Article

TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Transportation Research Record
2018, Vol. 2672(6) 31–39
© National Academy of Sciences:
Transportation Research Board 2018
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0361198118774159
journals.sagepub.com/home/trr
SAGE

Travel Behavior in TODs vs. Non-TODs: Using Cluster Analysis and Propensity Score Matching

Keunhyun Park[1], Reid Ewing[1], Brenda Case Scheer[1], and Shabnam Sifat Ara Khan[2]

Abstract

As a compact, mixed-use, and walkable district near a high-quality transit station, transit-oriented development (TOD) has arisen as a sustainable form of urbanism to minimize automobile dependency and maximize ridership. Existing travel behavior studies in the context of TOD, however, are limited in terms of small sample size, lack of consistency in TOD classification, and failure to control for residential self-selection. This study examines various travel outcomes—vehicle miles traveled (VMT), auto trips, transit trips, and walk trips—in different types of station areas in eight U.S. metropolitan areas using cluster analysis and propensity score matching. Using cluster analysis with three built environment factors—activity density, land use diversity, and street network design (i.e., D variables)—this study classifies existing 549 station areas as TOD, transit-adjacent development (TAD), and Hybrid types. After controlling for residential self-selection, the result shows that a TOD motivates its residents to walk more and take transit more while driving less. The significant difference between TOD and TAD in both VMT and the number of automobile trips means that TOD makes the personal vehicle trips shorter (39% reduction) and fewer (35% reduction). Travel behaviors in the Hybrid type are also examined for the potential outcomes of gradual and practical changes.

Expenditures on transportation have increased from the sixth largest share (less than 2%) of household budgets in 1917 to the second largest share since the 1970s (17% in 2014) (*1*). Under this circumstance, transit-oriented development (TOD) has gained popularity worldwide as a sustainable form of urbanism by concentrating developments near a transit station to minimize automobile dependency and maximize ridership. A TOD project should give people alternative transportation options and, in turn, decrease their transportation cost.

Much of the literature verifies that TODs enhance the use of public transport and reduce car usage (*2–7*). Existing TOD studies, however, have limits in terms of (1) small numbers of study sites, (2) lack of systematic methodology to distinguish TOD from other types of station areas, and (3) lack of control for the impact of residential self-selection on travel behavior. There are some exceptions to the above limitations (*8*), but no study overcomes all three limitations. As a result, it is hard to generalize the findings from the literature to planning practice. Also, the current distinctions between TODs and transit-adjacent developments (TADs), an opposite concept of TOD, limit the practical implications for transit officials and planners. Finally, when it fails to control self-selection effect, the result might overestimate the impact of TOD urban form on travel behavior.

Thus, this study asks two research questions. First, how can station area types be distinguished? Second, how do travel behaviors vary among different station area types? To answer these questions, cluster analysis is used to classify station area types and propensity score matching (PSM) is used to control for residential self-selection. For better generalizability, the data are collected from household travel surveys in eight urban areas across the U.S. with exact XY coordinates for households and trip ends. For greater policy relevance, the data are used to analyze various travel outcomes such as an automobile, transit, and walk trips at household level for different station area types. By doing so, this study seeks to examine the pure impact of living in TOD—versus a TAD—on travel behavior.

There is broad interest in the planning and policy communities for accurate tools to predict the consequences of TOD on the generation of transit ridership and reduction of automobile usage. This analysis will help guide transportation planners and decision makers to evaluate TOD projects regarding their economic, social, and environmental impacts.

[1]Department of City & Metropolitan Planning, University of Utah, Salt Lake City, UT
[2]City of Spearfish, South Dakota, Spearfish, SD

Corresponding Author:
Address correspondence to Keunhyun Park: keunhyun.park1@gmail.com

Literature Review

TOD/TAD Classification

Bernick and Cervero define TOD as "a compact, mixed-use community, centered around a transit station that, by design, invites residents, workers, and shoppers to drive their cars less and ride mass transit more" (*9*, p. 5). Kamruzzaman et al. state that a TOD is a neighborhood that is served by public transit services and offers amenities such as density, walkable, well-connected street patterns, and diversified land uses (*10*). TAD is often defined as a failure of a TOD. A TAD is a noncompact, segregated neighborhood development that calls for automobile use instead of inviting walk trips (*11–13*).

