Income moderates the nonlinear influence of built environment attributes on travel-related carbon emissions[1]

Tao Tao, University of Minnesota

Haotian Zhong, New York University Shanghai, Renmin University

Abstract

Policymakers have adopted built environment policies to modify people's travel behavior and the related emissions. However, few studies have examined the interactive impact between income level and built environment attributes on travel-related carbon emissions (TCE), and only several studies consider their nonlinear relationships. With data from the Twin Cities, US, this study estimated the nonlinear effects of built environment attributes and demographics on TCE. It further examined the interactive impacts between household income and built environment attributes. The findings highlight that demographics exert a greater influence on TCE than the built environment. Employment status, job accessibility, and gender are the most important predictors. Besides individual nonlinear relationships, household income and built environment attributes have salient interactive impacts on TCE. The results suggest that providing environment friendly and affordable transportation choices to low-income population, switching to clean energy vehicles, and offering more matched job opportunities to low-income population near their residence are promising to create a sustainable transportation system.

Keywords: carbon emission reduction; environment-friendly transportation system; travel behavior; land use; machine learning

1 Introduction

Greenhouse gas (GHG) emission continues to increase and negatively influence the climate (Global Carbon Bedget, 2023). GHG generated from transportation accounts for a large share of the total GHG emissions. In the US, the transportation sector produces approximately 27% of the GHG emissions in 2020 (EPA, 2022). Therefore, it is important to discourage people's dependence on auto use to reduce GHG emissions and their influence on the environment. A considerable number of scholars in the field of urban planning are extensively researching the impact of policies related to the built environment on shaping people's travel behavior and the consequent GHG emissions (Cao & Yang, 2017; Choi & Zhang, 2017; Wu et al., 2019).

However, there are two research gaps in the existing literature. First, most studies only focus on the entire population when studying the built environment effects on travel-related emissions but fall short of examining the heterogenous effects across the subpopulations with different income levels in the region. As people from different income groups differ in their travel behavior and respond differently to the built environment policies, results of the whole population might not be applicable to those low-income and/or high-income groups. This creates equity concerns. Second, the nonlinear relationships between the built environment and travel-related emissions have not received sufficient attention from scholars and practitioners yet. More studies considering nonlinear relationships are still needed to supplement the understanding of the complex relationships between built environment attributes and travel-related emissions.

To address these research gaps, this study applies a machine learning approach called Extreme Gradient Boost (XGBoost) to estimate the contributions of built environment attributes and demographic variables on people's daily travel-related carbon emissions (TCE) and the related nonlinear relationships with the data from the Twin Cities area in US. In addition, this study examines the interactive effects between household income and built environment attributes on TCE. Specifically, this study plans to answer two research questions:

1) How do built environment attributes and demographics influence TCE when considering nonlinear relationships?

2) What are the interactive impacts between household income and built environment attributes on TCE?

This study makes two significant contributions to the field. Firstly, it delves into the individual nonlinear effects of the built environment and demographic factors on travel-related carbon emissions, thereby enriching the limited research available in this area. Secondly, it investigates the interplay between household income and built environment attributes in influencing travel-related carbon emissions. The findings from this exploration offer valuable insights for formulating policies aimed at developing a sustainable and equitable transportation system.

The rest of this paper is organized as follows. Section 2 reviews the literature on the impact of built environment attributes on travel-related emissions. The data and methodologies used in this study are introduced in Section 3. Section 4 presents and discusses the results. The conclusions and policy implications are provided in the last section.

2 Literature review

The impacts of built environment attributes on driving have been extensively studied in the literature. Cervero and Kockelman (1997) found that "3Ds", including density, diversity, and design, could reduce vehicle trips. Furthermore, Ewing and Cervero (2010) proposed "5Ds", including density, diversity, design, distance to transit, and destination accessibility, and reviewed how these attributes could affect travel. Their review suggested that destination accessibility and street network design are the two variables most correlated with vehicle miles travelled (VMT). Many studies on the built environment and driving chose to follow this rule and included attributes from "5Ds" in their models (Stevens, 2017).

Recent studies have shifted their focus from the impact of the built environment on driving to its direct impact on travel-related emissions. This change in focus stems from the understanding that driving alone does not determine the total amount of travel-related emissions. These emissions are influenced by multiple travel modes, including cars and transit, and are determined by the share, distance, and emission rate of each mode. Travel-related emissions can be studied at both disaggregated levels (e.g., person or household level (Shao et al., 2023; Wu et al., 2019)) and aggregated levels (e.g., traffic analysis zone or census block group level (Boarnet et al., 2017; Credit & Lehnert, 2023; Feng et al., 2022)). Studies found that "5Ds" are important built environment attributes correlated with travel-related emissions. Specifically, their results show that density (Cao & Yang, 2017; Choi & Zhang, 2017), land use mix (Choi & Zhang, 2017; Wang et al., 2013), and street network connectivity (Wang et al., 2013; Xu et al., 2018) are all negatively correlated with travel-related emissions. On the contrary, distance to city center (Cao & Yang, 2017; Wang et al., 2017) and distance to transit stop (Barla et al., 2011; Boarnet et al., 2017) have positive relationships with travel-related emissions.

Studies on built environment and travel-related emissions applied two major types of approaches to estimate the impacts, including statistical and machine learning models. Statistical models usually assume the variables follow specific probability distributions, such as Gaussian and Poisson distributions (i.e., linear and Poisson regression). For instance, Barla et al. (2011) employed a linear regression to explore the correlations between built environment attributes and personal average daily travel-related GHG emissions, controlling for respondents' demographics, in Quebec, Canada. Machine learning models, instead, use advanced approaches, such as vector support machines and decision trees, to analyze impacts. Compared to statistical models, machine learning models make limited assumptions about the variable distributions and offer better predictive capabilities. For example, Shao et al. (2023) used the gradient boosting decision tree approach to estimate the relationships between the built environment and personal average daily travel-related carbon emissions in Zhongshan, China.

Spatial dependence has also been considered in studies on travel-related emissions aggregated at specific area unit. Spatial dependence indicates the phenomenon where observations located in close geographical proximity are more likely to be similar to each other than to those further apart. For example, Feng et al. (2022) used the geographically weighted regression model to explore the spatial impact of built environment attributes on vehicle emissions at the traffic analysis zone level in Holingola, China. Another advanced example is that Credit and Lehnert (2023) employed a causal machine learning approach to estimate the effect of light rail on travel-related carbon emissions at the block group level in Maricopa, Arizona. With a difference-in-difference research design, they applied the causal forest model to estimate the effect while controlling for spatial dependence.

The literature, however, has seldom examined the heterogeneous built environment effects on travel-related emissions for people with different income levels. Many studies have shown that people from different income groups behave differently in their travel behavior. According to the 2017 National Household Travel Survey, households with incomes over 100,000 dollars on average have over 4,000 trips every year. For households with incomes less than 15,000 dollars, the number of annual person trips drops to approximately 1,500 (FHWA, 2018; Wang & Renne, 2023). In addition, the low-income population responds differently to the built environment policies compared with the higher-income population. This is because low-income people have limited choices in their travel modes and work opportunities (Blumenberg, 2017). For example, with the travel data from the Twin Cities, USA, Tao and Cao (2021) found that people living in high-income areas only reduce their vehicle usage when transit supply increases to a certain threshold. However, people living in low-income areas keep reducing their driving amount when transit supply increases and do not have a threshold. As travel-related emissions are closely related to people's travel behavior, the effects of built environment attributes on travel-related emissions are expected to be different for people with various income levels and, thus, need more studies.

Moreover, there is a growing consensus among scholars that the associations between attributes of the built environment and travel behavior should be defined as nonlinear relationships (Boarnet, 2017; Van Wee & Handy, 2016). Nonlinear relationships indicate the incremental impact of a single built environment variable on travel behavior does not remain static, but rather, it varies based on the specific value of that built environment variable. Machine learning approaches have been widely adopted in recent literature due to their flexibility in estimating nonlinear relationships. Numerous empirical investigations have substantiated the existence of nonlinear relationships between the built environment and travel behavior (Ding et al., 2018; Sabouri et al., 2020; Wang & Ozbilen, 2020). For example, Ding et al. (2018) applied the gradient boosting decision tree approach to explore the relationships between built environment attributes and weekly driving distance in Oslo, Norway. They found that distance to city center is nonlinearly associated with weekly driving distance. The relationship showed that distance to city center has a moderately positive effect on weekly driving distance when it changes from 0 to 12 km, but this effect increases dramatically when distance to city center exceeds 12 km. Hence, it is plausible to anticipate that the relationships between attributes of the built environment and travel-related emissions may exhibit a nonlinear pattern. Only a limited number of studies have explored the nonlinear relationships between the built environment and travel-related emissions (Gao et al., 2022; Shao et al., 2023; Wu et al., 2019; Yang & Zhou, 2020) and none of them have examined the interactive impact between built environment attributes and demographics. More studies are still needed to examine the generalizability of the results of existing studies.

3 Data and methods

3.1 Data

One of the important datasets applied in this study is the Travel Behavior Inventory (TBI) survey in Twin Cities (Figure 1), in the US (Metropolitan Council, 2019). The Metropolitan Council, which is the local metropolitan planning organization, carried out the TBI survey from October 2018 to September 2019. The survey mainly collected information about people's demographic attributes and travel diaries. To increase the participation of the survey, TBI provided three approaches for people to take part in the survey, including smartphone application, online survey, and call center. All people used the same questionnaire no matter how they participated in the survey. The survey was also available in multiple languages, including English, Spanish, Karen, Oromo, Somali and Hmong. Totally, there were 13,215 people from 6,558 households participated in the survey, providing rich information about the travel behavior in the Twin Cities area.

