Exploring the Importance of Transportation Infrastructure and Accessibility to Satisfaction with Urban and Suburban Neighborhoods: An Application of Gradient Boosting Decision Trees

Jason Cao[1], Xinyi Wu[1]

[1] Humphrey School of Public Affairs, University of Minnesota–Twin Cities (MN)

Findings

Residential neighborhood characteristics, especially those related to transportation infrastructure and accessibility, enable the daily lives of residents and presumably enhance their satisfaction with neighborhoods. Using 2011 data in the Twin Cities, this study employs the gradient boosting decision trees approach to examine the impact on neighborhood satisfaction of transportation infrastructure and accessibility, as well as other neighborhood attributes. It also explores how residents living in urban and suburban neighborhoods value neighborhood features differently. The results show that urban residents value transportation and accessibility and suburban residents value affordability, safety, and school quality.

RESEARCH QUESTIONS AND HYPOTHESES

Residential neighborhoods are a key anchor for daily activities and travel. How transportation infrastructure and accessibility facilitate daily living is crucial to the subjective well-being of individuals (Cao 2016; Ma, Kent, and Mulley 2018; Morris, Mondschein, and Blumenberg 2018). Therefore, it is important to examine the influence of transportation and accessibility on residential neighborhood satisfaction. It is also important to recognize that other neighborhood features (such as social interaction and safety) affect neighborhood satisfaction as well (Cao and Zhang 2016; De Vos, Van Acker, and Witlox 2016; Yin et al. 2016). Understanding the relative role of transportation and accessibility in enhancing neighborhood satisfaction is an intriguing matter. Interestingly, residents living in different types of neighborhoods value different neighborhood characteristics (Cao 2008, 2015). Environmental correlates of neighborhood satisfaction should therefore vary by neighborhood type.

This study applies the gradient boosting decision trees (GBDT) approach to examine the relationship between neighborhood characteristics and neighborhood satisfaction in the Twin Cities. It aims to answer the following questions:

1. How important are transportation infrastructures and accessibility to neighborhood satisfaction?

2. How do urban and suburban residents value neighborhood characteristics differently?

METHODS AND DATA

In 2011, we administered a mail-back survey to randomly selected households living in urban and suburban neighborhoods in the Twin Cities (Figure 1). Three urban neighborhoods, mainly developed before World War II, are similar in regional location, street patterns, and transit access. By contrast, curvilinear streets are prevalent and transit services are limited in suburban neighborhoods predominantly developed in the 1970s. There were 1,303 respondents, with a response rate of 22.2%. Among them, 946 live in urban neighborhoods and 357 are from suburban neighborhoods. Refer to Cao and Wang (Cao and Wang 2016) for details on the research design and data collection.

This study uses three sets of variables from the survey:

- **Neighborhood satisfaction.** Respondents were asked to indicate how well the characteristics of their neighborhood meet the current needs of their household on a scale ranging from "extremely poorly" (1) to "extremely well" (7). This is the dependent variable.

- **Perceived neighborhood characteristics.** Respondents reported how true 30 characteristics are for their current neighborhoods (Table 1) on a scale from "not at all true" (1) to "entirely true" (4).

- **Preferred neighborhood characteristics.** We asked respondents to indicate the importance of the 30 characteristics if/when they were looking for a new place to live on a scale from "not at all important" (1) to "extremely important" (4).

If individuals perceived neighborhood characteristics incongruent with their preferences (i.e., the preferences are not met and the characteristics are mismatched), dissatisfaction with neighborhood begins to accumulate (Kahana et al. 2003). Here, we computed mismatched neighborhood characteristics (independent variables in this study) as the difference between perceptions and preferences, respectively.

Utilizing the GBDT approach, we examined the effects of mismatched neighborhood characteristics on neighborhood satisfaction, using the R-based "gbm" package (Ridgeway 2007). The tree-based ensemble can draw on insights and techniques from both statistical and machine learning methods (Friedman 2001). Compared to traditional regression, the GBDT approach has a few advantages in the context of satisfaction studies (Ding, Cao, and Næss 2018; Dong et al. 2019):

- It produces prediction that is more precise.

- It does not require data to follow a particular distribution. This feature is particularly useful because most, if not all, satisfaction variables in the literature are skewed to the left.

- It can accommodate missing variable data. The traditional listwise deletion approach may generate estimation bias if data are not missing completely at random. It also lowers statistical power by reducing sample size (Peugh and Enders 2004).