The most frequently studied factors for distinguishing a TOD from other types of station areas have been residential and employment density (*10*, *14–20*), land use diversity (*10*, *14*, *17*, *19–21*), and street connectivity (*14–16*, *21–25*). Recent studies deal with all three factors in the analysis (*10*, *14*, *17*). There are several ways to distinguish TOD from TAD, such as cluster analysis (*10*, *21*) or scoring system (*14–17*).

Existing studies differentiating TOD from TAD are limited. First, most studies cover only a single or few regions. Although Renne and Ewing study 54 regions across the U.S., the outcome variable is not overall travel behavior, but only the percentage of people who commute via public transportation (*14*). In contrast, the present study includes eight metropolitan areas in the U.S. with varying geographic and socioeconomic conditions to examine various travel outcomes. Second, unlike existing studies relying on straight-line catchment areas (*21*) or simple scoring systems (*14*), this study uses network distance from each station and cluster analysis. Finally, while Kamruzzaman et al. use a robust method of classification, their study analyzes all neighborhoods in a single city, Brisbane, Australia (*10*). Instead, this study uses the station-based approach as a focus of TOD and TAD because it deals with built environments of station areas and their impact on travel behavior, which has more direct implications for planning practice.

TOD and Travel Outcomes

Potential benefits of TOD are multiple, from promoting active modes of transportation to improving access to opportunities such as jobs or entertainment, to offering alternative mobility options and affordable housing, to reducing greenhouse gas emissions (*26*, *27*). Thus, TOD serves interrelated goals of making communities socially, economically, and environmentally more robust and sustainable.

To achieve these multiple goals, a TOD should first create settings that prompt people to drive less and ride public transit more (*3*). The Center for Transit Oriented Development identifies vehicle miles traveled (VMT) as the key performance measure for TOD (*28*). Lower VMT means that people walk, bike, and use transit more, and have more transportation options.

Much of the literature verifies that TODs enhance the use of public transport and reduce car usage (*2–7*). Based on data from 17 TOD projects, Cervero and Arrington show that residents living in TOD areas are two to five times more likely to commute by transit than their non-TOD counterparts (*29*). Nasri and Zhang find that people living in TOD areas tend to drive less, reducing their VMT by around 21–38%, compared with the residents of the non-TOD areas (*5*).

Cervero (*3*) finds evidence that many TOD ridership gains are a result of self-selection—individuals who wish to drive less may select transit-oriented environments. Many studies have found associations between attitudes and travel choices as evidence of residential self-selection (*30–32*). Thus, individuals' attitudes may confound the relationship between the TOD-type urban form and travel choices and, in turn, the effect of the built environment on travel may be overestimated (*33*).

From a review of 38 empirical studies, Cao et al. examine different methodological solutions to self-selection bias (*30*). Among the methodologies, the PSM method is highly recommended in a non-randomized observational study (*30*). The propensity score approach has recently been applied in travel behavior research (*34–38*), but rarely used in the context of station areas. The authors found only one master's thesis using PSM to explore the impact of TOD on travel mode choice (*39*). A detailed explanation of the PSM is presented in the following Research Design section.

Research Design

Study Regions

This study includes eight U.S. metropolitan regions meeting three criteria. First, they must have household travel survey data with XY coordinates for households and trip ends. Second, a region must provide land use databases at the parcel level so that land use mix can be calculated for the same years as the household travel surveys. Third, they must have had a rail-based transit system before the survey was conducted.

In the eight regions (Table 1), the household travel survey was conducted between 2006 and 2012, and there were 549 rail-based transit stations on the national TOD Database (Center for Transit Oriented Development, http://toddata.cnt.org/). Transit types include heavy rail (109 stations), commuter rail (148 stations), and light rail (272 stations). Boston has the greatest number of stations (n = 239), followed by Portland (n = 94) and Miami (n = 50), and Minneapolis-St. Paul has the least number (n = 20).

