risk Estimator, smaller values indicating better fitting models; DF is Degree of Freedom; USD is United States Dollar; AIC is Akaike Information Criterion. The McFadden Pseudo R-squared is computed as $\left[1 - {LL_{CONVERGENCE}}/{LL_{NULL}}\right]$ – and the ratio of the two likelihoods suggests the level of improvement over the intercept model offered by the full model.

5. DISCUSSION

Given that the models estimated under Strategy 2 can better account for the interrelated effects of multiple environmental variables, we base our discussion of the key findings in this section on the results shown in Table 5. As discussed in section 3.1, based on the GABLM models, we develop visualizations showing the relationship between the raw values of Z (environmental factors) plotted on x-axis and the probability of walking on y-axis (obtained from smoothed values in $f(Z)$) (Figures 3 and 4). Since the probability scale is based on inverse logit transform (footnote # 5), less than 50% indicates a negative coefficient in a logit model (decreasing the odds of walking, ceteris paribus) while greater than 50% indicates a positive coefficient (increasing the odds of walking, ceteris paribus). Figure 3 shows the contours of relationships for employment mix, residential density and count of transit stops, for aggregate data, low-income, medium-income, and high-income adults in metropolitan regions (Figure 3). Likewise, Figure 4 shows the contours of relationships for VMT and intersection density for aggregate data, low-income, medium-income, and high-income adults in metropolitan regions (Figure 4). Before discussing the findings, we note that the figures showing the non-linear contours for the built environment features and the overall interpretation of the findings are based on the common assumption of "ceteris paribus" or "all else being equal." That is, other variables are held constant when calculating the change in predicted probabilities because of change in a specific built environment feature. This assumption is not unique to this study and is at the core of interpreting inferential findings from statistical methods. However, we re-emphasize that real-world built environment measures tend to covary together - and while crucial to interpreting findings statistically, it is difficult to hold other variables constant in real-world situations.

5.1. Employment Mix, Residential Density, & Count of Transit Stops

The relationship between employment mix and likelihood of participation in transport walking follows a monotonically increasing (linear or nonlinear) pattern. That is, up until an employment mix of around 0.51 (see the grey contour in Figure 3 for employment mix), the positive correlations between transport walking and employment mix increases, but with the probability (or the chance of) of transport walking remaining below 0.5 (or 50%) (Figure 3). However, any increase in employment mix beyond 0.51 is associated with a greater than 50% likelihood of transport walking, with the maximum probability of transport walking equaling 0.57 associated with an employment mix near 1 (Figure 3). Thus, for employment mix in the model estimated with aggregate data (Table 4), a threshold of '> 0.51' can be defined, suggesting that an increase in employment mix beyond 0.51 will be associated with a greater than 50% likelihood of transport walking participation. While not directly comparable to the employment entropy measure used in the present study, Park et al. (2020) recommended a minimum of 0.2-0.5 job-population balance for achieving polycentricity (Park et al. 2020). Table 6 summarizes the thresholds deduced from the graphical contours. Likewise, for medium-income and high-income adults in metropolitan regions, thresholds of 0.60 and 0.59 can be deduced from Figure 3 (yellow and green contours in for employment mix for medium and high-income adults) and summarized in Table 6. While similar dependence structure is observed for low-income, the non-linear associations (and the corresponding threshold of 0.59, see Table 6) are statistically insignificant at any reasonable level of confidence. As can be seen, such thresholds beyond which an increase in employment mix could be associated with a more than 50% chance of transport walking cannot be obtained without an adequate characterization of the of complex relationships embedded in data.

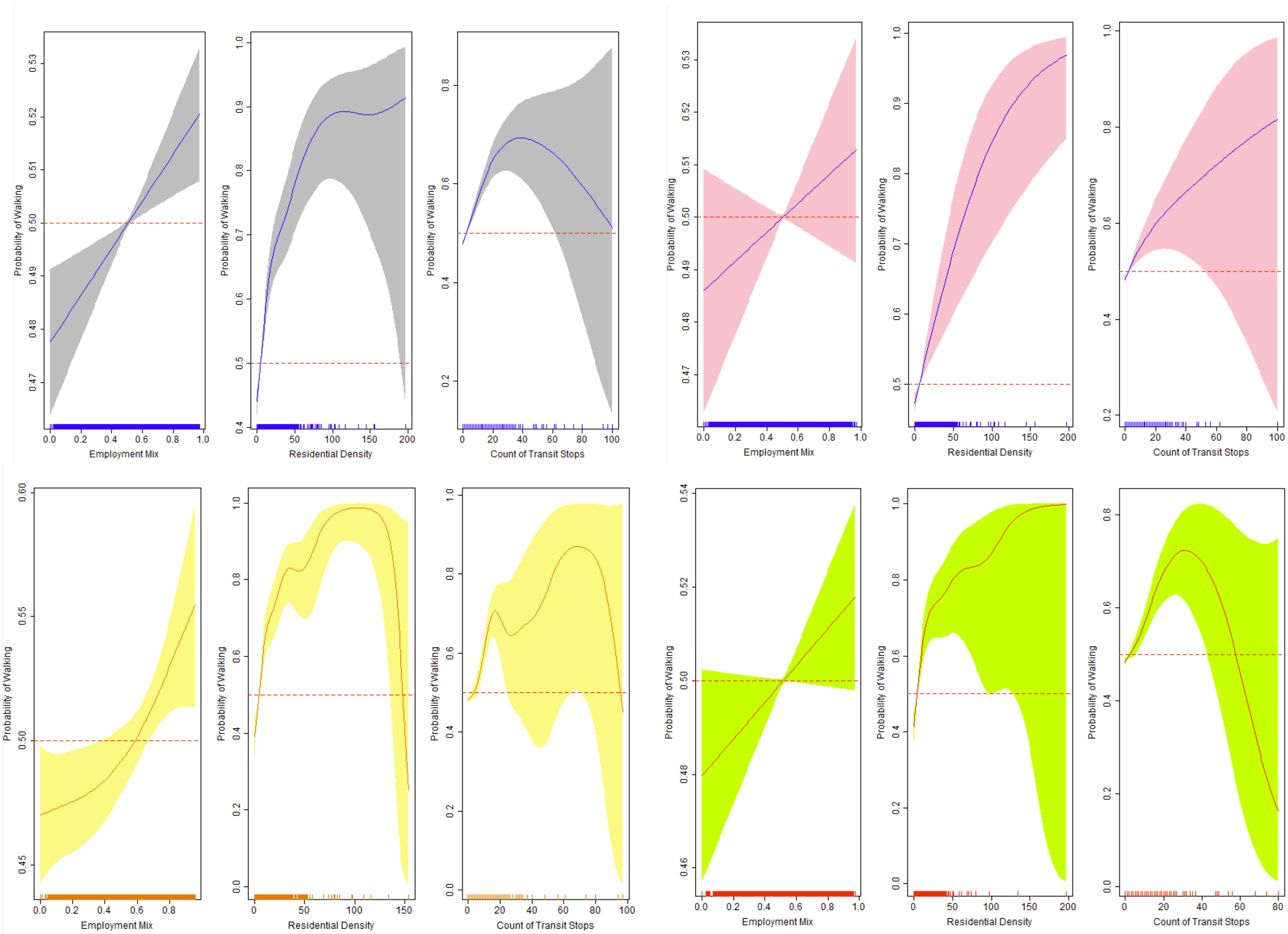


Figure 3: Dose-Response Relationships Between Objectively Measured Employment Mix, Residential Density, and Count of Transit Stops and Participation in Transport Walking for Aggregate Data (grey), Low-Income Adults (pink), Medium-Income Adults (yellow), and High-Income Adults (green)

Notes: Raw values of a built environment predictor are plotted on X-axis. Probability of transport walking obtained from Generalized Additive Model plotted on Y-axis. The horizontal red dotted line represents in the context under discussion a 50-50 chance of an adult in metropolitan regions participating in transport walking. The little vertical bars above the X-axis (the “rug”) represent the sample across the continuous range of an explanatory factor). The upper and lower lines in the shaded portions represent 2 standard errors above and below the estimate of the smooth being plotted.

