sustainability

Article

Exploring Associations between Multimodality and Built Environment Characteristics in the U.S.

Sangwan Lee

Nohad A. Toulan School of Urban Studies and Planning, Portland State University, 506 SW Mill Street, Portland, OR 97201, USA; sangwan@pdx.edu

Abstract: This study demonstrated associations between multimodality and built environment characteristics, and proposed policy implications for fostering multimodal travel behaviors. It conducted a U.S. nationwide analysis using ordinary least square regression and gradient boosting decision tree regressor models with American Community Survey 2015–2019 5-year estimates and the United States Environmental Protection Agency Smart Location Database version 3.0. Notable findings were as follows: First, built environment characteristics were found to be statistically significant predictors of multimodality across the U.S. Second, certain features were identified as having considerable importance, specifically including population density, regional accessibility, walkability index, and network density, all of which should be given particular attention by transportation and land-use planners. Third, the non-linear effects of built environment characteristics on multimodality suggested an effective range to encourage multimodal transportation choice behaviors in various situations. The findings can guide the development of effective strategies to transform the built environment, which may subsequently be used to minimize reliance on automobiles and promote people to travel more sustainably.

Keywords: multimodality; built environment; planning; U.S. nationwide analysis

check for updates

Citation: Lee, S. Exploring Associations between Multimodality and Built Environment Characteristics in the U.S. *Sustainability* **2022**, *14*, 6629. https://doi.org/10.3390/su14116629

Academic Editors: Karen Lucas and Emma Tsoneva

Received: 29 April 2022
Accepted: 26 May 2022
Published: 28 May 2022

1. Introduction

Multimodality refers to a modal variety that includes single-occupied vehicles as well as public transit, bicycling, walking, carpooling, and other modes; in other words, many perceive it as a counter-movement to autocentrism [1]. Multimodality has gained popularity in urban affairs and transportation planning studies [2], since it has been interconnected to economic, environmental, and social benefits [3–6]. Accordingly, there have been strategies for encouraging sustainable mobility patterns mainly at three levels, including the vehicle level, the transportation system (e.g., infrastructure), and the level of traffic participants (e.g., vehicle owners) [7]. However, the role and importance of the built environment on multimodality have not been sufficiently explored and considered from the planning perspective. Therefore, the purposes of this study were to (1) demonstrate associations between multimodality and built environment characteristics and (2) propose policy implications for fostering multimodal travel behaviors. The research questions were as follows: (1) whether there were significant relationships between multimodality and built environment characteristics across the United States, (2) whether the built environment characteristics were more important than other covariates in showing multimodal travel patterns, and (3) whether there were non-linear effects. Having these types of insights will ultimately aid in the development of appropriate plans and policies to promote multimodality more effectively. The following sections review previous literature, describe the research design, present findings, and conclude this study.

2. Literature Review

This study attempts to connect two bodies of the previous literature: (1) multimodality and (2) the relationship between travel behavior and the built environment. Thus, this section presents the two themes to offer a broader context of this study and then discusses the research gap and contribution of this work.

2.1. Multimodality

2.1.1. Needs for Multimodality

Transportation mode choice behavior is intertwined with many aspects of our lives, such as employment, housing, schools, shopping, and health [8]. As the automobility system has evolved and established itself as the dominant mode of transportation, it has pushed other modes of transportation, such as public transit and active transportation, to the sidelines [9]. The widespread use of automobiles has linked to obesity, traffic congestion, environmental pollution, urban sprawl, and social marginalization, which have been long acknowledged as severe consequences [6,10–14]. Moreover, the automobile-dependent society cannot meet the diverse transportation needs of different population groups, such as youths, seniors, adults unable to drive due to disability, and low-income households burdened by vehicle expenses.

Accordingly, planners and researchers are attempting to understand how a transition from automobile use toward more sustainable modes of transportation can be achieved [9,15,16]. These considerations underline the practical importance of knowing the circumstances under which people increase their usage of diverse transportation modes in their daily lives [17] to not only alleviate a variety of issues that an auto-dependent society can cause, but also bring benefits, such as quality of life [18].

2.1.2. Factors Influencing Multimodality

A small body of literature has explored factors influencing multimodality. For instance, they generally found that multimodality has been significantly associated with several socio-demographic characteristics, such as household income, employment status, education attainment, and race/ethnicity [19–21]. Additionally, previous literature has identified additional factors, including personal attitudinal features, car ownership, and current travel behaviors [22,23]. In addition, Astroza et al. [24] found that utilizing technology, such as smartphones, resulted in an expansion of the multimodal travel dimensions. Interestingly, Scheiner et al. [17] demonstrated significant associations between life-course events and multimodal travel behaviors; for instance, the multimodality of the parents increases when a child leaves the household.

2.2. Travel Behavior and Built Environment

Another body of literature on this study contributes to understanding how travel behavior and built environment characteristics have been connected [25–28]. The built environment in the literature has generally been operationalized in so-called D variables, which include density (e.g., population density), diversity (e.g., job–housing balance), design (e.g., intersection density), destination accessibility (e.g., regional centrality), and distance to transit.

Previous studies in the U.S. have observed the associations between built environments and travel behavior of different types of transportation, such as vehicles, public transportation, and active transportation. For instance, Sabouri et al. [29] found a significant association between vehicle ownership and built environment characteristics. Moreover, previous studies have acknowledged the significant association between bicycling and the D variables [30–32]. Additionally, existing research has explored the impact of the built environment on the ridership of shared mobility services, such as ride-hailing and bike-sharing services [33–35]. For instance, Malik et al. [36] observed that individuals in vibrant and walkable communities have a higher proportion of choosing ride-hailing. Additionally, population and employment density, transit density, university density, and the degree of mixed land use were positively and significantly associated with the trip generation of bike-sharing services [37,38]. In sum, generalizing this vast literature, the D variables that quantify built environment characteristics have been significantly associated with travel behavior at the individual or zonal levels in the U.S.

2.3. Research Gaps and Contribution of This Study

The literature review identified several research gaps. First, the connection between multimodality and the built environment features has not been fully explored. Second, a large body of research has examined the relationship between travel behavior of a particular transportation mode, such as ride-hailing services, and the built environment. Third, this strand of literature has investigated limited geographical scales (e.g., a case study of a city). Fourth, to the best of my knowledge, none of the previous literature has employed both econometrics models and machine learning algorithms to address a variety of questions.

Therefore, this study aims at enriching the body of literature by filling the critical gaps and investigating the following three research questions. First, were relationships between the built environment and multimodality significant at the Census Block Group level throughout the United States? Second, to what extent did the built environment characteristics play a role in showing multimodal travel behaviors? Third, were there non-linear effects of built environment features on multimodality? The answers to the three questions contribute to the existing body of research and inform policymakers about transforming travel behaviors in a more sustainable manner.

3. Materials and Methods

This section describes the study area, data, and methodological approaches used to address the three research questions in this study.

3.1. Study Area and Data Collection

This study area was the United States. A few large-scale studies conducted in the United States have focused on several metropolitan areas [35,39,40]. In this case, selection bias may exist since travel behaviors in highly populated areas such as metropolitan areas may differ from those in less densely populated areas. Therefore, this study did not use a certain population cap (e.g., metropolitan area with 200,000 persons or greater) to exclude certain areas.

This study used two publicly available data sources: (1) American Community Survey 2015–2019 5-year estimates (ACS) and (2) the United States Environmental Protection Agency Smart Location Database version 3.0 (SLD). The two data sets were appropriate since they contained crucial information for this study, such as commuter transportation mode choices and built environment characteristics with representative samples throughout the U.S. Unfortunately, since ACS does not collect data on trips that are not considered commutes, including trips for recreation, school, and personal obligations, this study only focused on commute trips. Given differences in commute and non-commute trip patterns in the U.S., it is one of the limitations of this study.