Although they were conducted by individual regions' organizations such as metropolitan planning organization (MPO), the regional household travel surveys have quite similar structure and questions, akin to U.S. DOT's National Household Travel Survey (NHTS). To gather comprehensive data on travel and transportation patterns, the survey data consistently include, but are not limited to, household demographic information, vehicle information, and data about one-way trips taken during a designated 24-hour period,

Table 1. Study Regions and Transit Stations[a]

No.	Region	Year (survey)	Heavy rail	Commuter rail	Light rail	Total	Household (½ mile)
1	Atlanta, GA	2011	38	0	0	38	138
2	Boston, MA	2011	49	121	72	239[b]	1586
3	Denver, CO	2010	0	0	36	36	152
4	Miami, FL	2009	22	4	24[c]	50	26
5	Minneapolis-St. Paul, MN	2010	0	4	16	20	97
6	Portland, OR	2011	0	7	87	94	304
7	Salt Lake City, UT	2012	0	1	36	37	114
8	Seattle, WA	2006	0	11	25	35[b]	16
	Total		109	148	272	549	2433

[a]This study includes only transit stations which had opened before the survey year.
[b]The total number of stations is not equal to the sum of the columns because some stations have two or more types of transit systems.
[c]Miami's People Mover, an automated guideway transit, is included under the LRT category.

including travel time, mode of transportation, and purpose of the trip. The survey data have exact XY coordinates, enabling geocoding of the precise locations of households and measurement of the lengths of trips, whereas the NHTS provides the location of households only at the Census Tract level. The regional survey data were acquired from individual MPOs or state DOTs with confidentiality agreements.

Data

Following the definition of TOD and the literature review, this study includes "activity density," "land use diversity," and "street network design" to classify station area types. For the density variable, population and employment data for traffic analysis zones (TAZ) were acquired from regional MPOs and summed to compute an overall activity density. Activity density is the sum of population and employment within the station area, divided by gross land area in square mile (*40*). For the diversity variable, an entropy index was computed.[1] Each region provided parcel maps, enabling calculation of the proportion of the area of each land use type—residential, commercial, and public—in a ½ mile (about 800 m) buffer from each station. For the street network design variable, the number of intersections per square mile was computed from street network shapefiles. Because these three built environment variables—activity density, land use entropy, and intersection density—vary in range, each variable was standardized to a mean of 0 and a standard deviation of 1.

In addition, a "distance to transit" variable was measured as a network distance from a household to the closest rail station, because that might be an important determinant of transit trips. Also, regional accessibility is another important variable to predict travel behaviors (*40*). That variable is defined as the percentage of jobs that can be reached within 30 minutes by transit, which tends to be highest at central locations and lowest at peripheral ones. This study used travel time skims and TAZ-level employment data acquired from regional MPOs.

Station areas were drawn as a ½-mile buffer in network distance from each station. Then, individual households were allotted to their nearest station based on network distance. The resulting pooled data set in station areas consists of 2,433 households in the eight regions. From the travel survey data in eight regions, VMT, automobile trips, transit trips, and walk trips by individual households were calculated. The survey data also include demographic variables such as household size, the number of employed, and household income.

Research Process and Methods

Step 1. TOD/TAD Classification: Cluster Analysis. Because the built environments around transit stations fall within a TOD–TAD spectrum, not a simple dichotomous scale, and there is no certain agreement of ideal built environments for TOD, identifying TODs and distinguishing them from TADs could be a difficult but important research step. Cluster analysis has been a preferred method for generating TOD typologies in previous studies (*21*, *23*, *41*).

Using cluster analysis, this study classifies station area types based on three built environment factors: activity density, land use diversity, and street network design. This approach enables grouping of existing station areas based on their actual built environment characteristics, rather than theoretical criteria of TOD or TAD. To be specific, this study uses a hierarchical clustering algorithm with Ward D2 distance measure. To determine the optimal number of clusters in a dataset, this study uses the "NbClust" package in R 3.3.1 software, which provides 26 validation indices of clustering such as Calinski and Harabasz index and Silhouette index (*42*).