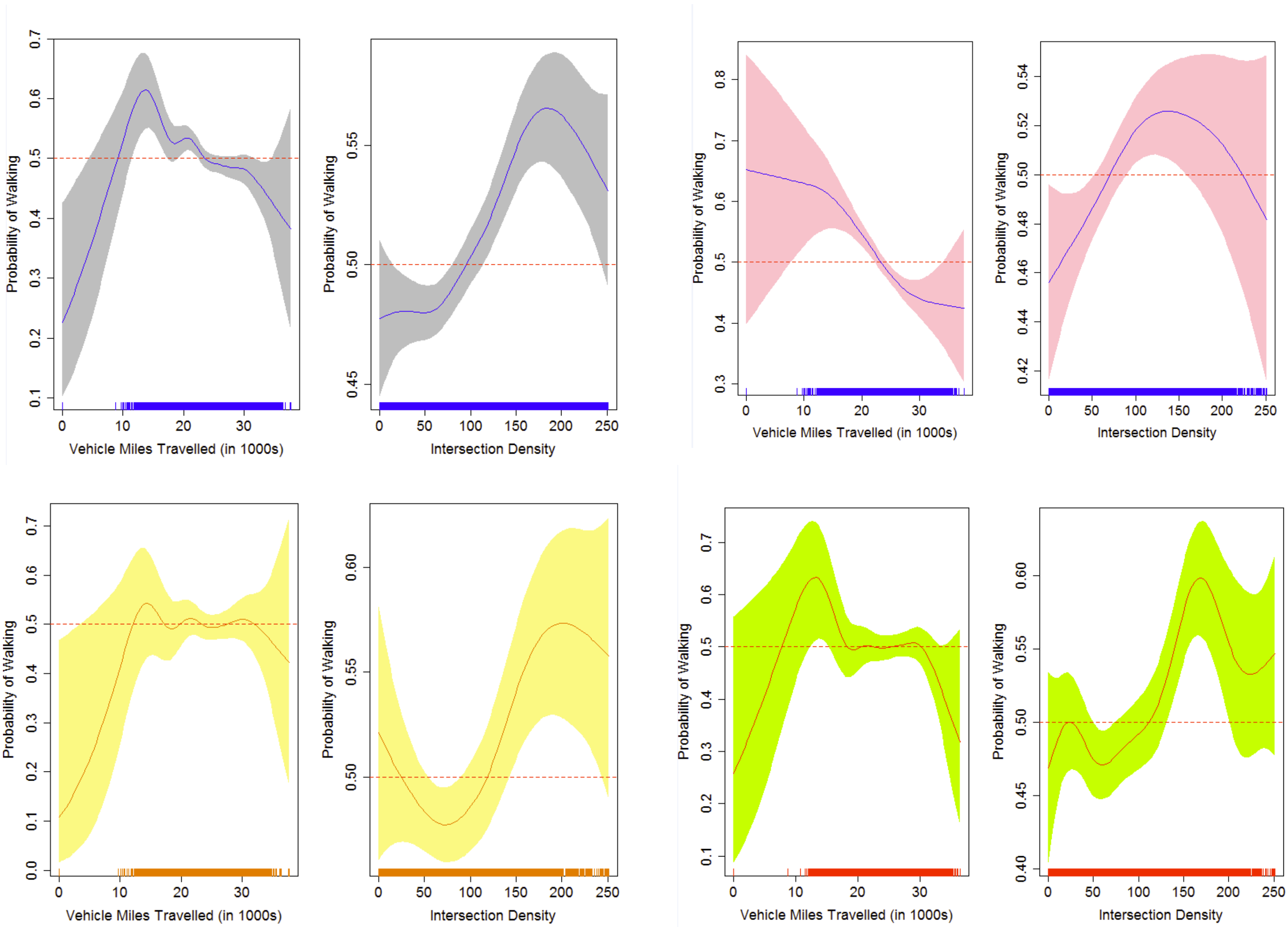


Figure 4: Dose-Response Relationships Between Objectively Measured Vehicle Miles Travelled (in1000s), Intersection Density, and Participation in Transport Walking for Aggregate Data (grey), Low-Income Adults (pink), Medium-Income Adults (yellow), and High-Income Adults (green)

Notes: Raw values of a built environment predictor are plotted on X-axis. Probability of transport walking obtained from Generalized Additive Model plotted on Y-axis. The horizontal red dotted line represents in the context under discussion a 50-50 chance of an adult in metropolitan regions participating in transport walking. The little vertical bars above the X-axis (the "rug") represent the sample across the continuous range of an explanatory factor). The upper and lower lines in the shaded portions represent 2 standard errors above and below the estimate of the smooth being plotted.

Table 6: Thresholds for Key Built-Environment Features

Key Environmental Variables	Chance	Metro All Adults	Metro Low-Income Adults	Metro Medium-Income Adults	Metro High-Income Adults
Employment Mix (Evenness of distribution of jobs across sector)	< 50%	<= 0.51	<= 0.59[a]	<= 0.6	<= 0.59*
	50 - 57%	> 0.51	> 0.59[a]	> 0.6	> 0.59*
Residential Density (Households /developable acre of land)	< 50%	[0 - 10]	[0 - 10]	[0 - 10]	[0 - 10]
	50 - 70%	[10.1 - 31]	[10.1 - 57]	[10.1 - 20]	[10.1 - 25]
	71 - 81%	[31.1 - 69]	[57.1 - 91]	[20.1 - 33]	[25.1 - 77]
	82 - 89%	[69.1 - 101]	[91.1 - 127]	[33.1 - 70]	[77.1 - 110]
	90 - 97%	---	[127.1 - 200]	[70.1 - 83]	[110.1 - 132]
	↪	[101.1 - 200]	---	[83.1 - 127]	[132.1 - 200][a]
	⇓	---	---	[127.1 - 150]	---
Access to Transit (# of transit bus stops/rail stations)	< 50%	[0 - 4]	[0 - 4]	[0 - 5]	[0 - 3]
	51 - 60%	[4.1 - 16]	[4.1 - 26]	[5.1 - 11]	[3.1 - 13]
	61 - 70%	[16.1 - 39]	[26.1 - 50]	[11.1 - 20]	[13.1 - 32]
Intersection Density (Intersections per square kilometer)	< 50%	<100	< 75	< 122	< 122
	51 - 60%	[100 - 200]	[75 - 150]	[122 - 201]	[122 - 179]
	⇓	> 200	> 150[a]	> 201	> 179*

Notes: Thresholds are deduced from the relationship contour plots in Figure 3 and 4; The thresholds/optimal ranges are those which result in an at least 50% chance of transport walking; All the threshold are statistically significant at 95% confidence level except those marked by (a) indicating statistically insignificant thresholds at any level of confidence. The ones marked by (*) indicate statistically significant at 90% confidence level; (---) indicates Not Applicable; (↪) indicates ceiling effect - i.e., increase in an environmental variable does not result in increase in probability of walking; (⇓) indicates overdose effects - i.e., an increase in residential density for a certain range in fact decreases the probability of walking.

As can be seen in Figure 3, the complex relationships between residential density, count of transit stops, and participation in transport walking exhibit a continuously increasing non-linear association up until a certain point but also exhibit ceiling effects and potential overdose effects.

Regarding the non-linear associations for residential density in the aggregate model (see grey contour for residential density in Figure 3), the correlations between residential density and transport walking increases (albeit non-linearly) across the entire range of this built environment feature. Compared to an effective DF of 1 (which would imply a linear relationship), the DF for residential density is 4.61 suggesting strong nonlinear patterns (Table 5). A sharp increase, in a piece-wise non-linear manner, is observed in the associations for residential density up until around 101 housing units per acre; a further increase beyond which is not associated with a significant increase in the probability of transport walking, indicating potential ceiling effects. Compared to a linear model, the estimated strong nonlinear dependencies are important as they lead to different changes in probability of transport walking along the range of residential density. To further elaborate on this, consider how increase in residential density up until 101 housing units per acre is correlated with the probability of walking. Table 6 shows the ranges of residential density corresponding to specific probability ranges for metro adults participating in transport walking. For instance, for the model based on aggregate data, any increase in residential density up until 10 housing units per acre is associated with a less than 50% chance of transport walking (Table 6). Any increase beyond 10 and up until 31 housing units is associated with a probability of transport walking ranging between 50% and 70% (see Table 6 and Figure 3). A further increase beyond 31 is associated with a subsequent increase in the chance of transport walking beyond 70%, reaching a maximum chance of 89% with a residential density of 101 housing units per acre. This is intuitive as increase in residential density could reflect the greater compactness of an urban area based on the amount of people living in an area and the type of housing available (Frank et al. 2003). Compared to less compact developments, people are likely to walk more in compact environmental configurations (Rodríguez et al. 2009).