This study did not use the latest data (i.e., ACS 2020) for the following reasons. First, ACS 2020 contained information derived from an interview sample of persons interviewed between March and December 2020, indicating that the data included some information since the COVID-19 outbreak. It can raise a validity issue of the final results due to the considerable influence of the pandemic on the transportation sector in the U.S. [41–44]. Additionally, according to Census Bureau [45], the pandemic adversely influenced the data collection process and may produce quality issues. Specifically, the sample obtained may not represent the entire population in the U.S. because the final number of interviews significantly reduced in 2020 [46].

The unit of analysis was the census block group (CBG), the smallest geographical unit in the nationwide data sets used in this study. Initially, data included 220,333 CBGs in the U.S. The sample was reduced due to missing values in either ACS or SLD for the large number of variables used in this study. Thus, the results of this study presented here were based on 206,380 valid CBGs.

3.2. Variables

This subsection summarizes the rationale for constructing and selecting dependent and independent variables. Table 1 describes the details of the variables, and Table 2 shows their descriptive statistics.

Table 1. Variables used in this study.

Name	Description	Equation	Data Source
	Dependent Variable		
Multimodality index	Entropy index for multimodality	Y	ACS
	Independent Variables of Interest		
Pop_den	The total population per acre at the census block group level in 10,000	X_1	ACS
Diversity_HH_job	Jobs to household balance in 1,000	X_2	SLD
Diversity_job	Entropy index for job diversity at the census block group level using the eight-tier employment categories, including retail, office, industry, service, entertainment, education, healthcare, and public sectors	X_3	SLD
Net_den	Network density in 10,000	X_4	SLD
Int_den	Intersection density in 10,000	X_5	SLD
Walkability index	Walkability index characterized by components of the built environment that influence the likelihood of walking	X_6	SLD
Job_proximity	Percentage of residents who take less than 10 min to commute in 10	X_7	SLD
Auto_accessibility	The relative regional accessibility measure by using the regional centrality index by auto	X_8	SLD
Transit_accessibility	The relative regional accessibility measure by using the regional centrality index by transit	X_9	SLD
	Independent Variables		
HH_size	Average household size	X_{10}	ACS
HH_income	Median household income in 10,000	X_{11}	ACS
White	Percentage of the residents who are non-Hispanic white in 10	X_{12}	ACS
Black	Percentage of the residents who are non-Hispanic black in 10	X_{13}	ACS
Asian	Percentage of the residents who are non-Hispanic Asian in 10	X_{14}	ACS
Single	Percentage of the residents who have not married in 10	X_{15}	ACS
Low education	Percentage of the residents who attained less than a bachelors' degree, including high school and college, in 10	X_{16}	ACS
No_car	Percentage of the residents who do not own a car in 10	X_{17}	ACS
Work_at_home	Percentage of the residents who work at home in 10	X_{18}	ACS

Note: All variables are at the census block group (CBG) level. Further details on independent variables can be found in the technical documentation of the two data sources. Source: American Community Survey 2015–2019 5-year estimates (ACS) and the United States Environmental Protection Agency Smart Location Database version 3.0 (SLD).

Table 2. Descriptive statistics of the variables (N = 206,380).

Variables	Mean	Median	Std. Dev	Min	Max
Multimodality index	0.555	0.525	0.27	0.000	1.588
Pop_den	0.635	0.264	1.53	0.000	81.131
Diversity_HH_job	0.001	0.001	0.01	0.000	1.631
Diversity_job	0.539	0.576	0.22	0.000	0.994
Net_den	0.001	0.001	0.00	0.000	0.012
Int_den	0.007	0.006	0.01	0.000	0.193
Walkability index	9.596	9.167	4.35	1.000	20.000
Job_proximity	1.293	0.961	1.22	0.000	10.000
Auto_accessibility	0.433	0.441	0.28	0.000	1.000
Transit_accessibility	0.112	0.000	0.20	0.000	1.000
HH_size	2.633	2.560	0.59	1.010	9.250
HH_income	6.707	5.917	3.63	0.249	25.000
White	7.313	8.232	2.63	0.000	10.000
Black	1.314	0.299	2.23	0.000	10.000
Asian	0.469	0.073	0.98	0.000	10.000
Single	3.319	3.076	1.42	0.000	10.000
Low education	7.008	7.521	2.04	0.000	10.000
No_car	1.402	0.620	1.92	0.000	10.000
Work_at_home	0.492	0.350	0.55	0.000	10.000

3.2.1. Dependent Variable: Operationalizing Multimodality

This study used the multimodality index as a dependent variable. This study used the entropy index to operationalize multimodality at CBGs, a widely-used matrix to represent diversity in a variety of fields [47,48].

The entropy index was appropriate in this study for the following reasons. First and foremost, the entropy index adequately assesses the evenness of the distribution across the shares of different transportation modes for commute trips, including single-occupied vehicles, carpooling, public transportation, active transportation (e.g., bicycling and walking), and others at CBGs. Second, a straightforward variation ratio to measure the share of trips made by certain transportation modes [18] can be inappropriate due to the absence of consideration of the distribution of diverse transportation mode choices. Third, previous literature has developed diverse indicators, including the Herfindahl index, Dalton index, and probability-based multimodality indicator [5,17,49]. However, given that a comparative study by Diana and Pirra [50] revealed that none of the indices consistently outperforms all the others in any situation, the choice of the entropy index over others may not pose a threat to the validity of the final models in this study.

The multimodality index (MI_i) is formalized as follows:

$$MI_i = -[S_{sov,i} \times \log(S_{sov,i}) + S_{cp,i} \times \log(S_{cp,i}) + S_{pt,i} \times \log(S_{pt,i}) + S_{at,i} \times \log(S_{at,i}) + S_{others,i} \times \log(S_{others,i})] \tag{1}$$

where $S_{sov,i}$ is the share of single-occupied vehicles for commute trips at CBG i, $S_{cp,i}$ is the share of carpooling, $S_{pt,i}$ is the share of public transportation, $S_{at,i}$ is the share of active transportation, and $S_{others,i}$ is the share of other modes of transportation.

3.2.2. Independent Variables

This study used nine independent variables of interest regarding built environment characteristics based on the D variables established in previous literature [26]: (1) density (i.e., population density), (2) diversity (i.e., job–household balance and job diversity), (3) design (i.e., intersection density, network density, and walkability index), and (4) destination accessibility (i.e., proximity to job and regional accessibility by car and transit). Unfortunately, this study dropped one important variable: the network distance to the nearest transit station, due to two-thirds of the missing values in the data set. Specifically, since SLD omitted CBGs that were further than three-quarter miles from a transit stop, the inclusion of the variable may lead to the validity issue of the results. Further details, such as the equations to operationalize the built environment indicators, advantages, and limitations, can be found in the SLD technical documentation [51].

This study controlled for nine covariates since multimodality has been significantly associated with the factors, such as household income, employment status, education attainment, car availability, and race/ethnicity [20,22]. This study did not apply the regional dummy variable to control the spatial heterogeneity among regional areas because it can lead to over-fitting issues, particularly in a machine learning algorithm. Figure 1 shows the Pearson correlation matrix between any two variables used in this study and indicates that a multicollinearity issue may be present. However, variance inflation factors (VIFs) of variables in Table 3 were lower than 10, confirming that the inclusion of all 18 variables in a model should not be cause for concern.