Step 2. Household Sample Selection: Propensity Score Matching. PSM has been widely used to overcome nonrandom assignment of treatment in the evaluation of social programs (*43*). Evaluation studies are often based on observational data, in

Table 2. Cluster Analysis Result and Descriptive Statistics

Cluster type	Number of stations	Activity density (/sq.mi.)		Entropy index		Intersection density (/sq.mi.)	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
TAD	107	10,319	11,751	0.30	0.19	110	58
Hybrid	382	21,210	19,764	0.75	0.15	194	79
TOD	60	135,327	51,025	0.70	0.24	386	110
Total	549	31,559	43,821	0.66	0.24	199	108

Note: TAD = transit-adjacent development; TOD = transit-oriented development.

which the assignment of treatment is not random. Accordingly, individuals in the treatment group are likely to differ systematically from those in the control group. For example, households living in suburban regions could be more affluent than their counterparts in downtown, a result of residential self-selection. Therefore, the observed difference in behavioral outcomes between the groups is confounded by residential self-selection. Statistically, this generates a biased estimate of treatment effect.

The propensity score is defined as the conditional probability of assignment to a particular treatment given a vector of observed covariates (*44*). In the context of TOD and TAD, the treated group is households living in TOD station areas, whereas the control group is those living in either TAD or Hybrid areas.

The PSM was implemented in R 3.3.1 using *MatchIt* package. First, a binary logit model was developed to estimate propensity score using the subsample of households living in TOD (treatment) and TAD (control). Household characteristics were chosen as independent variables—household size, the number of workers, household income, distance to the nearest transit station, regional job accessibility, and the regions—as potential sources of residential self-selection and confounding factors in travel outcome. Second, each household living in TOD was matched with those in TAD based on the propensity score. Caliper length of 0.03 is used for matching, meaning that for a treatment observation, a match in control observations whose propensity scores are within 0.03 of the score of the treatment observation was searched (*45*). Third, whether the matched residents in TOD were systematically different from those in TAD was evaluated. A *t*-test was used to assess whether demographics and locational factors are balanced between the matched groups.

The final goal of PSM is to compute the true impact of TOD/TAD on travel behavior. Once the matching was complete, the average treatment effects (ATE) of station area type on VMT, transit trips, and walk trips were calculated. The ATE is computed as the mean travel factors of the matched TOD households minus those of the matched TAD households.

Because this study has three area types—TOD, TAD, and Hybrid, in contrast to most studies with only one treatment and one control—this study applies PSM with multiple nominal treatments (*46*). In this case, the propensity score is estimated separately for each pair of control and treatment. For each pair of groups, a binary logit model was estimated using the subsample of respondents living in the groups.

TOD/TAD Classification

Table 2 shows the result of hierarchical clustering. Using the *NbClust* package in R 3.3.1 software, thirteen of the 26 validation indices suggest that three is the optional number of clusters. The first cluster (n = 107) is labeled as “TAD” because it has the lowest level of density, diversity, and intersection density. The second and largest one (n = 382) is classified as “Hybrid”, which has a low level of activity density and intersection density, but highest entropy index. The final cluster (n = 60) is named as “TOD” in terms that it has highest activity density and intersection density, and high level of land use mix level.

Sample households were selected as those living within half-mile network distance from stations. Individual households were allotted to their nearest stations based on network distance in order to assign the station types. TAD type has 251 households, and TOD and Hybrid type have 306 and 1,876 households, respectively (Table 3).

Table 3 shows that households living in TADs have more household members, more workers, and higher incomes than those living in TODs or Hybrids. ANOVA shows that the differences are significant. Regarding travel behavior, TAD households have much higher VMT and auto trips and lower transit and walk trips than those in TODs and Hybrids. The hybrid type is in the middle, except for their lowest household incomes and the highest level of transit trips on average.