However, a potential ceiling effect (statistically significant) can also be observed in the range of 101 to 200 housing units per acre (see Table 6), suggesting that an increase in residential density beyond 101 housing units per acre is likely associated with no increase in probability of walking or a relatively modest increase (note in Figure 3 the relatively flat region in the contour for this variable beyond 101 followed by a modestly increasing slope)[6]. Thus, in terms of an optimal threshold value, a conservative choice would be a residential density of 10 housing units per acre (as probability of transport walking increases beyond 50% with a greater than 10 housing units per acre in a certain CBG). However, more customized thresholds can be selected from Table 6 should a specific CBG in California fall in that range. For instance, if a CBG already has a residential density in the range of 10.1 – 31 housing units per acre (and which is associated with 50 to 70% chance of transport walking), a more context-specific target for this CBG could be a residential density of 31 housing units per acre, as further residential compactness beyond this point would be associated with an even higher increase in the probability of transport walking beyond 70% (see Table 6). As is evident, GABLMs help obtain a microscopic understanding of how specific ranges of built environment fabric correlate with the chance of transport walking.

When different GABLMs are estimated using stratified data based on income, the nature of relationship between residential density and probability of transport walking changes significantly, suggesting that there are significant context-specific differences among the adults in different income-groups (see the income group specific contours for residential density in

[6] Such differential dependencies (especially ceiling/overdose effects) of transport walking on residential density discussed above cannot be obtained from a linear model. Overall, higher residential densities could reflect other benefits associated with walking such as (energy) efficient use of road infrastructure, better housing affordability and likely vibrant street life. But when increased too much beyond a certain point, higher residential densities (like other environmental features) alone may not be the key determinants of increased physical activity (Forsyth et al. 2007, Oakes et al. 2007).

Figure 3). Subsequently, the thresholds for each of the income groups differ as well (see Table 6). Similar patterns are observed for the effect of count of transit stops as well. This clearly demonstrates that selection of thresholds should consider the socioeconomic factors and differences, and that a single threshold may not be generalizable across low, medium, and high-income adults in metropolitan regions. Also, as is evident, ignoring potential strong non-linear dependence structures can clearly hide such important information embedded in the data.

Coming to the non-linear associations of access to transit with transport walking participation, the likelihood of transport walking is greater than 50% for four transit stops, with maximum gains in the chance of transport walking participation observed in the range of 4-16 transit stops beyond which a ceiling effect is observed (Figure 3). While the overall pattern for different income groups remain similar (expect for medium income adults), the quantitative thresholds differ across the income groups (see Table 6). Again, while not directly comparable to the transit accessibility measure used in this study in addition to the differences in specification/formulation of statistical models, Park et al. (2020) found a sharp increase in the likelihood of walking beyond 125 transit stops per square mile with a ceiling effect observed around 175 transit stops per square mile (Park et al. 2020).

5.2. Intersection Density, Vehicle Miles Travelled, & Other Controls

The dependence of transport walking participation on intersection density exhibits a pattern where the associations for intersection density (and subsequently the likelihood of transport walking beyond 50%) steeply increases within certain range of intersection density. For example, within the range of 100 to 200 $intersections/km^2$, a steep positive slope can be observed compared to the range of 0-100 $intersections/km^2$ and > 200 $intersections/km^2$ where the slope is either increasing (albeit with likelihood of walking less than 50%) or decreasing (indicating overdose returns of connectivity), respectively (see Table 6 and the grey contour for intersection density in Figure 4). Again, a linear model would imply a constant association between intersection density and likelihood of transport walking. That is, the probability of transport walking increases by X% for the entire range of intersection density at times when the relationship between transport walking and intersection exhibits strong nonlinear patterns. A linear model would also preclude quantification of ceiling/overdose effects associated with increasing street connectivity (beyond a certain point) as identified by the GAM model (see Figure 5) – i.e., increase in intersection density beyond 200 $intersections/km^2$ is associated with a decrease in probability of transport walking. A linear model would ignore such overdose effects and suggest a constant increase in probability of transport walking beyond 200 $intersections/km^2$. The study by Park et al. (2020) observed a sharp increase in walk participation in the range of 75 – 160 intersections per square mile but with the likelihood of transport walk participation staying below 50%. The peak was observed around 300 intersections per square mile (Park et al. 2020). One possible explanation for the overdose effects of intersection density observed in this study could be the potential safety hazards around intersections. Increase in intersection density reflects better connectivity. However, an increase in intersection density beyond a certain point can also pose safety challenges for vulnerable road users since intersections are planned points of conflict in any roadway system and are the most dangerous locations for vulnerable road users (Carter et al. 2006, Wali et al. 2018a). Thus, the potential gains associated with greater (and pedestrian and bicyclist unfriendly) connectivity could be offset by the lower level of pedestrian and bicyclist safety at intersections. We also note that adding intersections seems unlikely to change anything unless accessibility is increased as a result. Accessibility is arguably better measured through network analysis, e.g., see (Sevtsuk et al. 2016). Through network analysis, Sevtsuk et al. (2016) demonstrated that the relationship between block size and walkability is non-linear. Through simulating 1000 randomly generated grids, the study confirmed the general hypothesis in the previous literature that smaller blocks (higher connectivity) typically achieve higher pedestrian accessibilities – pedestrian accessibility increases sharply with very small blocks and decreases slowly as block size increases. However, the exponential nature of the trend also highlighted that reducing block size may not improve accessibility in areas with large blocks with few plots. Admittedly, while not a network analysis, the present study also shows a strong non-linear association between intersection density (a surrogate for block size) and physical activity. In fact, as discussed, an 'overdose effect' of increasing intersection density (roughly reducing block size) is observed. As part of future work, using connectivity and accessibility measures obtained from network analysis can provide even deeper insights into the complex mechanisms through which physical activity and connectivity correlate.

Similar trends can be observed for low, medium, and high-income adults. However, the cut-points defining the thresholds vary significantly across the three income groups. For instance, based on the contours in Figure 4, the optimal ranges of intersection density for low, medium, and high-income are 75-150, 122-201, and 122-179, respectively (see Table 6). These findings again demonstrate that built-environment specific thresholds cannot be generalized across income groups, and in fact, vary significantly across adults of different income groups. The dependence of transport walking participation on vehicle miles traveled, and other controls are not discussed for brevity. As stated earlier, VMT is a key measure of car dependence and helps capture to some extent the key exposure pathway (such as air pollution) between built environment and active travel (Frank et al. 2019a). As a result, the observed CBG-level car dependence can be manifested through the VMT variable. However, we note that the other key

BE variables included will only reflect these relationships to the extent that they are not better captured by VMT. This does slightly limit interpretation of the model and may explain the complex dependence structures between VMT and active travel (Figure 4). Given the complexity in the relationship, VMT is not included in the summary thresholds table for key built-environment features (Table 6).

We reemphasize that most of the spline terms in aggregate as well as income-stratified GAMs are statistically significant (see Table 5). This means that “overall” the nonlinear associations are statistically significant. However, for certain ranges of the variables, the error bars encompass zero. That is, in a certain range of a variable, the non-linear contour is statistically insignificant. This is not surprising. Conceptually, we do not expect to have a non-linear contour statistically significant across the “entire range” of a specific variable. In some cases, the statistically insignificant portions of the non-linear contour lines could be reflecting limited sample size (e.g., see the contour for count of transit stops for low-income adults where the sample beyond 40 transit stops is limited – noticing the sparsity in the rug on the x-axis). In other cases, the statistically insignificant portions of the contour line reveal that the non-linear effect is *“statistically insignificant”* and not due to limited sample size (e.g., see the contour line for employment mix for low-income adults). Nonetheless, the thresholds shown in Table 6 are based on the ranges of variables where the non-linear effect is statistically significant. Around 55 thresholds are shown in Table 6 for different income groups and for different probability ranges. Only four are statistically insignificant at any level of confidence.