- It can address the multicollinearity problem. Multicollinearity could be an issue because some neighborhood characteristics measure a similar dimension of the built environment.

More importantly, the GBDT approach can quantify the relative importance of each independent variable in predicting response, which is a key objective of this study. However, a shortcoming of the GBDT approach is that it does not produce statistical inference.

FINDINGS

Table 1 presents the relative importance of all mismatched neighborhood characteristics to predict neighborhood satisfaction and compares important neighborhood characteristics between urban and suburban neighborhoods. All the relative importance sums to 100%.

Three neighborhood characteristics regarding transportation infrastructure and accessibility are among the top 10 important variables in urban neighborhoods. They are related to bike routes, proximity to workplace, and proximity to religious or civil buildings. By contrast, parking infrastructure is the only transportation and accessibility variable among the top 10 important characteristics in suburban neighborhoods. Overall, transportation- and accessibility-related variables are important to satisfaction with urban neighborhoods, with a collective contribution of more than 30%. However, they have a limited impact on satisfaction with suburban neighborhoods, with a collective contribution of only 12%.

Affordability and crime rate appear in the list of the top five important neighborhood characteristics for both urban and suburban neighborhoods. This highlights their critical role in affecting neighborhood satisfaction. This finding is consistent with the important determinants of residential location choice (Bina, Warburg, and Kockelman 2006; Cao 2008). On the other hand, the relative importance of affordability and crime is substantially different between urban and suburban neighborhoods.

The divergences in influential neighborhood characteristics generally reflects the different conditions of urban and suburban neighborhoods and their residents' differing needs. Because housing stock is older in urban neighborhoods than in suburban neighborhoods, housing quality and upkeep play a key role in neighborhood satisfaction in urban neighborhoods, more so than in suburban neighborhoods. Furthermore, Minneapolis is well-known for its biking culture and attracts bicyclists to reside there, so bike
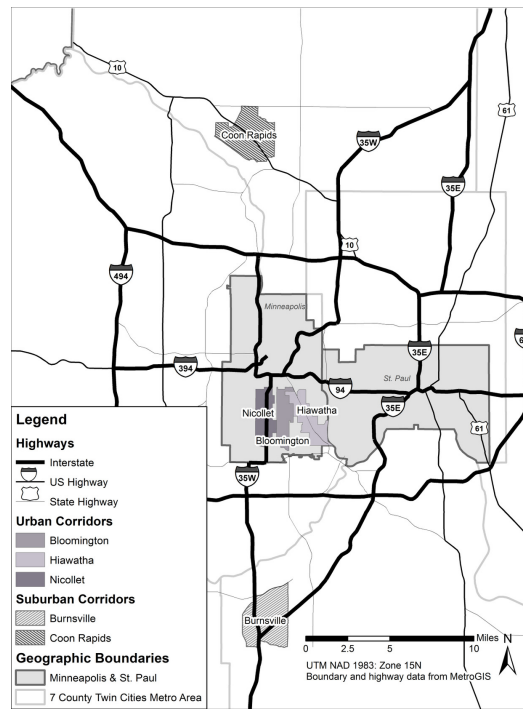
Figure 1: Location Of Urban And Suburban Corridors

infrastructure plays an important role. Suburban households have more children than urban households. Accordingly, school quality is more crucial to suburban residents than urban residents.

In summary, urban residents tend to value transportation and accessibility features of residential neighborhoods, whereas suburban residents tend to emphasize affordability, safety, and school features.

ACKNOWLEDGMENTS

Table 1: The Importance Of Neighborhood Characteristics To Neighborhood Satisfaction