Household Sample Selection: Propensity Score Matching

The cluster analysis results show that households living in TAD tend to be more affluent, have more cars, live in a larger household, and be more auto-oriented than their counterparts in TOD. Residential self-selection theory says, however, that the households living in TAD might live there because they are auto-oriented. Therefore, the true difference in travel outcomes between TOD and TAD is estimated here by matching samples using PSM.

Table 3. Household Characteristics and Travel Behavior by Station Area Types: Average and ANOVA Analysis

Cluster type	No. of stations	HH samples	HH size	HH workers	HH income ($1000)	VMT	Auto trips	Transit trips	Walk trips
TAD	107	251	2.19	1.28	83.72	21.23	6.06	0.72	1.91
Hybrid	382	1,876	2.15	1.22	77.02	15.44	4.93	1.47	3.89
TOD	60	306	1.54	0.97	82.02	8.61	2.04	1.35	4.81
Total	549	2,433	2.07	1.19	78.34	15.18	4.68	1.37	3.80
F-statistic (ANOVA)	-	-	37.23 ***	14.1 ***	2.37 *	32.00 ***	47.19 ***	12.42 ***	30.37 ***

Note: TAD = transit-adjacent development; TOD = transit-oriented development.***: p<.01, **: p<.5, *: p<.1.

Table 4. Differences in Travel Behavior between Station Area Types after Matching

Area type pair	Travel outcomes	Observed difference	PSM ATE	ATE/observed difference ratio	Mean of control group	ATE/control ratio
TAD (control)–TOD	VMT	6.34***	5.49**	0.87	14.18	0.39
n = 564 (unmatched), n = 108 (matched)	Auto trips	4.02***	1.44**	0.36	4.06	0.35
	Transit trips	−0.64***	−0.74***	1.16	0.50	−1.48
	Walk trips	−2.86***	−1.89***	0.66	1.72	−1.10
Hybrid (control)–TOD	VMT	3.64***	0.48	0.13	9.53	0.05
n = 2,204 (unmatched), n = 350 (matched)	Auto trips	2.87***	0.84**	0.29	3.10	0.27
	Transit trips	0.11	−0.12	−1.09	1.39	−0.09
	Walk trips	−0.89***	−1.20***	1.35	3.95	−0.30
TAD (control)–Hybrid	VMT	2.70**	3.75*	1.39	23.69	0.16
n = 2,142 (unmatched), n = 364 (matched)	Auto trips	1.15***	0.00	0.00	6.70	0.00
	Transit trips	−0.75***	−0.35*	0.47	0.73	−0.48
	Walk trips	−1.97***	−0.75*	2.08	2.08	−0.36

Note: TAD = transit-adjacent development; TOD = transit-oriented development.***: p<.01, **: p<.5, *: p<.1 (t-test results).

With the six explanatory variables—household size, the number of workers, household income, distance to the nearest transit station, regional job accessibility, and the regions—household pairs in three area type pairs (TOD–TAD, TOD–Hybrid, and TAD–Hybrid) are matched. The PSM generates 54 household pairs (108 in total) in the TOD–TAD pair, 175 pairs in the TOD–Hybrid pair, and 182 pairs in the TAD–Hybrid pair.

After matching, whether the chosen residents in one type are systematically different from those in another type was evaluated. If they are different in terms of demographics, self-selection is still a concern. Unlike unmatched samples where all demographic variables are significantly different between pair groups, t-test results for matched samples show that residents in TOD and TAD do not differ by all covariates used in the PSM. Those variables are not statistically different in both TOD–Hybrid and TAD–Hybrid pairs as well (results are not shown).

Once the matching was complete, the ATE, the observed differences, and the ratio between them on VMT, auto trips, transit trips, and walk trips for each area pair were calculated. As an example for the TOD–TAD pair, the observed difference is the mean travel factors of all TOD households minus that of all TAD households in the original sample. The ATE is the difference in mean travel factors between the matched samples in TOD and TAD.