From the results presented above, it is clear that the thresholds vary among population groups based on income. From an application standpoint, we always deal with a mix of population groups instead of a single group. To deal with a mix of population groups, the thresholds may be selected based on average income of individuals living in a CBG or CBGs in a specific vicinity. For example, if the average income of individuals in a CBG is USD 50K – 100K (medium income group), then thresholds/optimal values for built environment (BE) features presented in this study for medium income group can be used. Admittedly, while this approach is not ideal, it can still be useful since the BE data considered in this study capture variations at the census block group level and income differences among individuals living in a CBG (the smallest geographical unit) are likely to be relatively lower.

Finally, given the cross-sectional data, the BE benchmarks quantified in this study are correlational in nature and thus must not be interpreted in a causal manner. Nonetheless, it can still help planners and engineers to make informed decisions about growth patterns since currently decisions are made on no such evidence (see section 7 about the usefulness of longitudinal evidence and subsequent challenges in the context of threshold development). The overall patterns of associations observed in this study (on-average positive associations between the environmental measures and walking) are generally in line with the travel behavior and planning research (including causal evidence). Also, we note that thresholds or "non-linearities" can be context specific (in our case specific to California). Given this context dependency, the benchmarks should be interpreted with caution when transferring the results to areas outside CA. The weather in California is largely balmy and thus caution must be exercised in the interpretation of the results (especially in areas outside California where the weather is generally more extreme). Perhaps a relevant landmark case was in 1980, Pushkarev and Zupan published *Urban Rail in America* conveying residential density thresholds as low as 10 or fewer households per gross acre to justify rail transit investment (Zupan et al. 1980). This finding was based on an assessment of the New York Metropolitan region where one end of the commute was Manhattan rendering huge impedance for driving due to congestion and rendering parking prohibitive. This was not transferrable to other regions where the relative utility of driving was higher. Methodologically, results from the current study suggest that differences in the ability to establish and apply thresholds is a function of the differences in methodological approaches applied.

It is obvious that the results in this study are based on data on metropolitan areas in California (CA) and as such we clearly acknowledge the potentially limited transferability (out of great caution) of the results to metro areas outside CA. The limited transferability could be since demographics are shifting all the time and differ across geography as well as due to differences in culture and attitudinal predispositions. However, the influence of changes in these factors on travel behavior is applicable to the broader planning and travel behavior research and not merely to the goal of threshold development. Also, the environmental features used in this study exhibit remarkable variability. In addition, common patterns are observed between the characteristics of the reference area (CA) and national trends in terms of general demographic factors, education, and living arrangements (see below). The great variability in the data used in this study, when combined with the fact that California is one of the most (or perhaps the most) diverse states in the nation, makes the results more transferable to other locations compared to an analysis based on data with lesser variability and from less diverse regions. One way of addressing the transferability concern could be to use a nationally representative georeferenced travel survey data linked with objective measures of the environment. The authors did not have access to such geocoded data. Until access to such georeferenced data is publicly available, the choice is between imperfect information vs no information. Also, while it may be seemingly appealing to develop thresholds using national level data, the threshold development in the national context is even further complicated by complex and important methodological concerns of spatial and unobserved heterogeneity (in addition to systematic heterogeneity). To this end, the policy recommendation is to apply the income-stratified GAM methodology proposed in this study to other regions where needed.

Related to general factors characterizing the reference area (California - CA) (https://www.census.gov/quickfacts/fact/table/US,CA/PST045219), about 22.7% of the population is 18 years and under (22.4% for the entire U.S.) and 14.3% of the population is 65 years or older (16% for the entire nation). Around 31% of the population across the U.S. have a bachelor's degree or higher (compared to 33.3% in CA). In terms of race, around 72.1% of the population in CA is white alone compared to national average of 76.5%. As expected, the percent of foreign-born persons is significantly higher in CA (26.9% for CA vs. 13.5% for the entire U.S.). Related to living arrangements, the average number of persons in a household is almost similar (2.96 persons per household in CA vs. 2.63 for the entire nation). For transportation, the mean travel time to work (in minutes) (workers age 16 years +) for the entire nation is 26.6 minutes compared to 29.3 minutes for CA. However, in terms of population geography, CA is significantly denser (population per square mile of 87.4 for the entire nation vs. 239.1 for CA). While California is unique in several respects, the statistics presented above may allow readers to establish broader similarity patterns with the entire nation based on these features.

6. CONCLUSIONS

This paper described the development and application of a rigorous and evidence-based methodology for developing place-based thresholds for the following selected set of objectively measured environmental features which support physical activity: Employment Mix (diversity), Residential Density (density), Access to transit stops (transit), and Intersection Density (connectivity). For the development of place-based thresholds, the study emphasized on a key element of systematic heterogeneity, i.e., systematic heterogeneity arising from nonlinear responsiveness of individuals' physical activity to environmental factors. To achieve the objective, an innovative machine-learning based Generalized Additive Modeling (GAM) framework was employed with P-smoothing/thin plate regression splines to empirically identify the potential thresholds from the data at hand. To employ the methodology, large-scale survey data from the 2013 California Household Travel Survey are used to model the likelihood of transport walking of adults in metropolitan areas. Furthermore, the travel behavior data were linked with the comprehensive objectively measured environmental data in the Robert Wood Johnson Foundation's National Environmental Database. To model the individual-level likelihood of transport walking as a function of environmental factors (at the census block group level) and other controls, Generalized Additive Binary Logit Models (GABLM) are developed that capture complex contours of non-linear dependencies in a statistically intuitive way. To account for likely differences arising due to socioeconomic status of the participants, the GABLMs are estimated using the aggregate data (i.e., adult participants in metro regions) as well as stratified data by income-level.

The study offers valuable methodologically derived evidence-based guidance for planners and engineers concerning investments in built environment to support active transportation. The objectively measured key built environment (BE) measures considered in this study can help create polycentric urban developments associated with greater physical activity. The findings of the present study offer an in-depth understanding of the systematically varying influence of the urban environment on walking. The study showed how the complex variations can be systematically harnessed to determine benchmarks for environmental factors as it relates to active travel. Derived from a systematic analysis, the thresholds for achieving a greater than 50% chance of transport walking are: (1) 8-tier employment mix value of greater than 0.51 (ranging between 0 and 1), (2) Net residential density of greater than 10 housing units per residential acre, (3) Greater than 4 transit bus stops/rail stations, and (4) intersection density in the range of 100 to 200 intersections per square kilometer. Furthermore, important dynamic effects (such as leveling-off, ceiling, and trigger effects) of environmental benchmarks on the likelihood of transport walking were found. While the benchmark values of key BE variables are quantified separately (to facilitate ease of application), the modeling approach accounts for the potential interrelationships/covariances among the BE variables framework (Crane 2000, Frank and Engelke 2001). The results also reveal that a single threshold is not generalizable across low, medium, and high-income adults, underscoring the importance of accounting for group specific socioeconomic factors in development of thresholds for environmental benchmarks. Overall, the new findings allow for an investigation of census block group level current built environment conditions and their facilitation of walking in the US. The benchmarks for environmental features developed in this study can serve as a useful guiding tool (and admittedly not as hard truth) for engineers, planners, and other relevant stakeholders to track health conditions of communities over time. Findings from the current study suggest that it is possible to identify dose-response relationships between built environment features and walking. Moreover, due to the inherent non-linear relationship between built environment features and walking, policy relevant thresholds do exist and there are methods available to examine the extent of transferability from one location to another (such as through Geographically Weighted Regressions or Bayesian spatial variants).