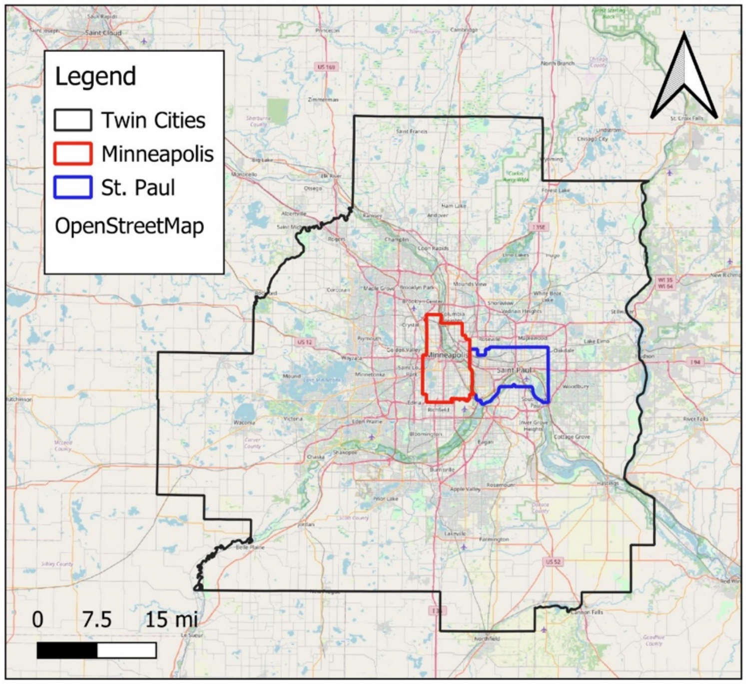


Figure 1. Study area

Individuals residing in Minnesota become eligible to apply for a driving license at the age of 16 or beyond. Consequently, this study incorporated participants who were at least 16 years of age. This study included 11,005 participants in the final analysis. Compared with the whole population (Table 1), the sample exhibits similar age and household income distributions. Additionally, our sample has a higher level of education and a slightly larger percentage of White individuals. However, the sample has a lower percentage of males. The home locations of these selected participants cover 1,867 of the 2,085 CBGs (90%) in the Twin Cities (Figure 2), showing a good spatial representativeness of the study area. In summary, the sample has similar demographic patterns and spatial distribution with the overall population. While some biases are present, they are not expected to significantly impact the modeling analysis and interpretation in this study. With a sufficient number of observations, the sample encompasses a diverse range for each factor, enabling a thorough examination of factors affecting travel-related carbon emissions.

Table 1. Comparison of demographic characteristics between sample and population

	Sample characteristics	Population characteristics [a]
Age (Median) [b]	45-54	47
Household income (Median) [b]	$75,000-$99,999	$92,000
Education (Median) [c]	Bachelor’s degree	One year of college
Male	45%	49%
White	86%	84%

Note:
[a] Population characteristics were calculated based on the 2014-2018 American Community Survey (ACS) 5-year data in the Twin Cities. We only included observations with age equal to or larger than 16.
[b] ACS provides detailed data on age and household income with exact values rather than categorized levels.
[c] ACS provides education in terms of categorized level: 1=Nursery school to grade 4; 2=Grade 5, 6, 7, or 8; 3=Grade 9; 4=Grade 10; 5=Grade 11; 6=Grade 12; 7=1 year of college; 8=2 years of college; 9=3 years of college; 10=4 years of college; 10=5+ years of college.

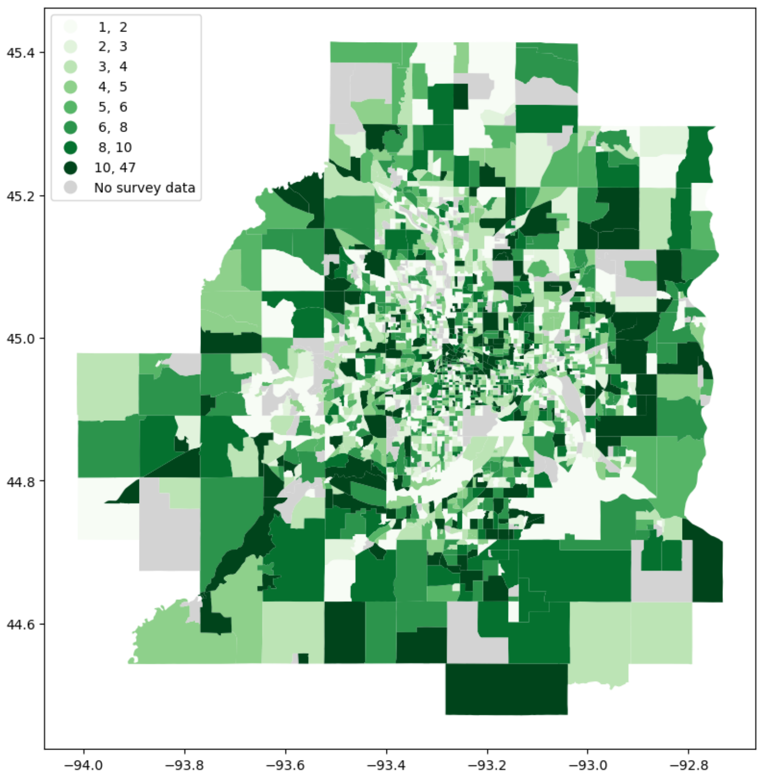


Figure 2. Distribution of the number of participants by CBGs

This study considered several important demographic variables about the participants, which were collected from the TBI survey (Table 2). Among them, driving license, worker, male, disability, white, and student are dummy variables. Household income, education level, and age are continuous variables.

As to the built environment variables, this study followed Ewing and Cervero (2010) and considered five types of variables, including density, diversity, design, distance to transit (e.g., transit supply), and destination accessibility (Table 2). Since the survey data only provided the residential locations at the census block group (CBG) level, the built environment variables were computed at CBG level accordingly. Job accessibility, which indicates the amount of employment opportunities that can be reached within 20-minute driving from the centroid of the CBG, was requested from the Accessibility Observatory. It was calculated with Equation (1) (Owen & Murphy, 2020).

$$A_i = \sum_j O_j f(C_{ij}) \quad (1)$$

A_i indicates the job accessibility by driving for ith location (i.e., the centroid of the CBG in this study). O_j indicates number of employment opportunities at jth location. C_{ij} indicates the travel time cost from ith to jth location. $f(C_{ij})$ is the weighting function. If C_{ij} is smaller than a travel time threshold t (i.e., 20 minutes in this study), $f(C_{ij})$ is 1; otherwise, $f(C_{ij})$ is 0. More information for job accessibility calculation could be found in Owen and Murphy (2020).

Land use mix indicates the diversity of land use in the CBG. It was calculated with Equation (2) (Song et al., 2013).

$$M = -\frac{\sum_i^k p_i \ln(p_i)}{\ln(k)} \quad (2)$$

M indicates the land use mix of the CBG. p_i indicates the percentage of ith land use type in the CBG. k indicates the number of land use types in the CBG. The land use dataset in the Twin Cities was requested from the Minnesota Geospatial Commons.

Intersection density indicates the number of four- or more-way intersections per square mile in the CBG. The intersection dataset was also requested from the Minnesota Geospatial Commons. Transit stop density indicates the number of transit stops per square mile in the CBG. The transit stop dataset was requested from the Metro Transit, the local transit agency, based on its operation schedule in September 2018. Population density indicates the number of people per acre in the CBG and was requested directly from the 2014-2018 5-year American Community Survey. Distance to St. Paul and distance to Minneapolis indicate the driving distance from the centroid of the CBG to downtown St. Paul and downtown Minneapolis, respectively. They were requested through the distance matrix API provided by the Google Maps platform.

We expect household income has a positive relationship with travel-related carbon emissions, as higher-income households may make more trip per person than lower-income ones (Wang & Renne, 2023). For built environment attributes, we expect distance to Minneapolis and distance to St. Paul could be positively correlated with travel-related carbon emissions. Literature has shown that distance to downtown has a positive contribution to people's driving distance and related carbon emissions (Ewing & Cervero, 2010; FHWA, 2018; Shao et al., 2023). On the contrary, we expect that job accessibility, land use mix, intersection density, transit stop density, and population density could be negatively correlated with travel-related carbon emissions. Areas with higher levels of these attributes tend to promote active travel and transit, both of which have lower per-person carbon emissions compared to driving (Ewing & Cervero, 2010; Tao & Cao, 2023).

As low-income households have limited options in work opportunities (Blumenberg, 2017), we expect that the interaction between household income and job accessibility may have additional effect on travel-related carbon emissions. For example, low-income people living in areas with a low level of job accessibility may produce additional carbon emissions because they are likely to seek working opportunities far away from their residences. Similarly, household income may have additional interactive effects with two distance to downtown variables. The imbalance distribution of opportunities between the downtown and suburban areas (Levinson & Keizek, 2018) may force low-income people to travel more and generate additional travel-related carbon emissions. Furthermore, as many captive transit riders are from low-income household (Wang & Renne, 2023), we expect that low-income people living in rich transit supply areas may generate additional travel-related carbon emissions because they tend to make more transit trips.

All the independent variables considered in this study are listed in Table 2. To minimize the impact of multicollinearity on our results, we examined the variance inflation factors (VIFs) for all independent variables. The analysis indicated that all VIFs were below 10, suggesting that multicollinearity is not a significant concern in this study.