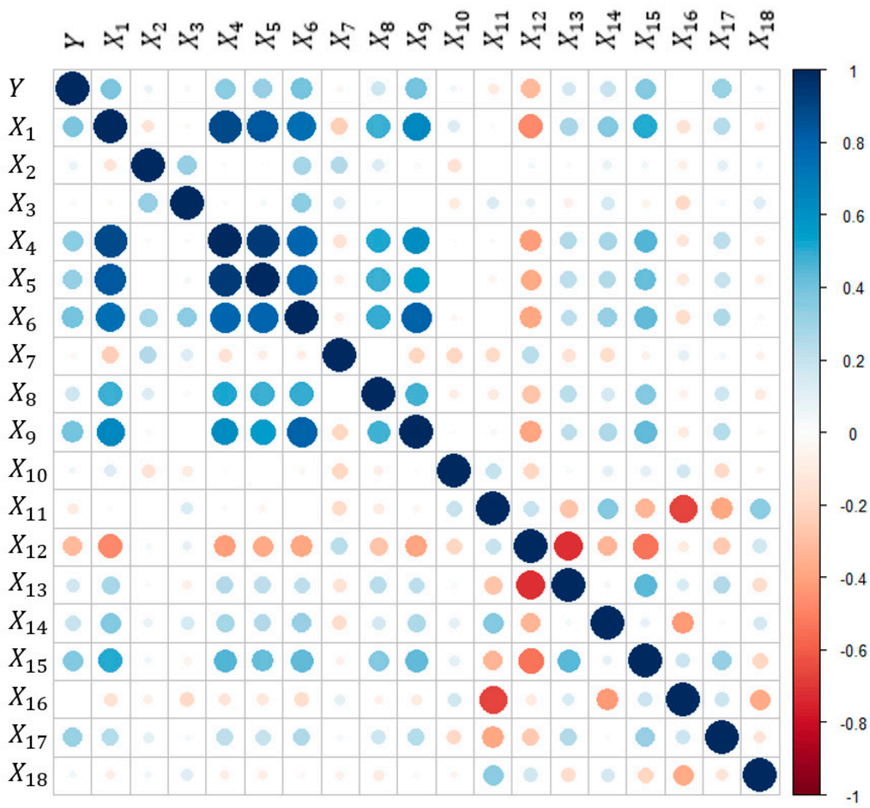

Figure 1. Pearson correlation matrix between variables.

3.3. Analytic Strategies

This study developed two models to answer the three research questions: (1) the ordinary least square regression model in econometrics and (2) the gradient boosting decision tree regressor model in machine learning.

3.3.1. Ordinary Least Square Regression

This study first employed the ordinary least square (OLS) regression model to explore (1) statistically significant relationships between the multimodality index, and multidimensional covariates, including built environment characteristics, (2) directions, and (3) magnitudes. The equation of OLS is as follows:

$$y = \beta_0 + \beta_1 X_1 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \beta_{11} X_{11} + \beta_{12} X_{12} + \beta_{13} X_{13} + \beta_{14} X_{14} + \beta_{15} X_{15} + \beta_{16} X_{16} + \beta_{17} X_{17} + \beta_{18} X_{18} + \varepsilon \quad (2)$$

where the dependent variable y is the multimodality index, and ε denotes the error term. The focus of the model was on the parameter estimate β_{i} for variables X_{i} described in Table 1.

Despite the possible spatial correlations observed in Figure 1, this study did not control for spatial autocorrelation in the final model since it may be less valuable to consider the spatial relationships when the unit of analysis is an administrative boundary (i.e., CBGs) [35]. Two regional job accessibility variables (X_8 and X_9) were instead included to capture the influence of contextual effects that may exist within it.

Moreover, this study did not use a multilevel regression, which is the widely-used method in previous literature to explore relationship between built environment and travel behavior [27], since there was no variance associated with the levels of the data used in this study.

3.3.2. Gradient Boosting Decision Tree Regressor

This study also used Gradient Boosting Decision Tree Regressor Model (GBDT) proposed by Friedman [52]. The machine learning algorithm has not been employed previously in the literature, although it is capable of estimating feature importance and non-linearity, which aids in answering the remaining two research questions and provides vital insights into this study. Noteworthy is the fact that the GBDT does not differentiate between causes and effects; rather, it draws associations between the dependent variable (target) and the covariates (input features).

The underlying process of the algorithm is to merge a series of weak base classifiers with different weights into a final one [53]. It is different from the traditional boosting algorithm since it causes global convergence of the algorithm by following the direction of the negative gradient [54]. Its generic procedure consists of several steps when $\{x_i, y_i\}_{i=1}^{n}$ assumes the dataset [55,56]. Specifically, the first step initializes the initial constant value of the algorithm β:

$$f_o(x) = \text{ arg min}_{\beta} \sum_{i=1}^{N} L(y_i, \beta) \tag{3}$$

Second, the gradient direction of residuals (called pseudo-residuals) is estimated for the number of iterations *m to M*:

$$y_i = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(X)-F_{m-1(x)}}, \ i = \{1, 2, \ldots, N\} \tag{4}$$

Third, the basic classifiers are used to fit sample data and produce the initial algorithm (also called base learner). The least squared approach finds the parameter of the algorithm a_m and fits the algorithm $h(x_i; a_m)$:

$$a_m = \text{ arg min}_{\alpha,\beta} \sum_{i=1}^{N} [y_i - \beta h(x_i; \alpha)]^2 \tag{5}$$

Then, the following loss function is minimized by solving the one-dimensional optimization problem:

$$\beta_m = \text{ arg min}_{\alpha,\beta} \sum_{i=1}^{N} L(y_i, F_{m-1}(x) + \beta h(x_i; \alpha)) \tag{6}$$

The fifth step updates the model:

$$F_m(x) = F_{m-1}(x) + \beta_m h(x_i; \alpha) \tag{7}$$

Finally, the final classification algorithm $F_m(x)$ is produced. The residuals steadily decrease during the stepwise process, and the loss approaches approximately the minimum.

After training the optimal GBDT algorithm with hyperparameters tuned in the grid-search, this study used two global model-agnostic explainable AI (XAI) approaches [57,58] to answer the remaining two questions: (1) permutation-based feature importance (PBFI) and (2) partial dependence plot (PDP).

This study calculated PBFI, which is the relative magnitude of the influence of input features on prediction performance [52,59,60]. It compares all input features and ranks those that contribute to reducing overall variance [61]. This study chose PBFI over impurity-based feature importance (IBFI) for the following reasons. First, PBFI normalizes the biases of IBFI, such as the inflation of the values with many categories [62]. Second, the values of IBFI for certain input features may be high, regardless of limited contribution to the prediction of the target value. PBFI alleviates the limitation of IBFI.

Furthermore, this study developed PDP to capture the non-linear relationship between input features and an output target [63–66]. Specifically, PDP estimates the expected effects of a certain input feature on the outcome target after marginalizing the influences of other independent variables [66–68]. PDP visualizes PD with a line graph; specifically, the curve in PDP shows the average predicted effect of the input feature. In the line graph, the x-axis shows the values of the input features, and the y-axis shows the corresponding marginal effects [69].

4. Results

This section is divided into four subsections. The first subsection presents a general distribution of multimodality across the United States before moving on to the presentation of the two models. The following three subsections correspond to one of the three research questions using either the ordinary least square regression (OLS) or gradient boosting decision tree regressor models (GBDT).

4.1. How Did Multimodal Travel Behaviors Vary across the U.S.?

Figure 2 depicts the spatial distribution of the multimodality index for commute trips across the United States mainland. The figure visually reveals a spatial concentration of the Census Block Groups (CBGs) in metropolitan areas with a higher degree of multimodality, such as San Francisco, New York, and Chicago. Interestingly, GBGs in states in the Mountain area, such as Idaho, Nevada, Utah, Arizona, and New Mexico, demonstrated a relatively higher rate of multimodality. Figure 3 can confirm the findings, given the somewhat higher proportion of people who use alternate modes of transportation, such as walking. The mean and median of the multimodality index in the U.S. were around 0.56 and 0.53, with a minimum of 0 and a maximum of 1.6 (see Table 2).