7. DIRECTIONS FOR FUTURE RESEARCH

Based on the findings from this study, there are several fruitful avenues for future research related to the development of benchmarks for environmental features as it relates to active travel. These include incorporating more objective measures of the built environment, enhancing data resolution, implementation of the methodology in other regions, and (ideally) conducting causal evaluations for development of environmental benchmarks.

Regarding environmental measures, more objectively measured BE variables could be considered in the future to paint an even more holistic picture of the systematically varying influence of urban form on physical activity. The analysis presented in this paper focused on determining thresholds for individual built environment components of a preexisting national walkability index created as part of work previously completed for Robert Wood Johnson Foundation. As such, the scope of this work was not to choose the components of walkability index created earlier but to study components of an index that already exists. Of the five components of the walkability index (intersection density, employment mix, residential density, transit count, VMT), higher values for VMT (more auto dependence) are indicator of less walkability (see section 3.3 for the rationale behind inclusion of VMT in National Walkability Index). In contrast, higher values for the other four variables have been found to increase walkability. Due to this reason, the VMT is the only variable in the walkability index that is subtracted (with -1 multiplied to VMT in calculating walkability index). In other words, VMT (auto dependence) to some extent could be a reverse proxy for the target variable (walking). As part of future work, the walkability index could be further enhanced, and additional objective measures of auto dependence could be considered and analyzed. Future studies could explore the use of more robust measures of transit accessibility and include objective measures of the natural environment.

The objective built environment data considered in this study are collected at the block-group level. CBG level environmental data are used because anonymity protections permit household location reporting in the CHTS (or any other national-level survey) only at the block-group level. Detailed spatial information and environmental data around participants' home and workplace locations of travel behavior and built environment data are not yet available across the entire nation; however, do exist within selected regions (Frank et al. 2019b). Due to the same reasons, previous studies have mostly used environmental data at census-tract/block-group level given the difficulty of obtaining data at a more fine-grained resolution (such as around home/work) for large and diverse surveys (Zhou and Kockelman 2008, Zhang et al. 2012, Park et al. 2020). However, active transportation (such as walking) can occur at quite different spatial scales which differ for home and work locations. Future research can apply more fine-grained environmental measures in developing potential place-based thresholds for built environment variables (Frank et al. 2019b). The goal of this study was to create an applicable model using a larger sample with large enough variability enabling us to stratify by income and other factors with sufficient power. Also, as part of future work, the characteristics of the CBGs can be traced in Google Street View to better understand the contexts associated with the threshold effects.

From a data standpoint, the development of benchmarks for objectively assessed environmental features is complicated by at least two contrasting preferences or needs. On one hand, greater variability in the data is sought so to enhance (to the extent possible) the transferability of the environmental benchmarks. This can be achieved by using large-scale regional surveys (such as the one used in this study). On the other hand, the need to consider more fine-grained and location-specific measures of the environment (such as around home/workplace) limits the analyses to small areas – the findings from which are least transferable to other contexts. To this end, perhaps a reasonable trade-off could be to apply the methodology presented in other regions to explore potential localized benchmarks for key BE measures associated with greater physical activity.

Finally, the usefulness of more longitudinal evidence to better understanding the determinants of active travel is well-recognized in the literature (Frank et al. 2019a). Longitudinal designs are undoubtedly ideal to isolate the individual specific and self-selection effects from behavioral and exposure related impacts of the environment. For small geographies, comprehensive longitudinal data with fine-grained measures of the environment, travel behavior, and even clinically assessed health biomarkers exist (Frank et al. 2019a, Frank et al. 2019b, Wali et al. 2021). To this end, causally derived environmental benchmarks for a plethora of objectively measured environmental features can be developed. However, the limited geographic scope of such longitudinal designs also poses a remarkable challenge from the perspective of threshold development where greater geographic coverage (data diversity) is paramount. Until large longitudinal travel behavior data are publicly available with sufficient geographic coverage, the choice is between no information about potential thresholds vs. imperfect information.

8. ACKNOWLEDGEMENTS

The authors would like to acknowledge the Robert Wood Johnson Foundation (RWJF), especially Dr. Tracy Orleans, Senior Program Officer and Carolyn Miller, Senior Program Officer for supporting this study. This research was conducted by Urban Design 4 Health, Inc. (UD4H) as part of research funded by RWJF (Grant ID 74518) -- National Environmental Database – Walk Index, Benchmarks & Investments. The authors would also like to thank RWJF for funding the compilation of the National Environment Database of built, natural, and social environment indicators used in this study (Grant ID 72859). The contents of this paper are the sole responsibility of the authors and do not necessarily represent official views of the RWJF. This is solely a UD4H project providing all the data and resources for conducting the major analysis. A portion of the first author's self-development time (20%) during his postdoctoral training at MIT was used to revise this paper for presentation at 2020 TRB Annual Meeting and submission to journal. Finally, the authors would like to sincerely thank the three anonymous reviewers whose detailed input on earlier versions led us to conduct additional analyses and significantly improve the manuscript.

9. AUTHORS' CONTRIBUTIONS

BW led and designed the study, performed all statistical analyses, synthesized the results, and wrote the original and revised manuscript. LDF helped develop the study design and contributed to the original and revised manuscript. JEC assisted with study design and analysis approach, as well as contributed to the manuscript. EHF acted as Project Manager on the RWJF National Environmental Database, developed the environmental variables used in this study and reviewed and contributed to the manuscript.

10. REFERENCES

Aziz, H. A., S. V. Ukkusuri and J. Romero (2015). "Understanding short-term travel behavior under personal mobility credit allowance scheme using experimental economics." Transportation Research Part D: Transport and Environment **36**: 121-137.

Bandura, A. (1986). "Social foundations of thought and action." Englewood Cliffs, NJ **1986**.

Bozdogan, H. (2000). "Akaike's information criterion and recent developments in information complexity." Journal of mathematical psychology **44**(1): 62-91.

Brownson, R. C., E. A. Baker, R. A. Housemann, L. K. Brennan and S. J. Bacak (2001). "Environmental and policy determinants of physical activity in the United States." American journal of public health **91**(12): 1995-2003.

Buehler, R. (2012). "Determinants of bicycle commuting in the Washington, DC region: The role of bicycle parking, cyclist showers, and free car parking at work." Transportation research part D: Transport and Environment **17**(7): 525-531.

Carter, D. J., W. W. Hunter, C. V. Zegeer, J. R. Stewart, H. F. Huang and A. Do (2006). Pedestrian and bicyclist intersection safety indices, United States. Federal Highway Administration.

Cervero, R. (2002). "Built environments and mode choice: toward a normative framework." Transportation Research Part D: Transport and Environment **7**(4): 265-284.

Cervero, R. and K. Kockelman (1997). "Travel demand and the 3Ds: Density, diversity, and design." Transportation Research Part D: Transport and Environment **2**(3): 199-219.

Cervero, R. and J. Murakami (2010). "Effects of built environments on vehicle miles traveled: evidence from 370 US urbanized areas." Environment and planning A **42**(2): 400-418.

Crane, R. (2000). "The influence of urban form on travel: an interpretive review." Journal of Planning literature **15**(1): 3-23.

Ding, C., X. Cao and Y. Wang (2018b). "Synergistic effects of the built environment and commuting programs on commute mode choice." Transportation Research Part A: Policy and Practice **118**: 104-118.

Ding, C., X. J. Cao and P. Næss (2018a). "Applying gradient boosting decision trees to examine non-linear effects of the built environment on driving distance in Oslo." Transportation Research Part A: Policy and Practice **110**: 107-117.

Ding, D., J. F. Sallis, T. L. Conway, B. E. Saelens, L. D. Frank, K. L. Cain and D. J. Slymen (2012). "Interactive effects of built environment and psychosocial attributes on physical activity: a test of ecological models." Annals of Behavioral Medicine **44**(3): 365-374.

Eom, H.-J., G.-H. J. T. R. P. D. T. Cho and Environment (2015). "Exploring thresholds of built environment characteristics for walkable communities: Empirical evidence from the Seoul Metropolitan area." **40**: 76-86.