Table 2. Variable description, data sources, and descriptive statistics (N=11,005)

Variable	Description	Data sources	Mean	Standard deviation
Daily average carbon emission	Personal daily average carbon dioxide emission in pound on weekdays	This study	15.33	19.9
Built environment variables (measured at the census block group (CBG) where the participant live)				
Job accessibility	Number of jobs in thousand that can be reached by auto in 20 minutes	Accessibility Observatory [a]	509,862	313,292
Land use mix	The entropy index of land use mix	MGC [b]	0.57	0.21
Intersection density	Number of four- or more-way intersections per square mile		42.31	45.76
Transit stop density	Number of transit stops per square mile as of Sep. 2018	Metro Transit	27.07	34.28
Population density	Number of people per acre	ACS [c]	8.33	10.11
Distance to St. Paul	The driving distance in miles from the centroid of the CBG to downtown Saint Paul	Google map [d]	16.88	9.29
Distance to Minneapolis	The driving distance in miles from the centroid of the CBG to downtown Minneapolis		14.22	8.86
Demographic variables				
Driving license	A dummy variable indicating whether the respondent has a driving license	TBI survey	0.94	0.24
Worker	A dummy variable indicating whether the respondent is employed (full-time, part-time, self-employed)		0.67	0.47
Male	A dummy variable indicating whether the respondent is male		0.45	0.5
Household income	The respondent's household income 1 = Under $15,000 2 = $15,000-$24,999 3 = $25,000-$34,999 4 = $35,000-$49,999 5 = $50,000-$74,999 6 = $75,000-$99,999 7 = $100,000-$149,999 8 = $150,000-$199,999 9 = $200,000-$249,999 10 = $250,000 or more		5.94	2.06
Education level	Educational background of the respondent 1 = Less than high school 2 = High school graduate/General educational development 3 = Some college 4 = Vocational/technical training 5 = Associate degree 6 = Bachelor's degree 7 = Graduate/post-graduate degree		5.22	1.79
Disability	A dummy variable indicating whether the respondent has a disability		0.04	0.21
Age	Age category of the respondent 3 = 16-17 4 = 18-24		7.04	1.79

Variable	Description	Data sources	Mean	Standard deviation
	5 = 25-34 6 = 35-44 7 = 45-54 8 = 55-64 9 = 65-74 10 = 75 or older			
White	A dummy variable indicating whether the respondent is Caucasian		0.86	0.34
Student	A dummy variable indicating whether the respondent is a student		0.08	0.27

Note:
[a] Accessibility Observatory at the University of Minnesota: https://access.umn.edu/
[b] Minnesota Geospatial Commons: https://gisdata.mn.gov/
[c] American Community Survey 2014-2018 5-year estimates: https://www.census.gov/programs-surveys/acs/news/data-releases.html
[d] Google Maps platform (distance matrix API): https://developers.google.com/maps/documentation/distance-matrix/overview

3.2 Methods

The dependent variable considered in this study is personal daily average carbon dioxide emissions on weekdays. We focus on weekday trips because a significant portion of the participants did not have travel information on weekends. Participants using smartphone applications can report seven days of travel trips. Participants using online survey and call center can only report one weekday of their travel (Metropolitan Council, 2019). According to the survey data, among all 11,005 participants included in the current analysis, only 6,982 (63%) of them have travel information available on weekends. In this case, considering weekends may introduce a bias to the results.

For simplicity, this study used travel-related carbon emissions (TCE) in the rest of the paper when necessary. With the travel diary information from the TBI survey, this study used Equation (3) to calculate TCE.

$$TCE = \frac{\sum_{1}^{I}\sum_{1}^{J_i}\frac{D_{ij}^{k}R_k}{N_{ij}}}{I} \tag{3}$$

In Equation (3), j ($j = 1,2,\ldots,J_i$) indicates the jth trip on the day and J_i indicates the number of trips reported on ith day by a participant. i ($i = 1,2,\ldots,I$) indicates the ith travel day and I indicates the number of travel days reported by a participant. k ($k = 1,2,\ldots,4$) indicates the index of the travel mode of a trip. The TBI survey provides more than 10 travel modes for the participants to choose from. In this study, we aggregated them into four types, including driving, bus, light rail, and active travel (i.e., biking and walking). D_{ij}^{k} presents the distance of jth trip on ith day with travel mode k. The TBI survey provides detailed information for each trip reported by the participants, including starting time, ending time, number of travelers, distance, and travel mode. We removed trips with abnormal travel speed or distance. Then, we summarized number of trips for each day (J_i) and number of travel days (I) for each participant.

N_{ij} is the number of travelers of jth trip on ith day, which is available in the TBI survey. For some auto trips, multiple people shared the same vehicle. This study assumed the people in the same trip equally shared the associated carbon dioxide emission. This assumption improved the accuracy of the calculation of personal carbon emission compared with several previous studies (Shao et al., 2023; Wu et al., 2019). For auto trips, N_{ij} represents the number of travelers reported in the TBI survey. If this information is unavailable for a given auto trip, N_{ij} is assigned the average number of travelers for all auto trips with valid data, which is 1.9 people in this study. For bus, light rail, and active travel trips, N_{ij} is set to 1.

R_k is the carbon dioxide emission rate for travel mode k. FTA (2010) provided carbon dioxide emission rates for different travel modes calculated with the data from 2008. Given the facts that power plants have reduced their use of coal to generate electricity and private vehicles have increased significantly in fuel efficiency (i.e., mile traveled per gallon), the carbon emission rates from 2008 were outdated when this study was conducted. Therefore, this study updated the carbon emission rates with the data in 2019 through the methodology provided by FTA (2010). The carbon emission rates used in this study are listed in Table 3. Note that carbon emission rates might be different for different makers and types of vehicles. However, this study did not consider these factors as the trip information provided by TBI survey did not provide the information of the vehicles used in the auto trips. Note that this study did not consider the carbon emission rate for electric private vehicles as the share of electric vehicles in 2019 was very low in Minnesota. According to the US Department of Energy (DOE, 2018), there were only 7,700 electric vehicles registered by 2018, which accounted for approximately 0.1% of all auto registrations in Minnesota.

Table 3. Carbon dioxide emission rates for different travel modes

Travel mode	**Carbon dioxide emission rate (pounds per passenger mile)**
Driving (Single occupied vehicle)	0.88
Bus	0.64
Light rail	0.43
Active travel	0

Figure 3 presents the distribution of the personal average daily TCE by the CBGs where the participants live across the Twin Cities. A clear pattern is that people living farther away from the urban area are associated with larger amounts of carbon emissions. Aside from distance to downtown, it is difficult to descriptively identify the impacts of other built environment attributes on TCE.

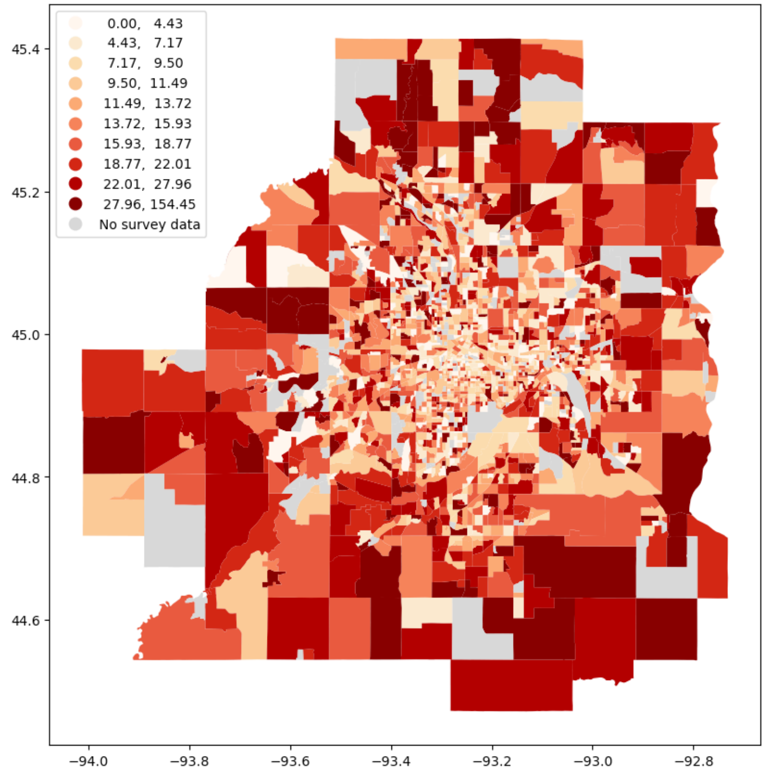


Figure 3. Distribution of personal average daily TCE at the CBGs where the participants live

This study applied the Extreme Gradient Boost (XGBoost) approach to estimate the nonlinear relationships between TCE and two types of independent variables, i.e., built environment attributes and demographics. XGBoost (Chen & Guestrin, 2016) applies the parallel tree boosting method to significantly to increase the accuracy and speed when fitting the sample. In general, XGBoost applies the decision tree method to split the sample into several subsamples according to certain criteria and combines multiple decision trees into a large complex model.