4.2. Was Multimodality Associated with Built Environment Characteristics in the U.S.?

As shown in Table 3, the ordinary least square model (OLS) produced relatively well-fitting results (adjusted R-squared of 0.309). More importantly, most of the coefficients were statistically significant. Notable findings in the model concern nine built environment characteristics: after controlling for other variables, this study found that the built environment characteristics, except for job and household balance (Diversity_HH_job), were statistically significant predictors of the extent to which multimodality exists in the United States. In particular, the multimodality index was found to be significantly and positively linked with population density. Intriguingly, network density was inversely associated with multimodal travel behavior (estimate of -5.998), implying that adding additional lines on the roads may encourage the usage of private automobiles. In contrast, given that multimodality was found to be positively and significantly associated with intersection density, the ratio of four-way intersection density may be able to increase the likelihood that people will choose alternative modes of transportation, such as walking, which corroborates the arguments of Jacobs [70]. Furthermore, multimodal census block groups (CBGs) tended to have a higher walkability rating and be closer to places of employment. Additionally, whereas regional accessibility by automobile was found to be negatively connected to multimodality (estimate of −0.080), the accessibility by public transit showed a positive relationship (estimate of 0.146). Also crucial in having multimodality at CBGs across the U.S. were control variables such as household size, income level, race/ethnicity, educational achievement, and car ownership.

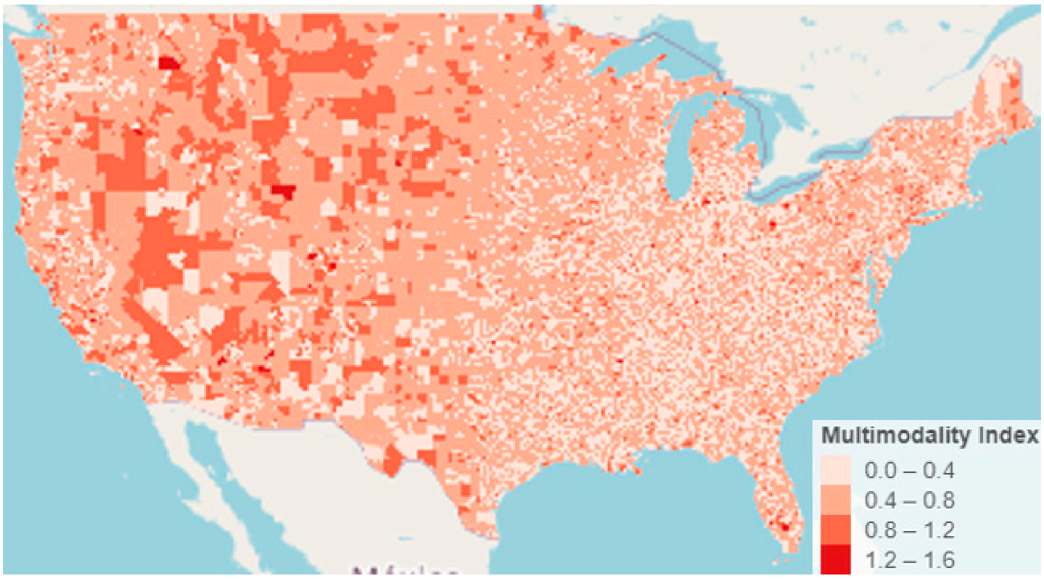

Figure 2. Spatial variation of the multimodality index in the United States mainland.

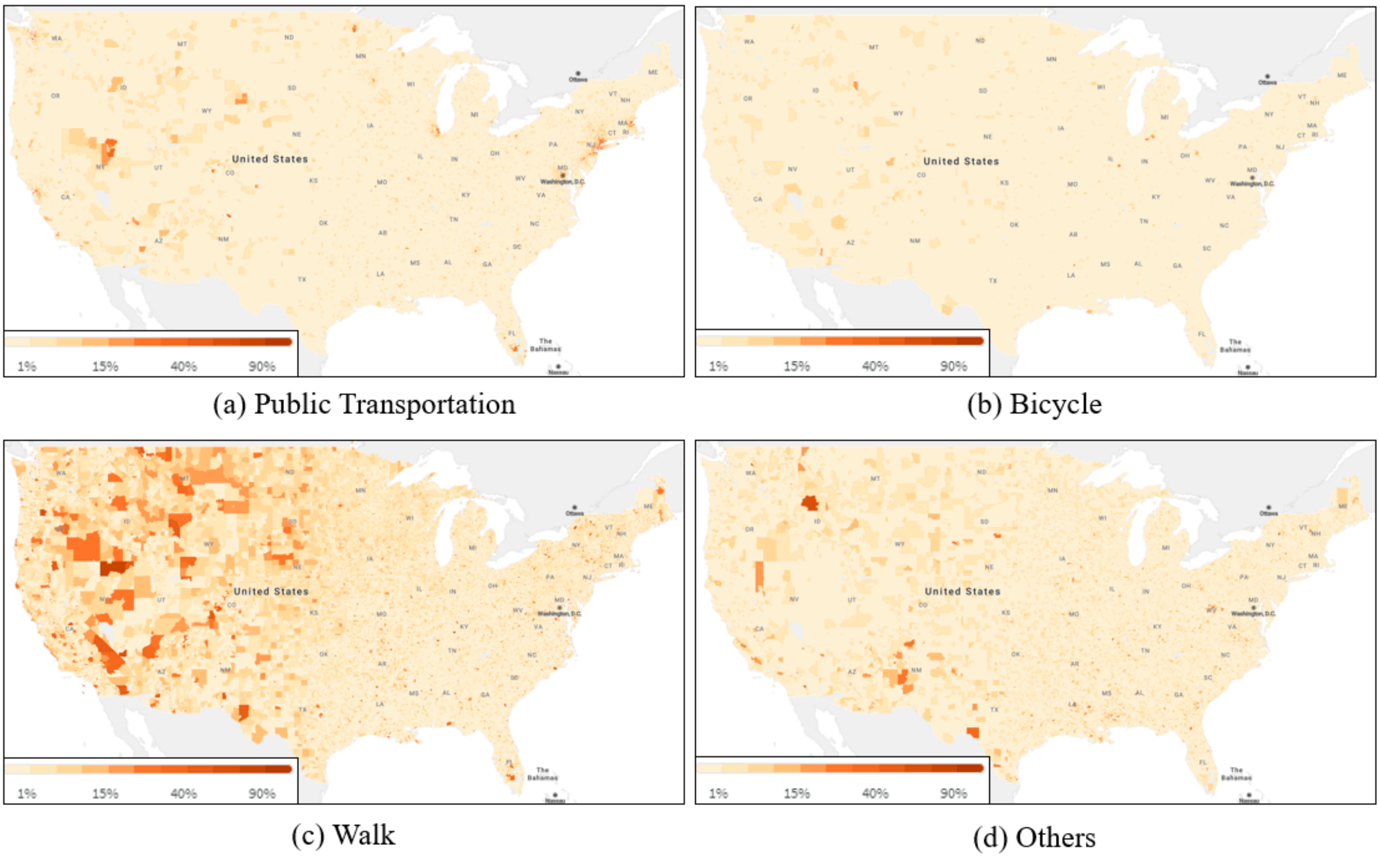

Figure 3. Shares of alternative transportation modes for commute trips in the United States mainland (Source: Social Explorer).

Table 3. Results of the ordinary least square model.