Ewing, R., G. Tian, J. Goates, M. Zhang, M. J. Greenwald, A. Joyce, J. Kircher and W. Greene (2015). "Varying influences of the built environment on household travel in 15 diverse regions of the United States." Urban Studies **52**(13): 2330-2348.

Forsyth, A., J. M. Oakes, K. H. Schmitz and M. Hearst (2007). "Does residential density increase walking and other physical activity?" Urban studies **44**(4): 679-697.

Fotheringham, A. S. (1981). "Spatial structure and distance-decay parameters." Annals of the Association of American Geographers **71**(3): 425-436.

Frank, L., P. Engelke and T. Schmid (2003). Health and community design: The impact of the built environment on physical activity, Island Press.

Frank, L. D., M. A. Andresen and T. L. Schmid (2004). "Obesity relationships with community design, physical activity, and time spent in cars." American journal of preventive medicine **27**(2): 87-96.

Frank, L. D. and P. O. Engelke (2001). "The built environment and human activity patterns: exploring the impacts of urban form on public health." Journal of planning literature **16**(2): 202-218.

Frank, L. D., N. Iroz-Elardo, K. E. MacLeod and A. Hong (2019a). "Pathways from built environment to health: A conceptual framework linking behavior and exposure-based impacts." Journal of Transport & Health **12**: 319-335.

Frank, L. D., J. L. Kuntz, J. E. Chapman, E. H. Fox, J. F. Dickerson, R. T. Meenan, B. E. Saelens, D. R. Young, J. Boone-Heinonen and S. P. Fortmann (2019b). "The Health and economic effects of light rail lines: design, methods, and protocol for a natural experiment." BMC public health **19**(1): 200.

Frank, L. D. and G. Pivo (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., T. L. Schmid, J. F. Sallis, J. Chapman and B. E. Saelens (2005). "Linking objectively measured physical activity with objectively measured urban form: findings from SMARTRAQ." American journal of preventive medicine **28**(2): 117-125.

Greene, W. H. and D. A. Hensher (2010). Modeling ordered choices: A primer, Cambridge University Press.

Guo, Y. and S. Y. He (2020). "Built environment effects on the integration of dockless bike-sharing and the metro." Transportation Research Part D: Transport and Environment **83**: 102335.

Hensher, D. A., J. M. Rose and W. H. Greene (2005). Applied choice analysis: a primer, Cambridge University Press.

Hensher, D. A. and C. Sullivan (2003). "Willingness to pay for road curviness and road type." Transportation Research Part D: Transport and Environment **8**(2): 139-155.

Hu, S., P. Chen, H. Lin, C. Xie and X. Chen (2018). "Promoting carsharing attractiveness and efficiency: An exploratory analysis." Transportation Research Part D: Transport and Environment **65**: 229-243.

Huston, S. L., K. R. Evenson, P. Bors and Z. Gizlice (2003). "Neighborhood environment, access to places for activity, and leisure-time physical activity in a diverse North Carolina population." American journal of health promotion **18**(1): 58-69.

Kaczynski, A. T., L. R. Potwarka, B. J. Smale and M. E. Havitz (2009). "Association of parkland proximity with neighborhood and park-based physical activity: variations by gender and age." Leisure Sciences **31**(2): 174-191.

Kang, B., A. V. Moudon, P. M. Hurvitz, B. E. J. T. r. p. D. t. Saelens and environment (2017). "Differences in behavior, time, location, and built environment between objectively measured utilitarian and recreational walking." **57**: 185-194.

Kirtland, K. A., D. E. Porter, C. L. Addy, M. J. Neet, J. E. Williams, P. A. Sharpe, L. J. Neff, C. D. Kimsey Jr and B. E. Ainsworth (2003). "Environmental measures of physical activity supports: perception versus reality." American journal of preventive medicine **24**(4): 323-331.

Koohsari, M. J., H. Badland and B. Giles-Corti (2013). "(Re) Designing the built environment to support physical activity: bringing public health back into urban design and planning." Cities **35**: 294-298.

Leslie, E., N. Coffee, L. Frank, N. Owen, A. Bauman and G. Hugo (2007). "Walkability of local communities: using geographic information systems to objectively assess relevant environmental attributes." Health & place **13**(1): 111-122.

Li, L., Z. Wang, L. Chen and Z. Wang (2020). "Consumer preferences for battery electric vehicles: A choice experimental survey in China." Transportation Research Part D: Transport and Environment **78**: 102185.

Ma, L. and J. Dill (2015). "Associations between the objective and perceived built environment and bicycling for transportation." Journal of Transport & Health **2**(2): 248-255.

McGinn, A. P., K. R. Evenson, A. H. Herring, S. L. Huston and D. A. Rodriguez (2007). "Exploring associations between physical activity and perceived and objective measures of the built environment." Journal of urban health **84**(2): 162-184.

McIntosh, J., R. Trubka, J. Kenworthy and P. Newman (2014). "The role of urban form and transit in city car dependence: Analysis of 26 global cities from 1960 to 2000." Transportation Research Part D: Transport and Environment **33**: 95-110.

Oakes, J. M., A. Forsyth and K. H. Schmitz (2007). "The effects of neighborhood density and street connectivity on walking behavior: the Twin Cities walking study." Epidemiologic Perspectives & Innovations **4**(1): 16.

Owen, N., N. Humpel, E. Leslie, A. Bauman and J. F. Sallis (2004). "Understanding environmental influences on walking: review and research agenda." American journal of preventive medicine **27**(1): 67-76.

Park, K., R. Ewing, S. Sabouri, D.-a. Choi, S. Hamidi and G. Tian (2020). "Guidelines for a polycentric region to reduce vehicle use and increase walking and transit use." Journal of the American Planning Association (under production). .

Pinjari, A. and C. Bhat (2006). "Nonlinearity of response to level-of-service variables in travel mode choice models." Transportation Research Record: Journal of the Transportation Research Board(1977): 67-74.

Rodríguez, D. A., K. R. Evenson, A. V. D. Roux and S. J. Brines (2009). "Land use, residential density, and walking: the multi-ethnic study of atherosclerosis." American journal of preventive medicine **37**(5): 397-404.

Saelens, B. E., J. F. Sallis and L. D. Frank (2003). "Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures." Annals of behavioral medicine **25**(2): 80-91.

Sallis, J. F., R. B. Cervero, W. Ascher, K. A. Henderson, M. K. Kraft and J. Kerr (2006). "An ecological approach to creating active living communities." Annu. Rev. Public Health **27**: 297-322.

Sallis, J. F., L. D. Frank, B. E. Saelens and M. K. Kraft (2004). "Active transportation and physical activity: opportunities for collaboration on transportation and public health research." Transportation Research Part A: Policy and Practice **38**(4): 249-268.

Sarkar, C., C. Webster, M. Pryor, D. Tang, S. Melbourne, X. Zhang and L. Jianzheng (2015). "Exploring associations between urban green, street design and walking: Results from the Greater London boroughs." Landscape and Urban Planning **143**: 112-125.

Sevtsuk, A., R. Kalvo and O. Ekmekci (2016). "Pedestrian accessibility in grid layouts: the role of block, plot and street dimensions." Urban Morphology **20**(2): 89-106.

StataCorp, L. (2013). "Stata multilevel mixed-effects reference manual." College Station, TX: StataCorp LP.

USDA (2010). 2010 Rural-Urban Commuting Area (RUCA) Codes. US Department of Agriculture, Economic Research Service. URL: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/documentation/.

Van Dyck, D., E. Cerin, T. L. Conway, I. De Bourdeaudhuij, N. Owen, J. Kerr, G. Cardon, L. D. Frank, B. E. Saelens and J. F. Sallis (2012a). "Associations between perceived neighborhood environmental attributes and adults' sedentary behavior: findings from the USA, Australia and Belgium." Social science & medicine **74**(9): 1375-1384.

Van Dyck, D., E. Cerin, T. L. Conway, I. De Bourdeaudhuij, N. Owen, J. Kerr, G. Cardon, L. D. Frank, B. E. Saelens and J. F. Sallis (2012b). "Perceived neighborhood environmental attributes associated with adults' transport-related walking and cycling: Findings from the USA, Australia and Belgium." International journal of behavioral nutrition and physical activity **9**(1): 70.