XGBoost has several advantages compared to traditional statistical models such as linear and generalized linear models. First, XGBoost has no assumptions about the form of the relationships between the dependent and independent variables and, thus, is effective in estimating irregular nonlinear relationships. Second, XGBoost has a better prediction capability. Third, XGBoost can better handle outliers and missing values. Listwise deletion is the standard method for handling missing values in traditional models. However, it can lead to biased parameter estimates if the missing data is not random. Decision trees tackle this problem by grouping observations with missing values into one subsample when dividing a sample (Therneau & Atkinson, 2022). Moreover, decision trees depend solely on the rankings of the independent variable values, making them unaffected by extreme outliers (Therneau & Atkinson, 2022). With that said, this study can include more observations in the model, and this could further increase the estimation performance. Compared with other popular machine learning approaches used in built environment and travel behavior studies, such as gradient boosting decision trees (GBDT) and random forest (RF) (Aghaabbasi & Chalermpong, 2023), XGBoost is faster in fitting the data as it uses more regulations (Chen & Guestrin, 2016).

This study applied the five-fold cross validation approach to search for the best combination of parameters, including number of decision trees, tree depth, and learning rate. Number of decision trees indicates the number of decision trees combined into the model. More decision trees usually provide better fitting performance. Tree depth represents the complexity of one decision tree. Learning rate indicates how much portion of the result from each decision tree is combined into the final model. After the cross validation, the final model included 500 decision trees. The corresponding learning rate is 0.01, and tree depth is 3. In this study, we also compared the model performance of XGboost with RF and GBDT (see Section A1 in Appendix for more information). Overall, XGBoost performs better than the other two approaches.

To interpret the model results, this study applied two tools. First, this study employed the SHapley Additive exPlanations (SHAP) value (Chen et al., 2022) to measure the contribution of independent variables. SHAP value is originated from the concept of Shapely value based on the cooperative game theory (Shapley, 1953). Generally, SHAP value measures the marginal impact of the inclusion of one independent variable on the prediction performance for each observation. A positive SHAP value indicates a positive impact and vice versa. In addition, the larger the absolute value of the SHAP value, the larger impact the corresponding independent variable has on prediction. The average absolute SHAP values of an independent variable for all observations is the contribution of the corresponding independent variable. Second, this study used the accumulated local effect (ALE) plots to visualize the nonlinear relationship between travel-related carbon emissions and independent variables. In addition, ALE plots (Apley & Zhu, 2020) were used to show the interactive impact between household income and built environment variables on travel-related carbon emissions. Compared with another commonly used tool, partial dependence plots, ALE plots are better in terms of their capability to handle multicollinearity among independent variables (Molnar, 2020).

4 Results

4.1 Variable contributions

Table 4 presents the contribution of each independent variable considered in this study. The SHAP value gauges the average absolute value of marginal impact when including the corresponding independent variable (i.e., built environment attribute or demographic variable) in predicting TCE for all individuals in the sample. The percentage indicates the proportion of the SHAP value of the independent variable among the sum of SHAP values of all independent variables. This study also ranked the independent variables based on their contributions and calculated the collective contributions in terms of percentage for built environment attributes and demographics, respectively.

Collectively, demographics make 60% of the contribution in estimating travel-related carbon emission, which is larger than that of built environment attributes (40%). This result shows that demographics play a more important role than the built environment in affecting people's TCE on weekdays.

Working status is the most important variable among all independent variables, with a SHAP value of 2.08. This result indicates that, on average, the marginal impact of being employed on TCE is 2.08 pounds. The second and third most important variables are job accessibility and gender, respectively. Household income has an SHAP value of 0.39, which is ranked eighth among all 16 variables considered in this study.

Table 4. Contributions of independent variables in predicting TCE

Type	Variable	SHAP value	Proportion	Ranking	Sum
Built environment attributes	**Job accessibility**	1.55	17.6%	2	40.0%
	Distance to St. Paul	0.46	5.2%	6	
	Distance to Minneapolis	0.45	5.1%	7	
	Transit stop density	0.38	4.3%	9	
	Intersection density	0.25	2.8%	11	
	Land use mix	0.23	2.6%	12	
	Population density	0.22	2.5%	14	
Demographic attributes	**Worker**	2.08	23.6%	1	60.0%
	Male	0.88	10.0%	3	
	Driving license	0.67	7.6%	4	
	Education level	0.55	6.3%	5	
	Household income	0.39	4.4%	8	
	Age	0.33	3.7%	10	
	Disability	0.22	2.5%	13	
	White	0.15	1.7%	15	
	Student	0.01	0.1%	16	

4.2 Main effects

This section presents the main effects of all built environment attributes and important demographics in terms of ALE plots. The ALE plot of an independent variable illustrates how its centered effect varies by its value. The centered effect is calculated by subtracting the average ALE value from each of the individual ALE values, which will shift the entire ALE curve so that its average effect is zero. Centering helps in emphasizing how the ALE values deviate from their average value. To facilitate the comparison across plots, this study applied the same scale of y axis for all ALE plots. The x axis represents the corresponding built environment attribute. The rugs on the horizontal axis represent the distribution of the built environment attribute. Segments in the ALE plots with a higher number of observations are likely to yield more robust results.

Figure 4 presents the relationship between TCE and important built environment variables, including job accessibility, distance to St. Paul, distance to Minneapolis, and transit stop density. When job accessibility increases from 0 to approximately 800,000 jobs, it has a negative impact on TCE. This relationship is consistent with our expectation as people living in areas with higher job accessibility usually have shorter travel distances. However, after the threshold of 800,000 jobs, the impact becomes trivial, suggesting there is a threshold effect regarding job accessibility. Areas with job accessibility exceeding 800,000 jobs are primarily concentrated in the economically vibrant areas of the region, particularly in Minneapolis, with some also located in St. Paul. In these areas, more people switch to biking and walking to reach their destination instead of driving and transit. According to this result, increasing job accessibility to 800,000 jobs could significantly reduce the TCE in the CBG.

Distance to St. Paul has a reverse U-shape relationship with TCE. When distance to St. Paul increases from 0 to 15 miles, there is a fast-increasing trend. The city of St. Paul is one important job center in the Twin Cities area. For people living farther from it, their commuting distance becomes longer, and the associated driving or transit trips generate more carbon emissions. When the distance increases from 15 to 36 miles, the increase of this positive relationship becomes slower. This implies that the attraction of St. Paul becomes smaller in these areas and its influence on people's TCE starts to decrease. When the distance exceeds 36 miles, the relationship becomes negative, which is different with our expectation. This might be because of the dual-center structure of Twin Cities and some of these areas are closer to the city of Minneapolis. People living in those areas choose to work in Minneapolis and have a lower level of TCE.

Distance to Minneapolis generally has a positive relationship with travel-related carbon emission. The reason for this positive relationship is similar to that of distance to St. Paul, i.e., the attraction of employment center. However, Minneapolis is the economic center of the Twin Cities area and, thus, has a stronger attractiveness than other employment centers in the area. This relationship becomes trivial when the distance exceeds 7 miles. Note that there is a negative impact when the distance starts to increase from 0. This negative relationship might be because people living in the downtown areas have more short trips and generate more TCE.

The relationship between transit stop density and travel-related carbon emission is U-shape. Initially, transit stop density is negatively associated with TCE, mainly because transit travel modes (e.g., bus and light rail) are efficient in reducing TCE than driving. However, after transit stop density reaches about 40, it starts to become positively correlated with TCE, which is inconsistent with our expectation. This positive relationship is mainly because that more transit trips are associated with more carbon emissions.

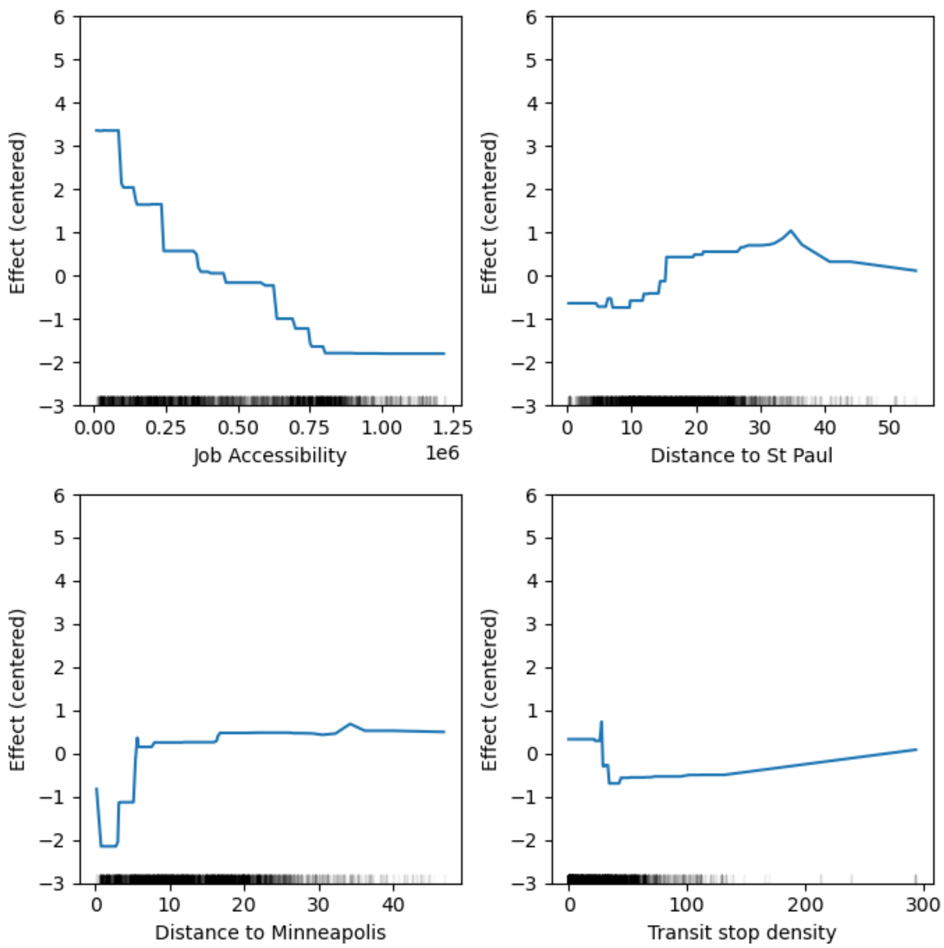


Figure 4. The relationships of important built environment variables

Figure 5 presents the relationship between household income and TCE. In general, household income also has a positive relationship with TCE, which is consistent with our expectation. However, this relationship becomes trivial when household income level exceeds 5 (i.e., 50,000 to 74,999 dollars). The relationships of other built environment attributes and important demographics are presented in Figure A1 and Figure A2 in Section A2 in Appendix, respectively.