Mismatched Neighborhood Characteristics	Urban Importance	Rank	Suburban Importance	Rank
High quality living unit	13.0	1	4.0	7
Good bicycle routes beyond the neighborhood	9.6	2	1.0	19
Affordable living unit	8.5	3	16.3	2
Low crime rate within neighborhood	8.4	4	18.4	1
High level of upkeep in neighborhood	7.9	5	3.0	9
Close to where I work	7.8	6	2.6	11
Lots of people out and about within the neighborhood	5.7	7	1.8	14
Religious or civic buildings (e.g., library) nearby	4.1	8	1.8	15
Quiet neighborhood	3.3	9	0.2	26
Good investment potential	3.0	10	2.4	12
Attractive appearance of neighborhood	2.9	11	1.1	16
Low level of car traffic on neighborhood streets	2.7	12	0.5	22
Parks and open spaces nearby	2.5	13	1.0	18
Economic level of neighborhoods similar to my level	2.5	14	0.8	20
Lots of interaction among neighbors	2.3	15	12.1	4
Lots of off-street parking (garages or driveways)	2.3	16	3.8	8
Safe neighborhood for kids to play outdoors	2.2	17	4.2	6
Safe neighborhood for walking	2.1	18	0.1	28
Easy access to a regional shopping mall	2.1	19	0.2	27
Large back yards	1.6	20	1.0	17
Shopping areas within walking distance	1.5	21	0.6	21
High quality K–12 schools	1.0	22	12.3	3
Good street lighting	0.9	23	0.1	30
Living unit on cul-de-sac rather than through street	0.7	24	4.4	5
Variety in housing styles	0.5	25	2.9	10
Easy access to downtown	0.3	26	0.4	23
Easy access to transit stop/station	0.2	27	0.1	29
Good public transit service (bus or rail)	0.2	28	0.2	25
Sidewalks throughout the neighborhood	0.1	29	2.2	13
Diverse neighbors in terms of ethnicity, race, and age	0.0	30	0.3	24

The unit of importance is percentage. Shaded variables are neighborhood characteristics related to transportation infrastructure and accessibility.

REFERENCES

Bina, M., V. Warburg, and K. Kockelman. 2006. "Location Choice Vis-à-Vis Transportation: Apartment Dwellers." *Transportation Research Record: Journal of the Transportation Research Board* 1977: 93–102.

Cao. 2008. "Is Alternative Development Undersupplied? Examination of Residential Preferences and Choices of Northern California Movers." *Transportation Research Record: Journal of the Transportation Research Board* 2077: 97–105.

Cao, X. 2015. "Residential Preference and Choice of Movers in Light Rail Neighborhoods in Minneapolis, Minnesota." *Transportation Research Record: Journal of the Transportation Research Board* 2494: 1–10.

———. 2016. "How Does Neighborhood Design Affect Life Satisfaction? Evidence from Twin Cities." *Travel Behaviour and Society* 5: 68–76.

Cao, X., and D. Wang. 2016. "Environmental Correlates of Residential Satisfaction: An Exploration of Mismatched Neighborhood Characteristics in the Twin Cities." *Landscape and Urban Planning* 150: 26–35.

Cao, and J. Zhang. 2016. "Built Environment, Mobility, and Quality of Life." *Travel Behaviour and Society* 5: 1–4.

De Vos, J., V. Van Acker, and F. Witlox. 2016. "Urban Sprawl: Neighbourhood Dissatisfaction and Urban Preferences." *Some Evidence from Flanders. Urban Geography* 37: 839–62.

Ding, C., X. Cao, and P. Næss. 2018. "Applying Gradient Boosting Decision Trees to Examine Non-Linear Effects of the Built Environment on Driving Distance in Oslo." *Transportation Research Part A* 110: 107–17.

Dong, Wei, Xinyu Cao, Xinyi Wu, and Yu Dong. 2019. "Examining Pedestrian Satisfaction in Gated and Open Communities: An Integration of Gradient Boosting Decision Trees and Impact-Asymmetry Analysis." *Landscape and Urban Planning* 185: 246–57.

Friedman, J.H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." *The Annals of Statistics* 29: 1189–1232.

Kahana, E., L. Lovegreen, B. Kahana, and M. Kahana. 2003. "Person, Environment, and Person–Environment Fit as Influences on Residential Satisfaction of Elders." *Environment and Behavior* 35: 434–53.

Ma, L., J. Kent, and C. Mulley. 2018. "Transport Disadvantage, Social Exclusion, and Subjective Well-Being: The Role of the Neighborhood Environment—Evidence from Sydney, Australia." *Journal of Transport and Land Use* 11: 31–47.

Morris, E.A., A. Mondschein, and E. Blumenberg. 2018. "Is Bigger Better? Metropolitan Area Population, Access, Activity Participation, and Subjective Well-Being." *Journal of Transport and Land Use* 11: 153–79.

Peugh, J.L., and C.K. Enders. 2004. "Missing Data in Educational Research: A Review of Reporting Practices and Suggestions for Improvement." *Review of Educational Research* 74: 525–56.

Ridgeway, G. 2007. "Generalized Boosted Models: A Guide to the Gbm Package."

Yin, J., X. Cao, X. Huang, and X. Cao. 2016. "Applying the IPA–Kano Model to Examine Environmental Correlates of Residential Satisfaction: A Case Study of Xi'an." *Habitat International* 53: 461–72.