Van Dyck, D., E. Cerin, T. L. Conway, I. De Bourdeaudhuij, N. Owen, J. Kerr, G. Cardon, L. D. Frank, B. E. Saelens and J. F. Sallis (2013). "Perceived neighborhood environmental attributes associated with adults' leisure-time physical activity: Findings from Belgium, Australia and the USA." Health & place **19**: 59-68.

Wagenmakers, E.-J. and S. Farrell (2004). "AIC model selection using Akaike weights." Psychonomic bulletin & review **11**(1): 192-196.

Wali, B., L. Frank, D. Young, B. Saelens, R. Meenan, J. Dickerson, E. Keast, J. Kuntz and S. Fortmann (2021). Linking Microscopic Built Environment with Objectively Measured Physical Activity and Health Care Expenditures from an Integrated Health Care System. Accepted for presentation at the 100th Annual Meeting of the Transportation Research Board, Washington DC, USA.

Wali, B., A. J. Khattak, H. Bozdogan and M. Kamrani (2018a). "How is driving volatility related to intersection safety? A Bayesian heterogeneity-based analysis of instrumented vehicles data." Transportation research part C: emerging technologies **92**: 504-524.

Wali, B., A. J. Khattak and T. Karnowski (2020). "The relationship between driving volatility in time to collision and crash-injury severity in a naturalistic driving environment." Analytic Methods in Accident Research **28**: 100136.

Wali, B., A. J. Khattak and X. Li (2018b). Exploring Non-Linear Dependencies in Correlates of Roadway Crashes: Application of Generalized Additive & Piecewise Linear Count Data Models. 97th Annual Meeting of Transporation Research Board. Washington DC.

Wali, B., A. J. Khattak, J. Waters, D. Chimba and X. Li (2018c). "Development of Safety Performance Functions: Incorporating Unobserved Heterogeneity and Functional Form Analysis." Transportation Research Record: Journal of the Transportation Research Board: 0361198118767409.

WHO (2010). Global recommendations on physical activity for health, World Health Organization.

Wood, S. N. (2017). Generalized additive models: an introduction with R, Chapman and Hall/CRC.

Zhang, L., J. Hong, A. Nasri and Q. Shen (2012). "How built environment affects travel behavior: A comparative analysis of the connections between land use and vehicle miles traveled in US cities." Journal of transport and land use **5**(3): 40-52.

Zhou, B. and K. M. Kockelman (2008). "Self-selection in home choice: Use of treatment effects in evaluating relationship between built environment and travel behavior." Transportation Research Record **2077**(1): 54-61.

Zupan, J., B. Pushkarev and R. P. Association (1980). Urban rail in America: An exploration of criteria for fixed-guideway transit, United States. Urban Mass Transportation Administration.

11. APPENDIX A

The results of GAMs based on partial segmentation approach are provided in Table A1 and Figure A1.

TABLE A1: Estimation Results of Discrete Outcome Generalized Additive Models with Interactions Between Income Indicator and Built Environment Smoothed Spline Effects

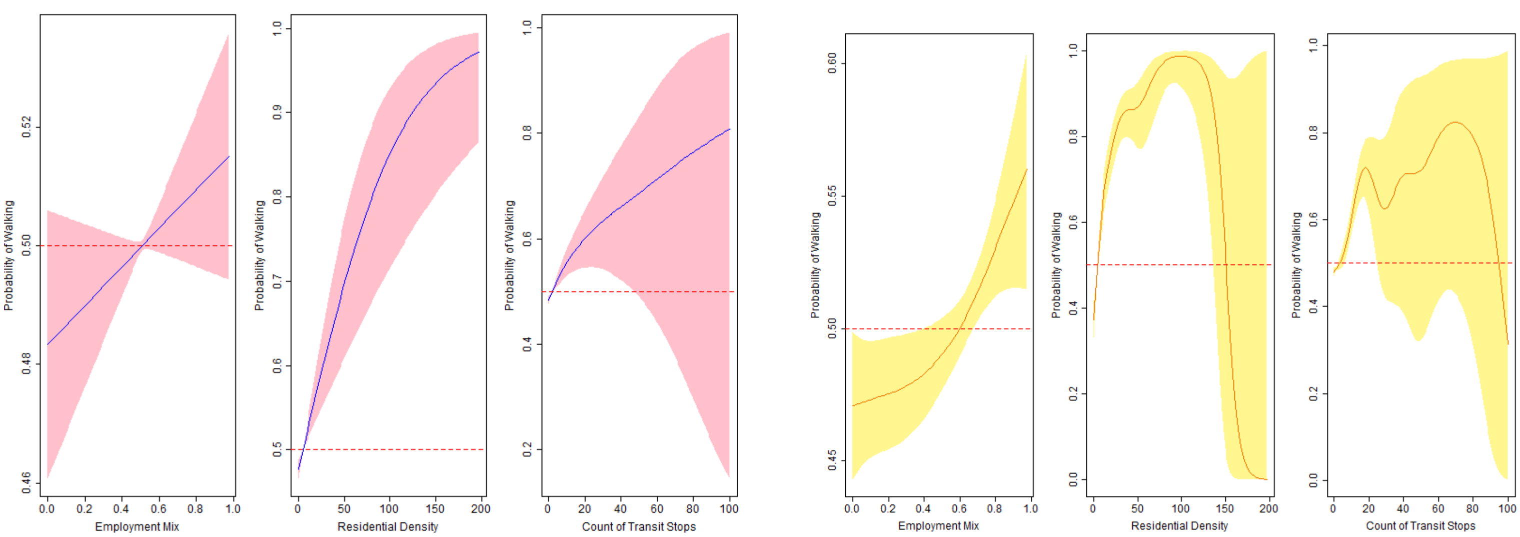
Variables	Aggregate Data (N = 44253) -- Spline by Income Interactions	
	Effective DF	**Chi-square (p-value)**
Walkability-Related Spline Effects		
Low-income		
Employment Mix	1.004	2.14 (0.1432)
Residential Density	1.003	17.90 (0.0000)
Number of Transit Stops	1.842	19.59 (0.0000)
Vehicles Miles Travelled (in 1000s)	3.097	45.69 (0.0000)
Intersection Density	2.48	8.62 (0.0421)
Medium-income		
Employment Mix	1.866	10.97 (0.0054)
Residential Density	6.044	108.88 (0.0000)
Number of Transit Stops	5.536	54.56 (0.0000)
Vehicles Miles Travelled (in 1000s)	1.811	2.52 (0.4051)
Intersection Density	3.66	22.12 (0.0004)
High-income		
Employment Mix	1.001	3.69 (0.0548)
Residential Density	4.99	35.47 (0.0000)
Number of Transit Stops	3.167	39.46 (0.0000)
Vehicles Miles Travelled (in 1000s)	7.241	16.58 (0.0400)
Intersection Density	6.902	37.44 (0.0004)
Controls	**β**	***z-score***
Intercept	-0.51	-5.37
Medium-income (50K - 100K USD)	-0.18	-4.48
High-income (100K USD or more)	0.06	1.40
Age: 25 - 34 years old	-0.37	-5.97
Age: 35 - 44 years old	-0.46	-7.79
Age: 45 - 54 years old	-0.42	-7.59
Age: 55 - 64 years old	-0.47	-8.43
Population (in 1000s)	-0.11	-8.77
Rented home (1/0)	0.29	8.06
Employed (1/0)	-0.34	-10.56
College degree (1/0)	0.42	12.90
Kids, multi parent (1/0)	-0.25	-5.37
Kids, single parent (1/0)	-0.29	-2.65
Household size	0.18	9.10
No disability (1/0)	0.16	2.80

Student count in household	-0.04	-2.24
License count in household	-0.56	-25.39
Male (1/0)	-0.12	-4.47

TABLE A1: Estimation Results of Discrete Outcome Generalized Additive Models with Interactions Between Income Indicator and Built Environment Smoothed Spline Effects (Continued)

Summary Statistics	***Value***
Log-likelihood at convergence	-17344.70
Akaike Information Criterion (AIC)	**34828.65**
AIC of GAM based on aggregate data (see Table 5 in revised manuscript).	**34967.13**
AIC Summation of three separate income-based GAMs (see Table 5 in the revised manuscript).	**(34633.2) 11447.50+9945.88+13239.82**
McFadden Pseudo R-Squared	0.11
UBRE Score	-0.21

Notes: UBRE is Un-biased Risk Estimator, smaller values indicating better fitting models; DF is Degree of Freedom; USD is United States Dollar; AIC is Akaike Information Criterion. The McFadden Pseudo R-squared is computed as $\left[1 - {}^{LL_{CONVERGENCE}}/_{LL_{NULL}}\right]$ – and the ratio of the two likelihoods suggests the level of improvement over the intercept model offered by the full model.