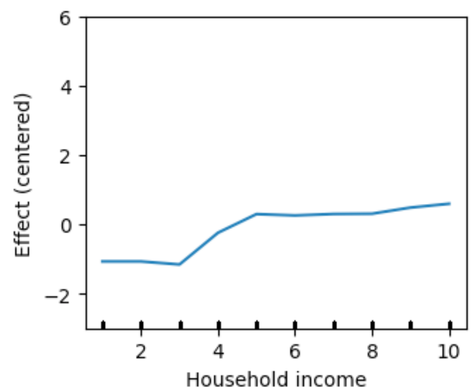


Figure 5. The relationship of household income

4.3 Interactive effects

This section presents the interactive effects between household income and important built environment variables on TCE. The interactive effects are shown as 3D ALE plots. The plot has three axes. The x axis represents household income. The y axis represents the built environment variable. The z axis presents the interactive effect of the two variables on TCE. Note that the interactive effect indicates the marginal effect generated by the interaction between the two variables and is not related to their main effects. To facilitate the readiness of the plots, this study used different colors to represent the value of the interactive effect. Darker red indicates a larger positive effect. Darker blue indicates a larger negative effect. White indicates no effect. This study also used the same scale of z axis for all plots to facilitate comparison across plots. One exception is the plot of household income and distance to St. Paul as their interactive effect is larger than other interactions.

Figure 6 presents the interactive effect between household income and job accessibility on TCE. There is a positive effect when both household income and job accessibility are in their lower values (i.e., household income smaller than or equal to approximately 3, i.e., 25,000-34,999 dollars, and job accessibility smaller than about 200,000 jobs). Figure 7 shows CBGs with job accessibility under 200,000 jobs and where participants with incomes at or below level 3 reside. This result suggests that, for people with lower household income, when they live in the areas with lower job accessibility, they tend to generate extra TCE. Low-income population usually have much fewer employment choices than high-income population. When there are fewer job opportunities they can access near their home locations, they are more likely to seek job opportunities in farther locations. In addition, most low-income jobs provide limited options to work at home, which forces people to commute.

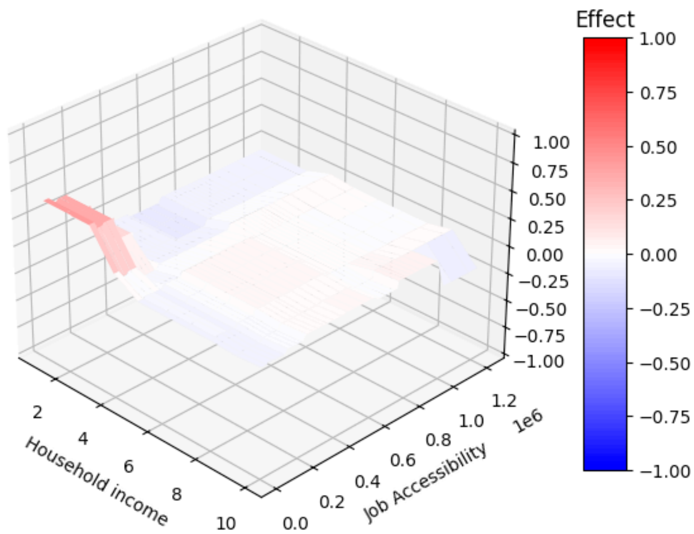
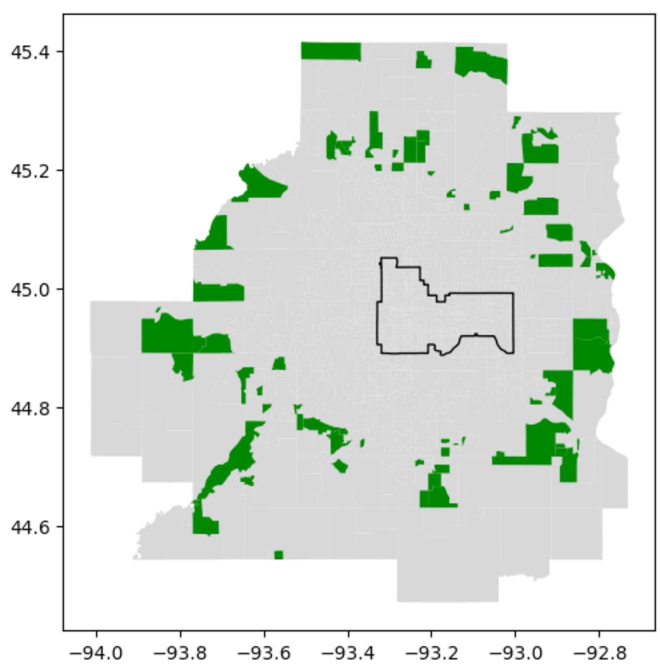


Figure 6. The interactive effects between household income and job accessibility



Notes: The CBGs in green indicates the areas where job accessibility is smaller than 200,000 jobs within 20 minutes of driving and the participants with their household income smaller than or equal to level 3 ($25,000-$34,999) live. The black line indicates the boundary of the two major cities – Minneapolis and St. Paul.

Figure 7. Distribution of the low-job-accessibility CBGs where low-income participants live

The interactive effect between household income and distance to St. Paul is shown in Figure 8. In regions where household incomes are below the third level and the distance from St. Paul exceeds 30 miles, a notable positive impact is observed. Lower-income people, when living in the areas farther from St. Paul, have a higher propensity to produce more TCE. Figure 9 presents CBGs 30 miles from downtown St. Paul and where participants with incomes at or below level 3 (i.e., 25,000 to 34,999 dollars) live. Note that those locations are also far away from Minneapolis. Lower-income people have fewer opportunities to work at home (Tao & Cao, 2021) and are less inclined to engage in online shopping (Cao et al., 2011; Saphores & Xu, 2021). Consequently, they often undertake more journeys for both commuting and shopping purposes, especially when residing at greater distances from urban centers, compared to their higher-income counterparts.

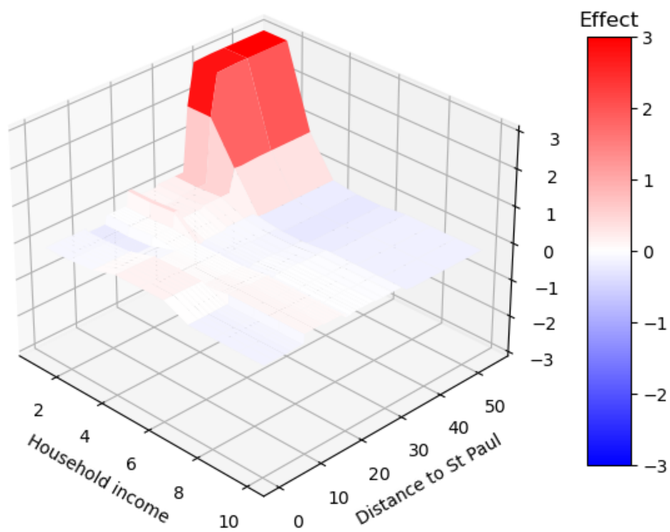
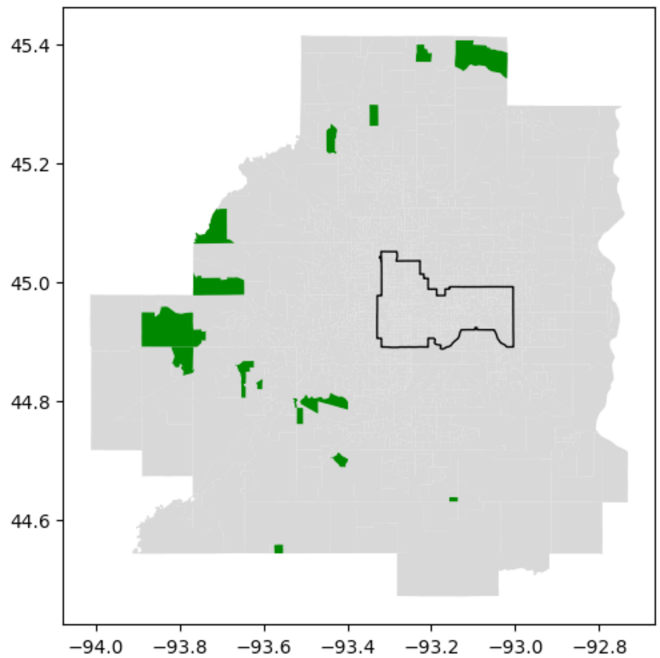


Figure 8. The interactive effects between household income and distance to St. Paul



Notes: The CBGs in green indicates the areas more than 30 miles from downtown St. Paul and where the participants with their household income smaller than or equal to level 3 ($25,000-$34,999) live. The black line indicates the boundary of the two major cities – Minneapolis and St. Paul.