From the 3rd to 7th columns, Table 4 shows observed difference in mean in the original sample, ATE in matched sample, ratio of ATE over observed difference, mean value of control group after matching (the control group is TAD in the first and third pair and Hybrid in the second pair), and ratio of ATE over control mean, respectively. Regarding VMT in TOD–TAD pairs, after matching (i.e., controlling residential self-selection), TAD households tend to drive 5.49 miles per day more than TOD residents ("PSM ATE" column in Table 4). The significant effect on VMT of living in TOD accounts for approximately 87% of the observed influence ("ATE/observed difference ratio" column in Table 4); that is, 13% of the observed difference may result from residential self-selection.

The mean difference in automobile trips between matched TAD and TOD households is 1.44 and statistically significant. That is, if a randomly selected household moves from a TAD to a TOD, a decrease in the number of driving by 1.44 trips per day would be expected. The effect on auto trips of living in TOD accounts for approximately 36% of the observed influence. On average, the matched sample households in TAD drove 4.06 times per day. Thus, the effect of living in TOD itself represents a considerable 35% decrease (1.44 fewer trips) in daily auto trips.

In addition, the probability of walking or taking transit significantly decreases from TAD to TOD. After matching, an average household living in TOD takes 0.74 transit trips more (or 148% higher) than that in TAD. Likewise, the average household in TOD talks 1.89 walk trips more (or 110% higher) than that in TAD, and both differences are statistically significant.

When Hybrid and TOD areas are compared, the number of auto trips is significantly higher (0.84 more auto trips), and the number of walk trips is significantly lower (1.20 fewer walk trips) in Hybrid areas. The effect of living in TOD itself represents a 27% decrease in daily auto trips and a 30% increase in daily walk trips, compared with living in a Hybrid type. VMT and transit trips are not significantly different between Hybrid and TOD types.

In the case of the TAD–Hybrid pair, except auto trips which are not significant, VMT, walk trips, and transit trips are slightly significant. Compared with TAD areas, VMT is lower (3.75 more miles), the number of transit trips is higher (0.35 more transit trips), and the number of walk trips is higher (0.75 more walk trips) in Hybrid areas. The effect of living in the Hybrid type itself represents a 16% decrease in VMT, a 48% increase in daily transit trips, and a 36% increase in walk trips, compared with living in a TAD type.

From the results of travel outcomes between three pairs of station area type, the gradual and cumulative changes could be examined. When a household moves from a TAD type to a Hybrid type, or a local government improves a TAD to a Hybrid by increasing its density, land use mix, or walkability, the average household is expected to have slightly shorter auto trips and more walk trips and transit trips. Then if the household moves to a TOD, or the station area is developed to a TOD, the household is expected to have fewer auto trips and more walk trips. Cumulatively, from TAD to TOD, a household is estimated to have significantly shorter and fewer auto trips and more transit and walk trips.

Discussion and Conclusions

The clustering approach in this study classified existing station areas into TOD, TAD, and Hybrid types in terms of built environment factors—density, diversity, and street network design. As a result, 11% of the 549 stations in eight regions were labeled as TOD as being dense, diverse, and walkable. One-fifth were named as TAD as having opposite urban form of TOD. The other 70% of the stations were classified as Hybrid. Land use mix was a key factor to distinguish TAD from Hybrid, whereas density and street design played important roles to differentiate TOD and Hybrid.

Station area types vary among the literature according to classifying methods, factors, and regions. This study has an advantage as the area types are drawn from all stations in eight urban areas in the U.S., whereas the majority of studies are limited to one or few regions. Renne and Ewing (*14*) cover 54 regions across the U.S., but this study uses a more objective and systematic approach—hierarchical cluster analysis—and analyzes more travel outcome variables using disaggregate data.