Variables	Estimate	Std. Error	*t*-Value	*p*-Value	VIF
Constant	0.359	0.007	48.780	<0.001	-
Pop_den	0.010	<0.001	24.190	<0.001	1.535
Diversity_HH_job	−0.082	0.063	−1.294	0.196	1.017
Diversity_job	−0.023	0.003	−7.851	<0.001	1.592
Net_den	−5.998	1.140	−5.262	<0.001	6.031
Int_den	2.223	0.123	18.091	<0.001	3.936
Walkability index	0.011	<0.001	45.797	<0.001	4.443
Job_proximity	0.014	<0.001	32.686	<0.001	1.180
Auto_accessibility	−0.080	0.002	−36.339	<0.001	1.559
Transit_accessibility	0.146	0.003	42.948	<0.001	1.935
HH_size	0.024	0.001	22.794	<0.001	1.606
HH_income	0.001	<0.001	3.826	<0.001	2.995
White	−0.018	<0.001	−37.876	<0.001	6.550
Black	−0.015	<0.001	−30.829	<0.001	4.905
Asian	0.013	0.001	17.815	<0.001	2.202
Single	0.037	<0.001	78.524	<0.001	1.857
Low education	−0.002	<0.001	−5.323	<0.001	2.960
No_car	0.025	<0.001	81.355	<0.001	1.410
Work_at_home	0.048	0.001	49.115	<0.001	1.180
		Model Statistics			
Observations			206,380		
R^2			0.309		
Adjusted R^2			0.309		

4.3. To What Extent Did the Built Environment Characteristics Play a Role in Showing Multimodal Travel Patterns?

The trained optimal gradient boosting decision tree regressor model (GBDT) found in the grid-search produced the R-squared of 0.410, explained variance of 0.411, and negative mean absolute error of −0.162 in the 10-fold cross-validation. Table 4 presents the permutation-based feature importance (PBFI) analysis findings based on the trained GBDT, which measures the relative contribution of factors within the total contribution of 100%. The nine factors regarding the built environment were important in showing multimodality at CBGs in the United States, with 40.8% of the feature importance. However, not all built environment elements were the input features of outstandingly high value to predict the extent of multimodality. Particularly, population density, regional accessibility, walkability index, and network density accounted for approximately 35.1% of the total importance of all independent variables, and interestingly, the population density was placed second with a significance of 15.9%, which was much greater than the relevance of the other built environment elements. Moreover, control variables such as race/ethnicity, marital status, and car ownership scored significantly higher than the other variables; mainly, non-Hispanic white and black proportion was the dominant factor with a 30.3% contribution.

4.4. Were There Non-Linear Effects of Built Environment Factors on Multimodality?

This subsection presents partial dependence plots in Figures 4 and 5, where independent variables and their corresponding marginal effects on predicted probability are marked on the x-axis and y-axis while accounting for the average influences of all other variables in the trained GBDT [71]. This subsection illustrates the non-linear effects of independent variables on multimodality that traditional linear regression in Table 3 cannot fully capture.

Table 4. Results of the feature importance in the gradient boosting decision tree regressor.

Variables	Impurity-Based Feature Importance		Permutation-Based Feature Importance	
	Importance	Rank	Importance	Rank
	Built environment characteristics			
Pop_den	0.206	1	0.159	2
Diversity_HH_job	0.019	15	0.010	17
Diversity_job	0.011	18	0.004	18
Net_den	0.036	7	0.034	9
Int_den	0.018	16	0.015	16
Walkability index	0.117	3	0.051	7
Job_proximity	0.033	8	0.028	10
Auto_accessibility	0.026	11	0.048	8
Transit_accessibility	0.108	4	0.059	6
	Neighborhood characteristics			
HH_size	0.024	13	0.022	12
HH_income	0.031	9	0.020	15
White	0.043	6	0.178	1
Black	0.018	16	0.125	3
Asian	0.023	14	0.021	13
Single	0.101	5	0.061	5
Low education	0.025	12	0.021	13
No_car	0.133	2	0.116	4
Work_at_home	0.029	10	0.027	11

Note: This table included impurity-based feature importance for comparison purposes.

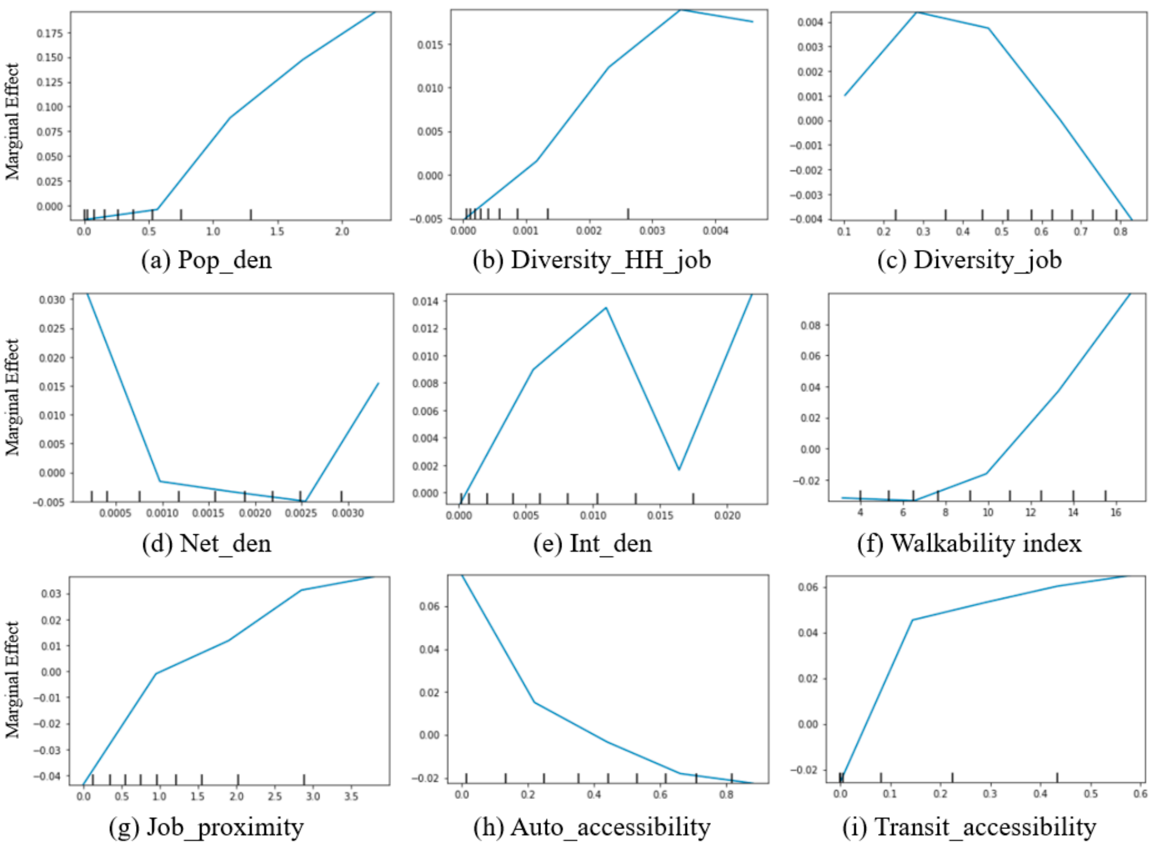

Figure 4. Non-linear effects of built environment characteristics on multimodality index.