Figure 9. Distribution of the rural CBGs where low-income participants live

Figure 10 presents the interactive effect between household income and distance to Minneapolis on TCE. The plot shows that, in downtown Minneapolis (i.e., where distance to Minneapolis is very small), lower-income individuals tend to generate more TCE, while those with higher incomes tend to produce less. Figure 11 presents the CBGs less than 3 miles from the downtown Minneapolis, including the downtown and areas near the downtown. Downtown Minneapolis and its nearby areas, being the economic hub of the Twin Cities, host a variety of high-income job opportunities in sectors such as banking, insurance, and consulting (MPLS Downtown Council, 2021). Consequently, higher-income residents living downtown are more likely to find jobs that match their skills in close proximity, leading to shorter commutes. In contrast, lower-income individuals may often need to seek employment in farther areas, resulting in longer travel distances.

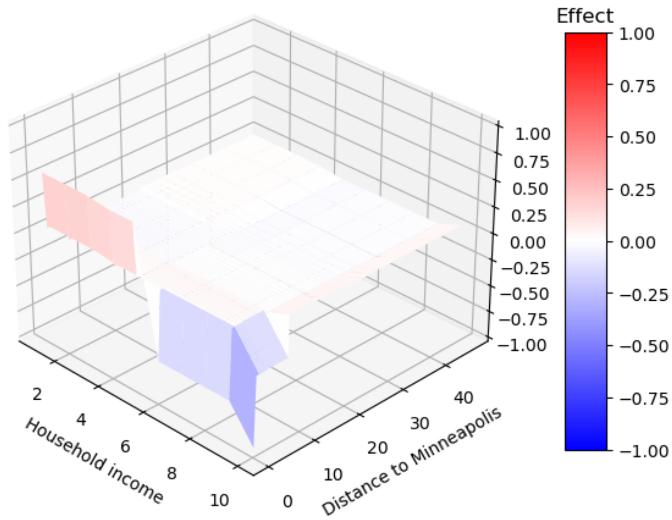
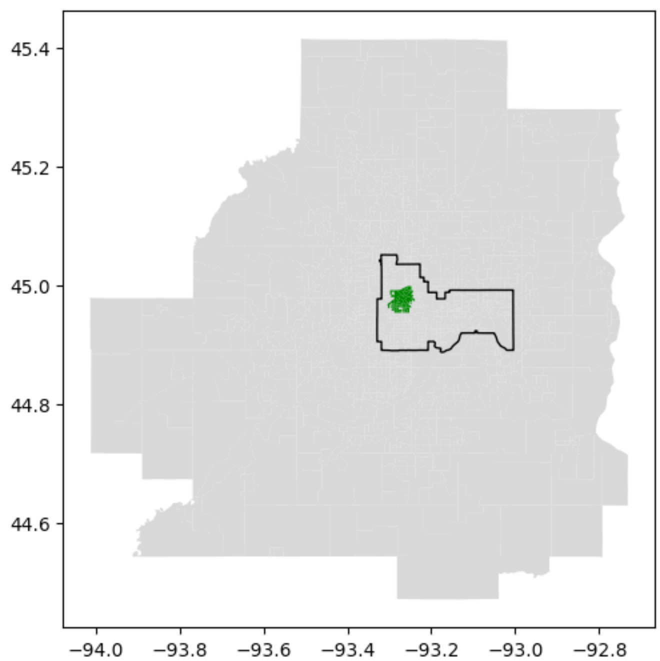


Figure 10. The interactive effects between household income and distance to Minneapolis



Notes: The CBGs in green indicates the areas within 3 miles from downtown Minneapolis, including the downtown and areas near downtown Minneapolis. The black line indicates the boundary of the two major cities – Minneapolis and St. Paul.

Figure 11. Distribution of the downtown CBGs

Figure 12 present the interactive effect between household income and transit stop density on TCE. When transit stop density is higher than around 120 stops per square mile near their home locations, lower-income people tend to have additional TCE while higher-income people tend to have less TCE. This difference is mainly because low-income people are captive transit riders and use more transit than high-income people. These additional transit trips generate more TCE. Figure 13 shows the CBGs with their transit stop density higher than 120 stops per square mile. All these CBGs are situated within the urban areas of Minneapolis and St. Paul. Low-income residents in these areas often face longer commutes to suburban areas to find service jobs due to the job-housing imbalance (Levinson & Keizek, 2018). Living in areas with more transit options encourages these residents to use public transit for long-distance travel, which results in additional TCE.

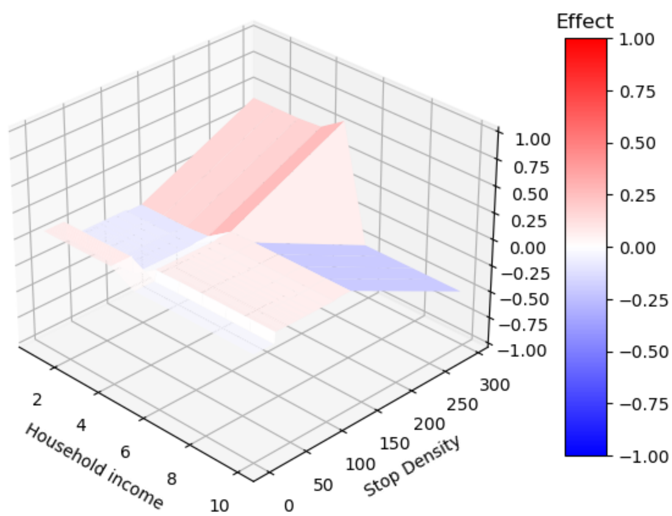
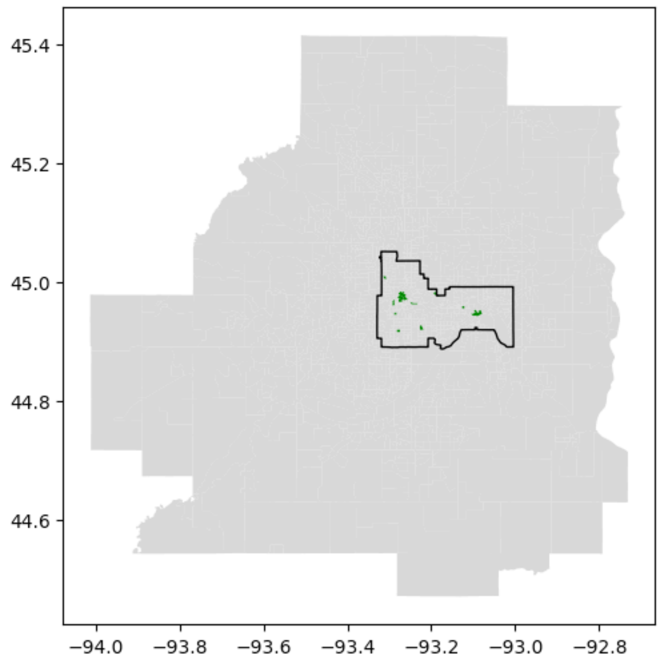


Figure 12. The interactive effects between household income and transit stop density



Notes: The CBGs in green indicates the areas with more than 120 transit stops per square mile. The black line indicates the boundary of the two major cities – Minneapolis and St. Paul.

Figure 13. Distribution of the rich-transit-supply CBGs

5 Conclusions

With the data from the Twin Cities area in US, this study employed the XGBoost approach to estimate the contributions of built environment attributes and demographics to predicting TCE and their nonlinear relationships. Besides the main effects, this study estimated the interactive effect between household income and important built environment variables on TCE.

Demographic variables collectively have a larger contribution than built environment attributes in estimating TCE. Individually, six demographic variables, including employment status, gender, driver's license, education level, household income, and age, are among the ten most important variables according to their SHAP values. This result confirms the important role of demographics in influencing TCE. For built environment attributes, job accessibility, distance to St. Paul, distance to Minneapolis, and transit stop density are important variables, showing that distribution of job opportunities, the city structure, and transit supply are important in affecting people's TCE.

Furthermore, the main effects show salient nonlinear relationships. The interactive effects indicate that the interaction between household income and built environment attributes can lead to additional impacts on TCE. Positive interactive impacts are observed among lower-income residents living in areas with limited job accessibility, greater distances from downtown St. Paul, within downtown Minneapolis, and in locations with abundant transit options. Conversely, negative interactive impacts are noted for higher-income individuals residing in downtown Minneapolis and areas well-served by public transit.

This study offers several important policy implications for creating sustainable and equitable transportation systems. First, more environment friendly and affordable transportation options, such as sharing electric vehicles, should be provided to low-income population, especially those living in areas with low job accessibility, downtown areas, and areas farther from the city center. One example is the Evie carsharing program in the Twin Cities[2]. This program started its operation in 2022 and provided all-electric and free-floating vehicles. This program offered discounted prices (i.e., Access PLUS) for the low-income population. Currently, the program covers the downtown area in Minneapolis, which could help address the additional TCE generated by low-income households living there. In the future, the program could cover more places with lower job accessibility and farther from the urban center.

Second, transit agencies should switch to using vehicles with clean energy such as electricity. Transit systems need more riders to ensure their efficiency in energy use and reduction in carbon emissions. However, transit ridership has been struggling in recent years in the US. For example, the transit ridership in the Twin Cities area has been decreasing since 2016 (Metro Transit, 2023). The pandemic also made another huge hit on the ridership. The energy consumption per passenger mile of the transit system has increased since then. The main result in this study showed that more transit supply is associated with more travel-related carbon emission. Furthermore, low-income people living in areas with more transit supply tend to produce additional carbon emissions via transit trips. Therefore, the local transit agencies should switch to vehicles with clean energy, such as electric vehicles. Electric vehicles are more efficient in energy use and could help significantly reduce carbon emissions. For example, the

[2] Evie Carshare: https://eviecarshare.com/

Zero-Emission Bus Transition Plan by the Metro Transit (2022) in the Twin Cities plans to replace 20% of its vehicles with electric ones.