Household characteristics and travel behaviors from household travel survey data were matched to each station area type, and this study found that residents living in different types are different from each other. Households in TAD tend to be more affluent, have more cars, live in a larger household, and be more auto-oriented than their counterparts in TOD. Regarding travel behavior, TAD households have much higher VMT and lower walk and transit trips than those in TODs and Hybrids. The average number of daily automobile trips shows the most dramatic differences, in that TAD households generate three times more auto trips than TOD households (6.06 vs. 2.04). The large difference in mode share between TOD and TAD (e.g., auto mode shares in TAD and TOD are 68% and 25%, respectively) is observed in other studies (*14*, *47*), some of which is much less dramatic—approximately 70% (TOD) vs. 85% (non-TOD) (*17*, *48*).

In this study, PSM enabled matching of samples to control for residential self-selection. Although the differences in travel outcomes become less dramatic after controlling self-selection, the matched sample still shows that TOD motivates its residents to walk more and take transit more while using personal vehicles less. The significant difference between TOD and TAD in both VMT and the number of auto trips means that TOD makes personal vehicle trips shorter (39% deduction) and fewer (35% deduction).

By considering the in-between hybrid type, this study may provide some practical implications. For example, when a local government and transit authority develop a TAD-type station area which is sprawled, single use, and not walkable, into a Hybrid type, mainly by adding different land uses, they could expect small increases in transit and walk trips. A Hybrid type of station area could be changed into a TOD type by adding density and decreasing block sizes, which would result in less driving and more walking by residents. The cumulative change from TAD to TOD could encourage its residents to drive less with shorter distances, walk more, and take transit more, which will have positive impacts on the city's environment, society, and economy.

This study has three main limitations. First, station area classification might generate different results if the input is changed; for example, if different regions or different built environment factors are included. The result also depends on the clustering method. The classification relying on cluster analysis is not able to provide specific guidance such as benchmark thresholds of density, land use mix, or street connectivity (for benchmark threshold examples, see Renne and colleagues [*14, 49*]). Nevertheless, the clustering approach in this study might reflect reality better than hypothetical benchmarks do. Also, the result shows that different station areas such as TOD and TAD have meaningfully different effects on the travel outcomes of their residents.

Second, PSM works only when all confounding factors are included in the analysis. This study, however, only includes the factors reflecting self-selection indirectly, and does not have residents' attitude information. The risk of not controlling all confounding factors is that the effect of residential self-selection on travel behavior might be under- or over-estimated. To the authors' knowledge, there are no such attitude data covering multiple regions in the U.S., but the result of this study needs to be checked for external validity by additional TOD studies including residential preference data in specific regions.

Third, in theory, the observed covariates in the propensity score equation are measured before the treatment whereas the outcome is measured after the treatment (*44*). In the context of this study, the data collection point for household characteristics and location factors needs to be before the station area is developed, whereas the travel outcome data should be collected after the development. This requires longitudinal data. However, because the regional household travel surveys are conducted in different years in each region, it is not plausible to put all longitudinal data into one analysis. Although this study uses cross-sectional data to control the temporal differences across regions and stations, further research needs more advanced methods, such as natural experiments with longitudinal data.

Nevertheless, as first-of-its-kind research using both cluster analysis and PSM in the exploration of TOD/TAD outcomes, this study provides evidence that a TOD, and even a Hybrid type of station area, could encourage its residents to use more active modes of transportation. An effort to create a transit-oriented neighborhood does not have to be a "megaproject." Gradual changes of a station area into a denser, more diverse, and more walkable environment would compensate in the form of sustainable travel behavior, which ultimately provides more environmental, social, and economic benefits.

Acknowledgment

This research was funded by the National Institute for Transportation and Communities (NITC) under grant number 859, a program of the Transportation Research and Education Center at Portland State University and a U.S. Department of Transportation university transportation center.

Author Contributions

The authors confirm contribution to the paper as follows. Study conception and design: Keunhyun Park, Reid Ewing, Brenda Scheer; data collection: Keunhyun Park, Shabnam Khan; analysis and interpretation of results: Keunhyun Park, Reid Ewing; draft manuscript preparation: Keunhyun Park. All authors reviewed the results and approved the final version of the manuscript.