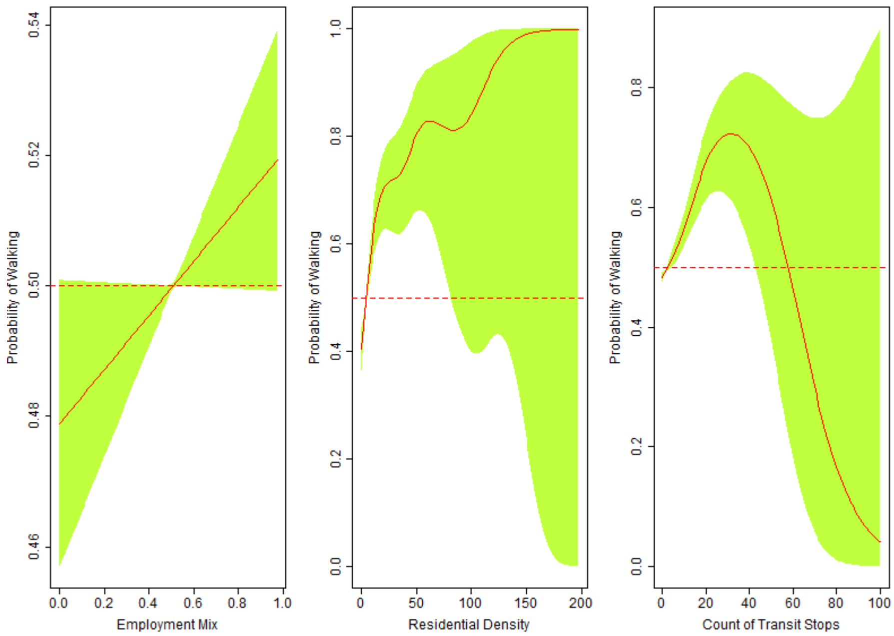
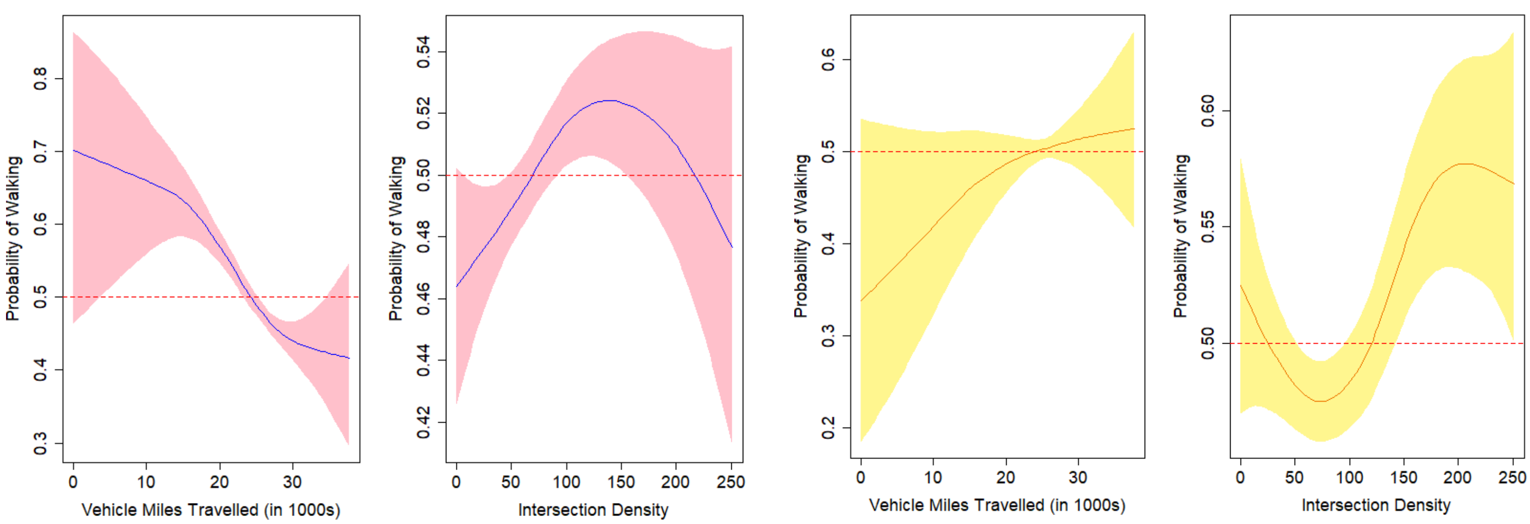


Figure A1: Dose-Response Relationships from GAMs Based on Income-Built Environment Spline Term Interactions for Low-Income Adults (pink), Medium-Income Adults (yellow), and High-Income Adults (green)

Notes: Raw values of a built environment predictor are plotted on X-axis. Probability of transport walking obtained from Generalized Additive Model plotted on Y-axis. The horizontal red dotted line represents in the context under discussion a 50-50 chance of an adult in metropolitan regions participating in transport walking. The horizontal bar above the X-axis (the "rug") represents the sample across the continuous range of an explanatory factor). The upper and lower lines in the shaded portions represent 2 standard errors above and below the estimate of the smooth being plotted.



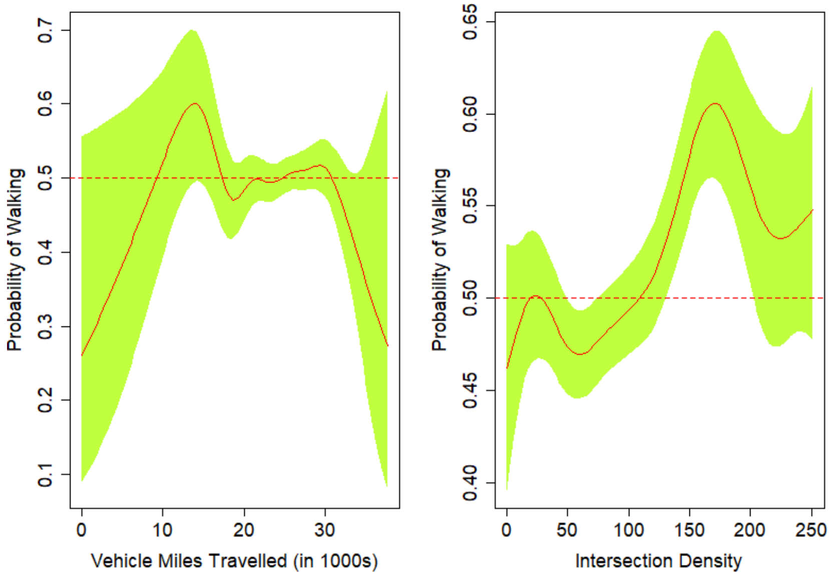


Figure A1 (Continued): Dose-Response Relationships from GAMs Based on Income-Built Environment Spline Term Interactions for Low-Income Adults (pink), Medium-Income Adults (yellow), and High-Income Adults (green)

Notes: Raw values of a built environment predictor are plotted on X-axis. Probability of transport walking obtained from Generalized Additive Model plotted on Y-axis. The horizontal red dotted line represents in the context under discussion a 50-50 chance of an adult in metropolitan regions participating in transport walking. The horizontal bar above the X-axis (the “rug”) represents the sample across the continuous range of an explanatory factor). The upper and lower lines in the shaded portions represent 2 standard errors above and below the estimate of the smooth being plotted.