Finally, more matched job opportunities should be provided to low-income people near their residential locations. The results of interactive effects showed that spatial mismatch forces the low-income population to seek jobs in areas farther from their home locations and make longer commuting trips. Providing matched employment opportunities near or in their residential areas could help shorten or even reduce these trips. This requires the coordination between different jurisdictions at the regional level to ensure that job opportunities and their nearby residents matched.

This study has several limitations. First, it did not account for spatial dependence. When working with TCE at a disaggregated level, XGBoost and similar tree-based machine learning methods are unable to control for spatial dependence. Second, the study identified correlations rather than causations, as we lack longitudinal records of people's travel behavior and related factors. Future studies could explore the causal relationships when the related longitudinal dataset is available. Finally, the findings of this study are specific to the Twin Cities area, and caution should be exercised when applying them to other contexts. Future research could replicate this study in different cities to assess the generalizability of the results.

Appendix

A1 Model performance comparison

We compared the model performance of three most frequently used machine learning approaches on built environment and travel behavior studies, including the eXtreme Gradient Boosting (XGBoost), gradient boosting decision tree (GBDT), and random forest (RF), according to a recently-published comprehensive literature review (Aghaabbasi & Chalermpong, 2023). We applied three evaluation measures, including R-squared (R^2), mean absolute error (MAE), and root mean squared error (RMSE). We carried out a five-fold cross validation to tune the parameters. After determining the best combination of parameters, we calculated the evaluation measures for each approach. The results are listed in Table A1 below. Overall, XGBoost performs better than the other two approaches.

Table A1. Model performance comparison

	R^2	Standard Deviation	MAE	Standard Deviation	RMSE	Standard Deviation
RF	0.064	0.011	12.648	0.299	19.324	1.601
GBDT	0.063	0.015	12.628	0.306	19.331	1.593
XGBoost	0.076	0.015	12.393	0.210	19.062	1.773

A2 Supplemental ALE plots

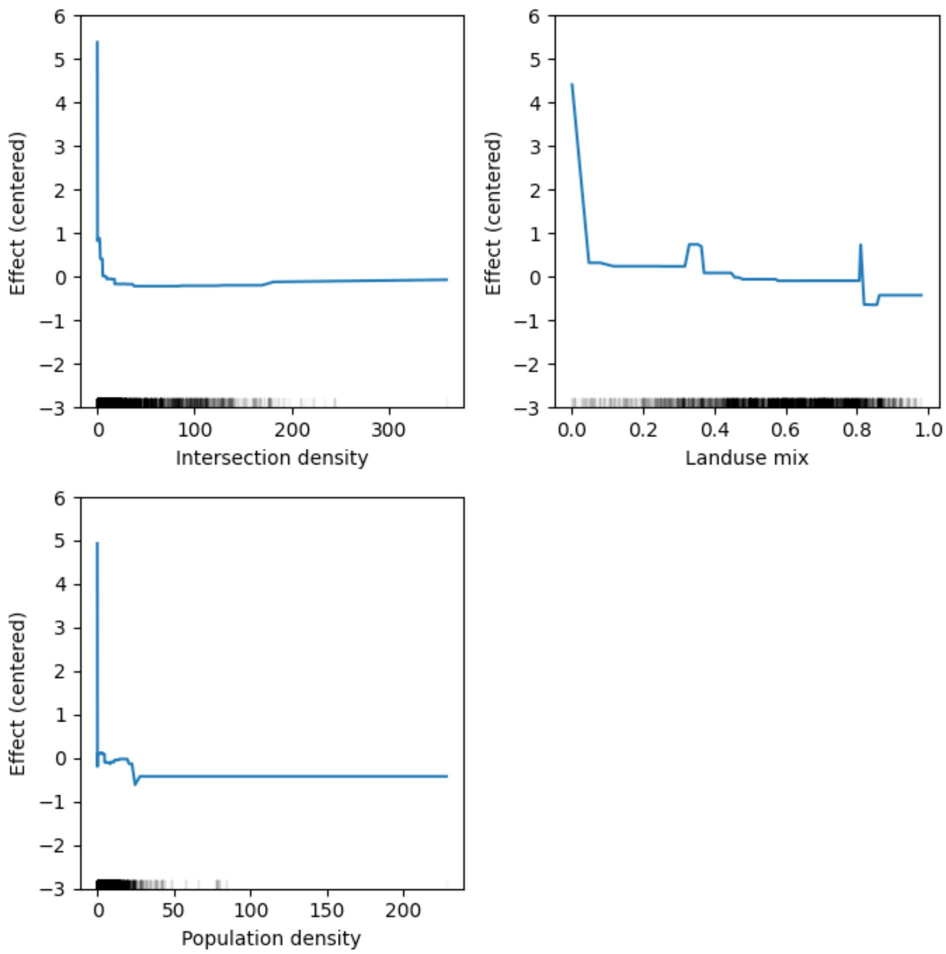


Figure A1. The relationships of other built environment variables

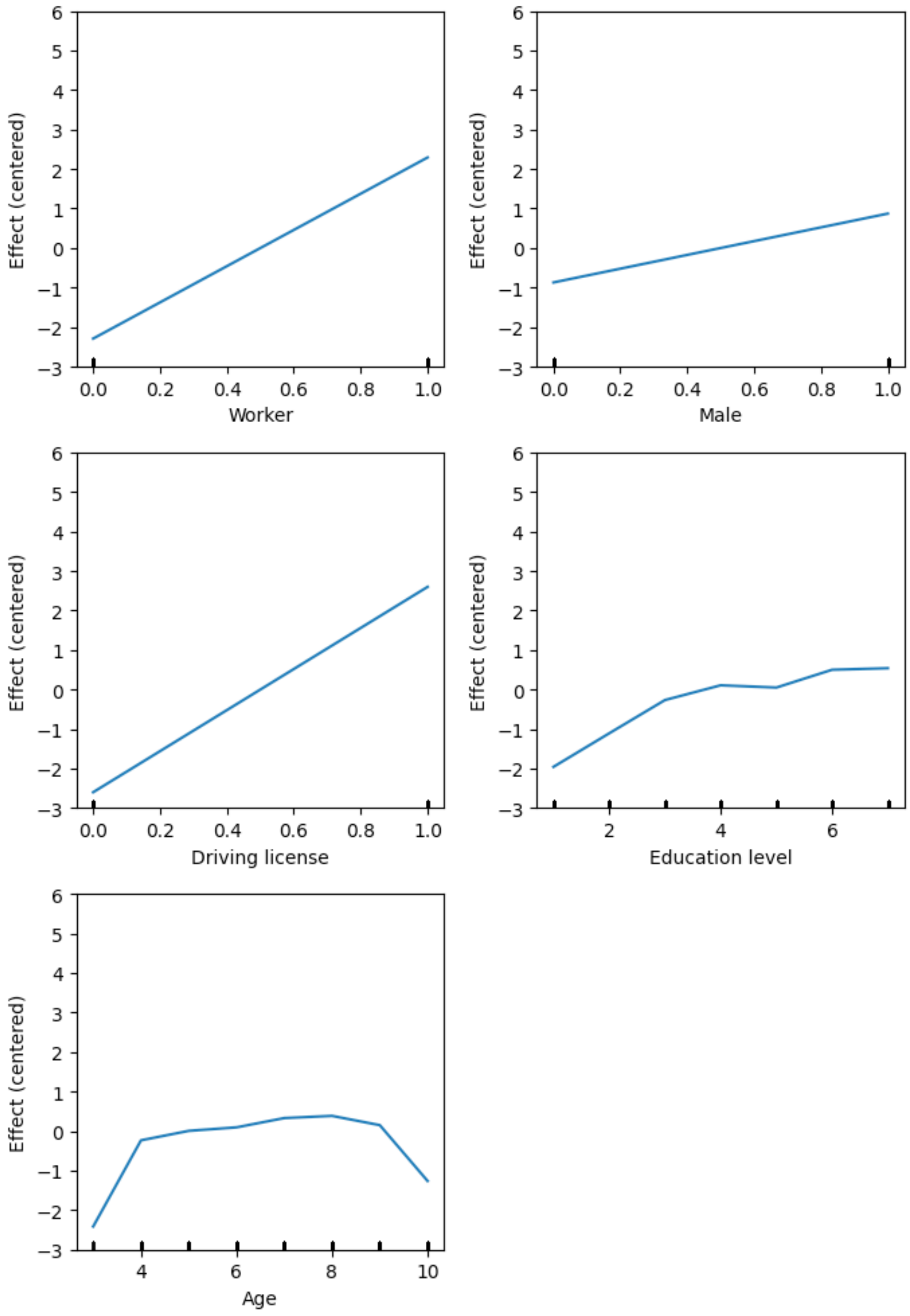


Figure A2. The relationships of other important demographic variables

References

Aghaabbasi, M., & Chalermpong, S. (2023). Machine learning techniques for evaluating the nonlinear link between built-environment characteristics and travel behaviors: A systematic review. *Travel Behaviour and Society*, *33*. https://doi.org/10.1016/j.tbs.2023.100640

Apley, D. W., & Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *82*(4), 1059-1086. https://doi.org/10.1111/rssb.12377

Barla, P., Miranda-Moreno, L. F., & Lee-Gosselin, M. (2011). Urban travel CO2 emissions and land use: A case study for Quebec City. *Transportation Research Part D: Transport and Environment*, *16*(6), 423-428. https://doi.org/10.1016/j.trd.2011.03.005

Blumenberg, E. (2017). Social equity and urban transportation. In G. Giuliano & S. Hanson (Eds.), *The geography of urban transportation* (4th ed.). The Guilford press.