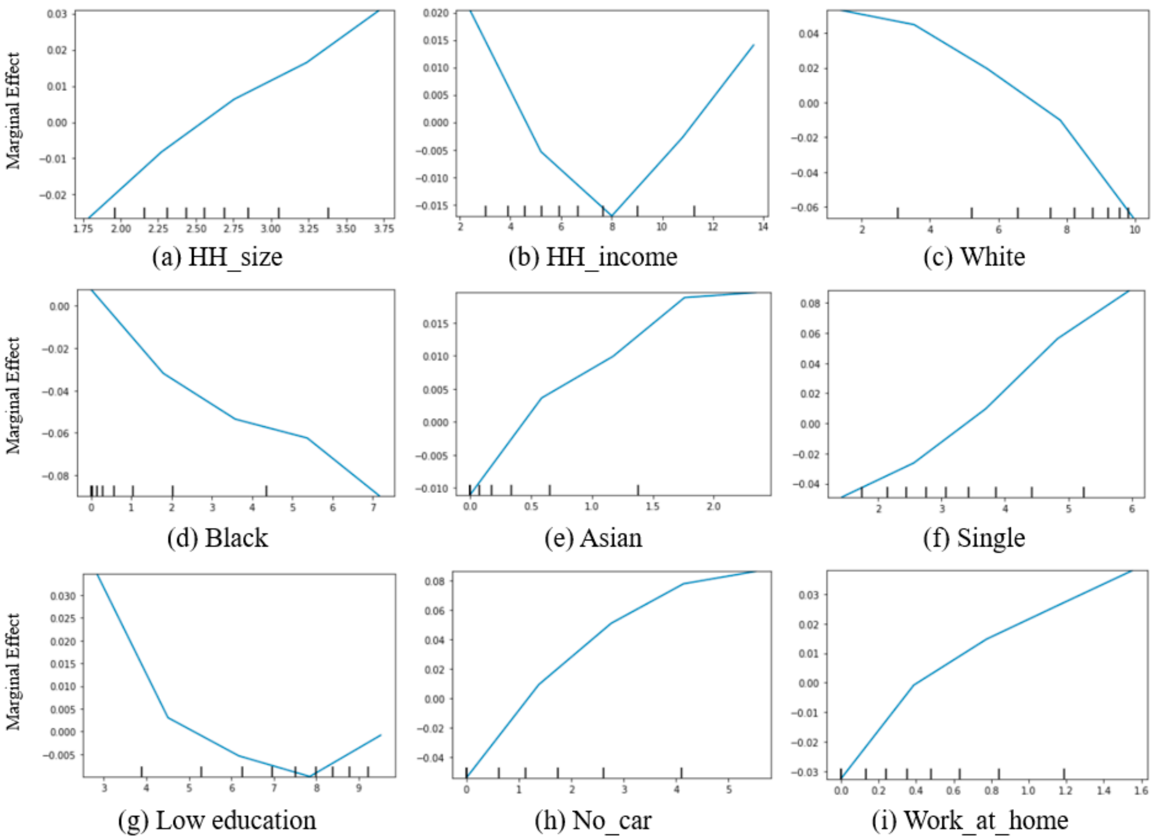

Figure 5. Non-linear effects of neighborhood characteristics on multimodality index.

Figure 4a demonstrates that the influence of population density, which was found to be the most influential factor among built environment characteristics in Table 4 on multimodality index, was relatively stable roughly until 50,000 persons per acre, while it exerted significant and positive effects beyond the range. A similar pattern was observed in the walkability index in Figure 4f; specifically, a walkability index of more than 10 resulted in a considerable increase in the multimodality index. Figure 4i indicates that the increase in regional accessibility by transit rapidly increased the multimodality index to 0.1, although it had marginal effects thereafter. In the United States, regional accessibility by automobile had a relatively linear negative influence on multimodality, like the findings in Table 3. CBGs with medium or higher ranges of intersection density, as shown in Figure 4e, had a greater multimodality index than CBGs with a lower range of intersection density.

In addition, control factors such as household size, race/ethnicity, marital status, and car ownership, as shown in Figure 5, had roughly linear impacts on multimodality, like the OLS model. For example, CBGs with a lower proportion of non-Hispanic white and black residents had a lower multimodality index than others. The influence of household income and educational attainment on multimodality was not monotonous, in which the multimodality index decreased when the two factors ranged from lower values and thereafter increased with higher values (see Figure 5b,g). Interestingly, as individuals in the U.S. worked at home, multimodality at CBGs increased.

5. Conclusions and Discussion

Transportation mode choice behavior is intertwined with diverse aspects of our lives, such as employment, housing, schools, shopping, and health [8]. Since the automobility system has expanded and established itself as the dominant mode of transportation, it has pushed alternative transportation modes, such as public transit and walking, to the sidelines [9]. As a result, there have been several consequences, such as obesity, traffic congestion, environmental pollution, urban sprawl, and social marginalization [13]. Thus, planners are attempting to make a transition away from automobile use and toward more multimodal and sustainable forms of transportation [9,15,16] by using strategies mainly at three levels, the vehicle level, the transportation system (e.g., infrastructure), and the level of traffic participants (e.g., vehicle owners) [7]. However, since the significance and role of the built environment have not been sufficiently considered, they are currently facing difficulty in their efforts to fulfill their responsibilities to plan and deliver comprehensive, efficient, high-performing, multimodal transportation networks that are in line with the goals of the community. Therefore, this study demonstrated a link between multimodality and built environment characteristics in the U.S. and proposed policy implications for promoting multimodal travel behaviors using ordinary least square regression and gradient boosting decision tree regressor models.

Several findings of the U.S. nationwide analysis deserve further discussion. First, certain built environment characteristics were predictors of multimodality at census block groups (CBGs) throughout the U.S., with statistical significance and a relatively higher contribution. This suggests that planners who would like to encourage multimodal travel behavior should consider the features, particularly population density, regional accessibility, walkability index, and network density, when developing their land-use design strategies for the transportation system. Second, the salient non-linear effects of built environment characteristics on multimodality suggest an effective range to encourage multimodal transportation mode choice behaviors in various situations in the U.S. For example, the effective population density exceeded 50,000 people per acre in the U.S. Furthermore, even with a little improvement in regional accessibility by transit, individuals may significantly alter their travel behavior toward a more sustainable manner. However, a considerable increase in walkability is needed to promote multimodal travel behaviors meaningfully. Additionally, considering that regional accessibility via automobile demonstrated a significant disincentive toward multimodal travel behavior across all ranges, it is possible that infrastructure development and improvement, as well as strategies to improve the level of services provided for automobiles, may have a negative impact on encouraging individuals to use a variety of modes of transportation.

In sum, the findings of this study can be applied to a problem that decision-makers are currently facing: how to encourage multimodal travel behavior while also providing practical advice for developing appropriate plans. In addition, the integration of these findings and a variety of mobility management strategies, such as reforming price structures for transportation, may have a synergistic impact on the promotion of non-automobile travel behaviors. These considerations highlight the practical necessity of understanding the circumstances under which individuals increase their use of diverse modes of transportation in their lives.

This study acknowledges several limitations. For instance, since this study used census block groups (CBGs) as the unit of analysis, findings may face the so-called ecological fallacy that transfers relationships between covariates at aggregate scales to individuals [72]. Additionally, the first model did not control for spatial dependence, which may produce biased and inefficient estimates of covariates [73]. Moreover, an omitted variable bias may be present in the final results since this study did not include a comprehensive set of built environment features identified in previous studies [74]. Additionally, the researcher did not undertake a longitudinal analysis to draw causal inferences from the findings. The travel patterns of non-commute trips were not investigated in this study. Probably, the results may not be transferable to other countries because their travel behaviors differ significantly from those of the United States.

Beyond this study, future research needs to broaden the scope of this research by investigating whether favorable built environment characteristics aid multimodality in maintaining or even increasing their previous extent after the COVID-19 outbreak within the framework of resilience [75]. Additionally, further studies are needed to explore how the findings of this study are applied at the not national but local level with some case studies.

Funding: This research received no external funding.

Data Availability Statement: The two data sets used in this study are publicly available in websites: (1) https://www.census.gov/data/developers/data-sets/acs-5year.html (accessed on 10 December 2020) and (2) https://www.epa.gov/smartgrowth/smart-location-mapping (accessed on 1 June 2021).

Conflicts of Interest: The author declares no conflict of interest.

References

1. Adams, G.; Gerard, D. Smart growth and transportation: Opportunities and challenges for Austin. *Inst. Transp. Eng. ITE J.* **2000**, *70*, 30–34.
2. Altmann, A.; Toloşi, L.; Sander, O.; Lengauer, T. Permutation importance: A corrected feature importance measure. *Bioinformatics* **2010**, *26*, 1340–1347. [CrossRef] [PubMed]
3. Apley, D.W.; Zhu, J. Visualizing the effects of predictor variables in black box supervised learning models. *J. R. Stat. Soc. Ser. B Statistical Methodol.* **2020**, *82*, 1059–1086. [CrossRef]
4. Astroza, S.; Garikapati, V.M.; Bhat, C.; Pendyala, R.M.; Lavieri, P.S.; Dias, F.F. Analysis of the Impact of Technology Use on Multimodality and Activity Travel Characteristics. *Transp. Res. Rec. J. Transp. Res. Board* **2017**, *2666*, 19–28. [CrossRef]
5. Beckmann, J. Automobility—A Social Problem and Theoretical Concept. *Environ. Plan. D Soc. Space* **2001**, *19*, 593–607. [CrossRef]
6. Beenackers, M.A.; Foster, S.; Kamphuis, C.B.; Titze, S.; Divitini, M.; Knuiman, M.; van Lenthe, F.J.; Giles-Corti, B. Taking Up Cycling After Residential Relocation. *Am. J. Prev. Med.* **2012**, *42*, 610–615. [CrossRef]
7. Bhatt, U.; Xiang, A.; Sharma, S.; Weller, A.; Taly, A.; Jia, Y.; Ghosh, J.; Puri, R.; Moura, J.M.F.; Eckersley, P. Explainable machine learning in deployment. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, Barcelona, Spain, 27–30 January 2020. [CrossRef]
8. Buehler, R.; Hamre, A. The multimodal majority? Driving, walking, cycling, and public transportation use among American adults. *Transportation* **2014**, *42*, 1081–1101. [CrossRef]
9. Bureau, U.C. Census Bureau Announces Changes for 2020 American Community Survey 1-Year Estimates. Census.Gov. 2021. Available online: https://www.census.gov/newsroom/press-releases/2021/changes-2020-acs-1-year.html (accessed on 29 July 2021).
10. Caspi, O.; Noland, R.B. Bikesharing in Philadelphia: Do lower-income areas generate trips? *Travel Behav. Soc.* **2019**, *16*, 143–152. [CrossRef]
11. Chapman, J.; Fox, E.H.; Bachman, W.; Frank, L.D. *Smart Location Database Technical Documentation and User Guide*; The United States Environment Protection Agency: Washington, DC, USA, 2021. Available online: https://www.epa.gov/smartgrowth/smart-location-database-technical-documentation-and-user-guide (accessed on 1 June 2021).
12. Christopherson, S.; Michie, J.; Tyler, P. Regional resilience: Theoretical and empirical perspectives. *Camb. J. Reg. Econ. Soc.* **2010**, *3*, 3–10. [CrossRef]
13. Conley, J. *Car Troubles: Critical Studies of Automobility and Auto-Mobility*; Routledge: London, UK, 2016.
14. Diana, M.; Pirra, M. A comparative assessment of synthetic indices to measure multimodality behaviours. *Transp. A Transp. Sci.* **2016**, *12*, 771–793. [CrossRef]
15. Ding, C.; Cao, X.; Næss, P. Applying gradient boosting decision trees to examine non-linear effects of the built environment on driving distance in Oslo. *Transp. Res. Part A Policy Pract.* **2018**, *110*, 107–117. [CrossRef]
16. Dong, G.; Harris, R. Spatial Autoregressive Models for Geographically Hierarchical Data Structures. *Geogr. Anal.* **2014**, *47*, 173–191. [CrossRef]
17. Duany, A.; Plater-Zyberk, E.; Speck, J. *Suburban Nation: The Rise of Sprawl and the Decline of the American Dream*; Macmillan: New York, NY, USA, 2001.
18. El-Assi, W.; Mahmoud, M.S.; Habib, K.N. Effects of built environment and weather on bike sharing demand: A station level analysis of commercial bike sharing in Toronto. *Transportation* **2015**, *44*, 589–613. [CrossRef]
19. Ewing, R.; Cervero, R. Travel and the Built Environment: A meta-analysis. *J. Am. Plan. Assoc.* **2010**, *76*, 265–294. [CrossRef]
20. Ewing, R.; Cervero, R. "Does Compact Development Make People Drive Less?" The Answer Is Yes. *J. Am. Plan. Assoc.* **2017**, *83*, 19–25. [CrossRef]
21. Ewing, R.; Tian, G.; Goates, J.P.; Zhang, M.; Greenwald, M.J.; Joyce, A.; Kircher, J.C.; Greene, W.H. Varying influences of the built environment on household travel in 15 diverse regions of the United States. *Urban Stud.* **2014**, *52*, 2330–2348. [CrossRef]
22. Fishman, E. Bikeshare: A Review of Recent Literature. *Transp. Rev.* **2015**, *36*, 92–113. [CrossRef]
23. Friedman, J.H. Multivariate Adaptive Regression Splines. *Ann. Stat.* **1991**, *19*, 1–67. [CrossRef]
24. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
25. Geels, F.W. A socio-technical analysis of low-carbon transitions: Introducing the multi-level perspective into transport studies. *J. Transp. Geogr.* **2012**, *24*, 471–482. [CrossRef]
26. Gehrke, S.R.; Wang, L. Operationalizing the neighborhood effects of the built environment on travel behavior. *J. Transp. Geogr.* **2019**, *82*, 102561. [CrossRef]

27. Ha, J.; Lee, S.; Kwon, S.M. Revisiting the Relationship between Urban Form and Excess Commuting in US Metropolitan Areas. *J. Plan. Educ. Res.* **2018**, *41*, 294–311. [CrossRef]
28. Hagenauer, J.; Helbich, M. A comparative study of machine learning classifiers for modeling travel mode choice. *Expert Syst. Appl.* **2017**, *78*, 273–282. [CrossRef]
29. Handy, S.L.; Boarnet, M.G.; Ewing, R.; Killingsworth, R.E. How the built environment affects physical activity: Views from urban planning. *Am. J. Prev. Med.* **2002**, *23* (Suppl. S2), 64–73. [CrossRef]
30. Hastie, T.; Friedman, J.; Tibshirani, R. (Eds.) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: Berlin/Heidelberg, Germany, 2001; pp. 299–345. [CrossRef]
31. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2016.
32. Heinen, E.; Mattioli, G. Does a high level of multimodality mean less car use? An exploration of multimodality trends in England. *Transportation* **2017**, *46*, 1093–1126. [CrossRef]
33. Huang, N.; Lu, G.; Xu, D. A Permutation Importance-Based Feature Selection Method for Short-Term Electricity Load Forecasting Using Random Forest. *Energies* **2016**, *9*, 767. [CrossRef]
34. Hunecke, M.; Groth, S.; Wittowsky, D. Young social milieus and multimodality: Interrelations of travel behaviours and psychographic characteristics. *Mobilities* **2020**, *15*, 397–415. [CrossRef]
35. Hunter, R.F.; Garcia, L.; de Sa, T.H.; Zapata-Diomedi, B.; Millett, C.; Woodcock, J.; Pentland, A.; Moro, E. Effect of COVID-19 response policies on walking behavior in US cities. *Nat. Commun.* **2021**, *12*, 3652. [CrossRef]
36. Jacobs, J. *The Death and Life of Great American Cities*; Random House: New York, NY, USA, 1961; Available online: https://www.goodreads.com/work/best_book/1289564-the-death-and-life-of-great-american-cities (accessed on 28 April 2022).
37. Jacquemin, A.P.; Berry, C.H. Entropy Measure of Diversification and Corporate Growth. *J. Ind. Econ.* **1979**, *27*, 359. [CrossRef]
38. Jeon, C.M.; Amekudzi, A.A.; Guensler, R.L. Evaluating Plan Alternatives for Transportation System Sustainability: Atlanta Metropolitan Region. *Int. J. Sustain. Transp.* **2010**, *4*, 227–247. [CrossRef]
39. Jost, L. Entropy and diversity. *Oikos* **2006**, *113*, 363–375. [CrossRef]
40. Kennedy, D.; Bates, R.R.; Watson, A.Y. (Eds.) *Air Pollution, the Automobile, and Public Health*, 1st ed.; National Academies Press: Washington, DC, USA, 1988.
41. Kim, J.; Kwan, M.-P. The impact of the COVID-19 pandemic on people's mobility: A longitudinal study of the U.S. from March to September of 2020. *J. Transp. Geogr.* **2021**, *93*, 103039. [CrossRef]
42. King, D.M.; Jacobson, S. What Is Driving Obesity? A Review on the Connections Between Obesity and Motorized Transportation. *Curr. Obes. Rep.* **2017**, *6*, 3–9. [CrossRef] [PubMed]
43. Kroesen, M. Modeling the behavioral determinants of travel behavior: An application of latent transition analysis. *Transp. Res. Part A Policy Prac.* **2014**, *65*, 56–67. [CrossRef]
44. Kuhnimhof, T.; Buehler, R.; Wirtz, M.; Kalinowska, D. Travel trends among young adults in Germany: Increasing multimodality and declining car use for men. *J. Transp. Geogr.* **2012**, *24*, 443–450. [CrossRef]
45. Lambert, J.H.; Joshi, N.N.; Peterson, K.D.; Wadie, S.M. Coordination and Diversification of Investments in Multimodal Transportation. *Public Work. Manag. Policy* **2007**, *11*, 250–265. [CrossRef]
46. Lee, C.; Miller, J.S. A probability-based indicator for measuring the degree of multimodality in transportation investments. *Transp. Res. Part A Policy Prac.* **2017**, *103*, 377–390. [CrossRef]
47. Lee, D.; Mulrow, J.; Haboucha, C.J.; Derrible, S.; Shiftan, Y. Attitudes on Autonomous Vehicle Adoption using Interpretable Gradient Boosting Machine. *Transp. Res. Rec. J. Transp. Res. Board* **2019**, *2673*, 865–878. [CrossRef]
48. Litman, T. *Introduction to Multi-Modal Transportation Planning: Principles and Practices*; Victoria Transport Policy Institute: Victoria, BC, Canada, 2021; Available online: https://vtpi.org/multimodal_planning.pdf (accessed on 23 April 2021).
49. Lucas, K. Transport and social exclusion: Where are we now? *Transp. Policy* **2012**, *20*, 105–113. [CrossRef]
50. Lyons, T.; Ewing, R. Does transit moderate spatial mismatch? The effects of transit and compactness on regional economic outcomes. *Cities* **2021**, *113*, 103160. [CrossRef]
51. Malik, J.; Bunch, D.S.; Handy, S.; Circella, G. A deeper investigation into the effect of the built environment on the use of ridehailing for non-work travel. *J. Transp. Geogr.* **2021**, *91*, 102952. [CrossRef]
52. McCoy, K.; Andrew, J.; Glynn, R.; Lyons, W.; John, A.; Volpe National Transportation Systems Center (U.S.). Integrating Shared Mobility into Multimodal Transportation Planning: Improving Regional Performance to Meet Public Goals (DOT-VNTSC-FHWA-18-13; FHWA-HEP-18-033). 2018. Available online: https://rosap.ntl.bts.gov/view/dot/34689 (accessed on 1 February 2018).
53. McLaren, A.T. Families and transportation: Moving towards multimodality and altermobility? *J. Transp. Geogr.* **2016**, *51*, 218–225. [CrossRef]
54. Molnar, C. Interpretable Machine Learning: A Guide for Making Black Box Models Explainable. 2021. Available online: https://christophm.github.io/interpretable-ml-book/shapley.html (accessed on 29 March 2022).
55. Molnar, C.; Freiesleben, T.; König, G.; Casalicchio, G.; Wright, M.N.; Bischl, B. Relating the Partial Dependence Plot and Permutation Feature Importance to the Data Generating Process. *arXiv* **2021**, arXiv:2109.01433.
56. Muhs, C.D.; Clifton, K. Do characteristics of walkable environments support bicycling? Toward a definition of bicycle-supported development. *J. Transp. Land Use* **2015**, *9*, 147–188. [CrossRef]