Boarnet, M. (2017). Land use, travel behavior, and disaggregate travel data. In G. Giuliano & S. Hanson (Eds.), *The geography of urban transportation* (4th ed.). The Guilford Press.

Boarnet, M. G., Wang, X., & Houston, D. (2017). Can New Light Rail Reduce Personal Vehicle Carbon Emissions? A before-after, Experimental-Control Evaluation in Los Angeles. *Journal of Regional Science*, *57*(3), 523-539. https://doi.org/10.1111/jors.12275

Cao, X., & Yang, W. (2017). Examining the effects of the built environment and residential self-selection on commuting trips and the related CO 2 emissions: An empirical study in Guangzhou, China. *Transportation Research Part D: Transport and Environment*, *52*, 480-494. https://doi.org/10.1016/j.trd.2017.02.003

Cao, X. J., Xu, Z., & Douma, F. (2011). The interactions between e-shopping and traditional in-store shopping: an application of structural equations model. *Transportation*, *39*(5), 957-974. https://doi.org/10.1007/s11116-011-9376-3

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199-219. https://doi.org/10.1016/s1361-9209(97)00009-6

Chen, H., Lundberg, S. M., & Lee, S.-I. (2022). Explaining a series of models by propagating Shapley values. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-31384-3

Chen, T., & Guestrin, C. (2016). *XGBoost* Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,

Choi, K., & Zhang, M. (2017). The net effects of the built environment on household vehicle emissions: A case study of Austin, TX. *Transportation Research Part D: Transport and Environment*, *50*, 254-268. https://doi.org/10.1016/j.trd.2016.10.036

Credit, K., & Lehnert, M. (2023). A structured comparison of causal machine learning methods to assess heterogeneous treatment effects in spatial data. *Journal of Geographical Systems*. https://doi.org/10.1007/s10109-023-00413-0

Ding, C., Cao, X. J., & Næss, P. (2018). 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. https://doi.org/10.1016/j.tra.2018.02.009

DOE. (2018). *Vehicle Registration Counts by State*. Retrieved 07-28-2023 from https://afdc.energy.gov/vehicle-registration?year=2019

EPA. (2022). *Sources of Greenhouse Gas Emissions*. https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment. *Journal of the American Planning Association*, *76*(3), 265-294. https://doi.org/10.1080/01944361003766766

Feng, R., Feng, Q., Jing, Z., Zhang, M., & Yao, B. (2022). Association of the built environment with motor vehicle emissions in small cities. *Transportation Research Part D: Transport and Environment*, *107*. https://doi.org/10.1016/j.trd.2022.103313

FHWA. (2018). *Summary of travel trends (2017 National Household Travel Survey)*. https://nhts.ornl.gov/assets/2017_nhts_summary_travel_trends.pdf

FTA. (2010). *Public Transportation's Role In Responding To Climate Change*. https://www.transit.dot.gov/sites/fta.dot.gov/files/docs/PublicTransportationsRoleInRespondingToClimateChange2010.pdf

Gao, J., Ma, S., Peng, B., Zuo, J., & Du, H. (2022). Exploring the nonlinear and asymmetric influences of built environment on CO2 emission of ride-hailing trips. *Environmental Impact Assessment Review*, *92*, 106691. https://doi.org/10.1016/j.eiar.2021.106691

Global Carbon Bedget. (2023). *Fossil CO2 emissions at record high in 2023*. Retrieved 12-09-2023 from https://globalcarbonbudget.org/fossil-co2-emissions-at-record-high-in-2023/

Levinson, D., & Keizek, K. (2018). *Metropolitan Transport and Land Use Planning for Place and Plexus*. Routledge.

Metro Transit. (2022). *Zero-Emission Bus Transition Plan*. https://metrocouncil.org/Council-Meetings/Committees/Transportation-Committee/2022/January-24,-2022/0124_2022_5-attachment_ZEBTP-FinalReport.aspx

Metro Transit. (2023). *Metro Transit Ridership*. Retrieved 07-31-2023 from https://www.metrotransit.org/performance

Metropolitan Council. (2019). *2019 Travel behavior inventory household survey results*. https://gisdata.mn.gov/dataset/us-mn-state-metc-society-tbi-home-interview2019

Molnar, C. (2020). *Interpretable Machine Learning - A Guide for Making Black Box Models Explainable*. lulu.com. https://christophm.github.io/interpretable-ml-book/

MPLS Downtown Council. (2021). *2021 mpls downtown by the numbers*. Retrieved 07-31-2023 from https://www.mplsdowntown.com/2021-by-the-numbers/

Owen, A., & Murphy, B. (2020). *Access Across America: Auto 2018 Methodology*. https://cts-d8resmod-prd.oit.umn.edu/pdf/cts-20-04.pdf

Sabouri, S., Brewer, S., & Ewing, R. (2020). Exploring the relationship between ride-sourcing services and vehicle ownership, using both inferential and machine learning approaches. *Landscape and Urban Planning*, *198*. https://doi.org/10.1016/j.landurbplan.2020.103797

Saphores, J.-D., & Xu, L. (2021). E-shopping changes and the state of E-grocery shopping in the US - Evidence from national travel and time use surveys. *Research in Transportation Economics*, *87*. https://doi.org/10.1016/j.retrec.2020.100864

Shao, Q., Zhang, W., Cao, X., & Yang, J. (2023). Built environment interventions for emission mitigation: A machine learning analysis of travel-related CO2 in a developing city. *Journal of Transport Geography*, *110*. https://doi.org/10.1016/j.jtrangeo.2023.103632

Shapley, L. S. (1953). A value for n-person games. In *Contributions to the Theory of Games, Volume II, Annals of Mathematics Studies* (pp. 307-317). Princeton University Press.

Song, Y., Merlin, L., & Rodriguez, D. (2013). Comparing measures of urban land use mix. *Computers, Environment and Urban Systems*, *42*, 1-13. https://doi.org/10.1016/j.compenvurbsys.2013.08.001

Stevens, M. R. (2017). Does Compact Development Make People Drive Less? *Journal of the American Planning Association*, *83*(1), 7-18. https://doi.org/10.1080/01944363.2016.1240044
Tao, T., & Cao, J. (2021). Exploring the interaction effect of poverty concentration and transit service on highway traffic during the COVID-19 lockdown. *Journal of Transport and Land Use*, *14*(1), 1149-1164. https://doi.org/10.5198/jtlu.2021.1978
Tao, T., & Cao, J. (2023). Exploring nonlinear and collective influences of regional and local built environment characteristics on travel distances by mode. *Journal of Transport Geography*, *109*. https://doi.org/10.1016/j.jtrangeo.2023.103599
Therneau, T. M., & Atkinson, E. J. (2022). An Introduction to Recursive Partitioning Using the RPART Routines. https://cran.r-project.org/web/packages/rpart/vignettes/longintro.pdf
Van Wee, B., & Handy, S. (2016). Key research themes on urban space, scale, and sustainable urban mobility. *International Journal of Sustainable Transportation*, *10*(1), 18-24. https://doi.org/10.1080/15568318.2013.820998
Wang, K., & Ozbilen, B. (2020). Synergistic and threshold effects of telework and residential location choice on travel time allocation. *Sustain Cities Soc*, *63*, 102468. https://doi.org/10.1016/j.scs.2020.102468
Wang, X., Khattak, A., & Zhang, Y. (2013). Is Smart Growth Associated with Reductions in Carbon Dioxide Emissions? *Transportation Research Record: Journal of the Transportation Research Board*, *2375*(1), 62-70. https://doi.org/10.3141/2375-08
Wang, X., & Renne, J. L. (2023). Socioeconomics of urban travel in the U.S.: Evidence from the 2017 NHTS. *Transportation Research Part D: Transport and Environment*, *116*, 103622. https://doi.org/10.1016/j.trd.2023.103622
Wang, Y., Yang, L., Han, S., Li, C., & Ramachandra, T. V. (2017). Urban CO2 emissions in Xi'an and Bangalore by commuters: implications for controlling urban transportation carbon dioxide emissions in developing countries. *Mitigation and Adaptation Strategies for Global Change*, *22*(7), 993-1019. https://doi.org/10.1007/s11027-016-9704-1
Wu, X., Tao, T., Cao, J., Fan, Y., & Ramaswami, A. (2019). Examining threshold effects of built environment elements on travel-related carbon-dioxide emissions. *Transportation Research Part D: Transport and Environment*, *75*, 1-12. https://doi.org/10.1016/j.trd.2019.08.018
Xu, L., Cui, S., Tang, J., Yan, X., Huang, W., & Lv, H. (2018). Investigating the comparative roles of multi-source factors influencing urban residents' transportation greenhouse gas emissions. *Sci Total Environ*, *644*, 1336-1345. https://doi.org/10.1016/j.scitotenv.2018.07.072
Yang, W., & Zhou, S. (2020). Using decision tree analysis to identify the determinants of residents' CO2 emissions from different types of trips: A case study of Guangzhou, China. *Journal of Cleaner Production*, *277*, 124071. https://doi.org/https://doi.org/10.1016/j.jclepro.2020.124071