57. Nobis, C. Multimodality. *Transp. Res. Rec. J. Transp. Res. Board* **2007**, *2010*, 35–44. [CrossRef]
58. Noland, R.B.; Smart, M.; Guo, Z. Bikeshare trip generation in New York City. *Transp. Res. Part A Policy Prac.* **2016**, *94*, 164–181. [CrossRef]
59. Parker, M.E.; Li, M.; Bouzaghrane, M.A.; Obeid, H.; Hayes, D.; Frick, K.T.; Rodríguez, D.A.; Sengupta, R.; Walker, J.; Chatman, D.G. Public transit use in the United States in the era of COVID-19: Transit riders' travel behavior in the COVID-19 impact and recovery period. *Transp. Policy* **2021**, *111*, 53–62. [CrossRef]
60. Rao, H.; Shi, X.; Rodrigue, A.K.; Feng, J.; Xia, Y.; Elhoseny, M.; Yuan, X.; Gu, L. Feature selection based on artificial bee colony and gradient boosting decision tree. *Appl. Soft Comput.* **2018**, *74*, 634–642. [CrossRef]
61. Sabouri, S.; Park, K.; Smith, A.; Tian, G.; Ewing, R. Exploring the influence of built environment on Uber demand. *Transp. Res. Part D Transp. Environ.* **2020**, *81*, 102296. [CrossRef]
62. Sabouri, S.; Tian, G.; Ewing, R.; Park, K.; Greene, W. The built environment and vehicle ownership modeling: Evidence from 32 diverse regions in the U.S. *J. Transp. Geogr.* **2021**, *93*, 103073. [CrossRef]
63. Saha, D.; Alluri, P.; Gan, A. Prioritizing Highway Safety Manual's crash prediction variables using boosted regression trees. *Accid. Anal. Prev.* **2015**, *79*, 133–144. [CrossRef] [PubMed]
64. Scheiner, J.; Chatterjee, K.; Heinen, E. Key events and multimodality: A life course approach. *Transp. Res. Part A Policy Prac.* **2016**, *91*, 148–165. [CrossRef]
65. Shamshiripour, A.; Rahimi, E.; Shabanpour, R.; Mohammadian, A. How is COVID-19 reshaping activity-travel behavior? Evidence from a comprehensive survey in Chicago. *Transp. Res. Interdiscip. Perspect.* **2020**, *7*, 100216. [CrossRef] [PubMed]
66. Strobl, C.; Boulesteix, A.-L.; Zeileis, A.; Hothorn, T. Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinform.* **2007**, *8*, 25. [CrossRef] [PubMed]
67. Subramanian, S.V.; Jones, K.; Duncan, C. Multilevel Methods for Public Health Research. In *Neighborhoods and Health*; Kawachi, I., Berkman, L.F., Eds.; Oxford University Press: New York, NY, USA, 2003; pp. 65–111. [CrossRef]
68. Talmage, C.A.; Frederick, C. Quality of Life, Multimodality, and the Demise of the Autocentric Metropolis: A Multivariate Analysis of 148 Mid-Size U.S. Cities. *Soc. Indic. Res.* **2018**, *141*, 365–390. [CrossRef]
69. Urry, J. Mobile sociology1. *Br. J. Sociol.* **2000**, *51*, 185–203. [CrossRef]
70. U.S. Census Bureau. *American Community Survey Accuracy of the Data (2020)*; U.S. Census Bureau: Suitland, MA, USA, 2021; p. 25. Available online: https://www2.census.gov/programs-surveys/acs/tech_docs/accuracy/ACS_Accuracy_of_Data_2020.pdf (accessed on 17 March 2022).
71. Wells, E.M.; Small, M.; Spurlock, C.A.; Wong-Parodi, G. Factors associated with emerging multimodal transportation behavior in the San Francisco Bay Area. *Environ. Res. Infrastruct. Sustain.* **2021**, *1*, 031004. [CrossRef]
72. Xi, Y.; Zhuang, X.; Wang, X.; Nie, R.; Zhao, G. A Research and Application Based on Gradient Boosting Decision Tree. In *Web Information Systems and Applications*; Meng, X., Li, R., Wang, K., Niu, B., Wang, X., Zhao, G., Eds.; Springer: Cham, Switzerland, 2018; pp. 15–26. [CrossRef]
73. Yuan, X.; Abouelenien, M. A multi-class boosting method for learning from imbalanced data. *Int. J. Granul. Comput. Rough Sets Intell. Syst.* **2015**, *4*, 13. [CrossRef]
74. Zhao, P. The Impact of the Built Environment on Bicycle Commuting: Evidence from Beijing. *Urban Stud.* **2013**, *51*, 1019–1037. [CrossRef]
75. Zhou, Z.-H. *Ensemble Methods: Foundations and Algorithms*, 1st ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2012.