Note

1. The entropy index measures balance between three different land uses. The index ranges from 0, where all land is in a single use, to 1 where land is evenly divided among the three uses. Values are intermediate when buffers have more than one use but one use predominates. The entropy calculation is: entropy = – [residential share * ln(residential share) + commercial share * ln(commercial share) + public share * ln(public share)]/ln(3), where ln is the natural logarithm of the value in parentheses and the shares are measured in terms of total parcel land areas (*40*).

References

1. U.S. Bureau of Labor Statistics. *Consumer Expenditure Survey*. Inter-university Consortium for Political and Social Research, Ann Arbor, Mich., 2014.
2. Cervero, R. *Ridership Impacts of Transit-Focused Development in California*. Monograph 45. Institute of Urban and Regional Development, University of California, Berkeley, 1993.
3. Cervero, R. *TCRP Report 102: Transit-Oriented Development in the United States: Experiences, Challenges, and Prospects* Transportation Research Board of the National Academies, Washington, D.C., 2004.
4. Langlois, M., D. van Lierop, R. A. Wasfi, and A. M. El-Geneidy. Chasing Sustainability: Do New Transit-Oriented Development Residents Adopt More Sustainable Modes of Transportation? *Transportation Research Record: Journal of the Transportation Research Board*, 2015. 2531: 83–92.
5. Nasri, A., and L. Zhang. The Analysis of Transit-Oriented Development (TOD) in Washington, D.C. and Baltimore Metropolitan Areas. *Transport Policy*, Vol. 32, No. 1, 2014, pp. 172–179.
6. Olaru, D., and C. Curtis. Designing TOD Precincts: Accessibility and Travel Patterns. *Planning Practice and Research*, Vol. 15, No. 1, 2015, pp. 1567–7141.
7. Venigalla, M., and A. Faghri. A Quick-Response Discrete Transit-Share Model for Transit-Oriented Developments. *Journal of Public Transportation*, Vol. 18, No. 3, 2015, pp. 107–123.
8. Renne, J. L., T. Tolford, S. Hamidi, and R. Ewing. The Cost and Affordability Paradox of Transit-Oriented Development: A Comparison of Housing and Transportation Costs across Transit-Oriented Development, Hybrid and Transit-Adjacent Development Station Typologies. *Housing Policy Debate*, Vol. 26, No. 4–5, 2016, pp. 819–834.
9. Bernick, M., and R. Cervero. *Transit Villages in the 21st Century*. McGraw-Hill, New York, 1997.
10. Kamruzzaman, M., F. M. Shatu, J. Hine, and G. Turrell. Commuting Mode Choice in Transit Oriented Development: Disentangling the Effects of Competitive Neighbourhoods, Travel Attitudes, and Self-Selection. *Transport Policy*, No. 42, 2015, pp. 187–196.
11. Belzer, D., and G. Autler. *Transit Oriented Development: Moving from Rhetoric to Reality*. Brookings Institution Center on Urban and Metropolitan Policy, Washington, D.C., 2002.
12. Cervero, R., and M. Duncan. *Residential Self Selection and Rail Commuting: A Nested Logit Analysis*. Working Paper 604. University of California Transportation Center, Berkeley, 2002.

13. Dittmar, H., and G. Ohland. *The New Transit Town: Best Practices in Transit Oriented Development*. Island Press, Washington, D.C., 2004.
14. Renne, J. L., and R. Ewing. *Transit-Oriented Development: An Examination of America's Transit Precincts in 2000 & 2010*. UNOTI Publications Paper No. 17. University of New Orleans, 2013.
15. Laaly, S. *New Definition of Transit Oriented Development (TOD) Based on Context Sensitive Paradigm*. Doctoral dissertation. Morgan State University, 2014.
16. Pollack, S., A. Gartsman, M. Boston, A. Benedict, and J. Wood. Rating the Performance of Station Areas for Effective and Equitable Transit-Oriented Development. Presented at 93rd Annual Meeting of the Transportation Research Board, Washington, D.C., 2014.
17. Jeihani, M., L. Zhang, A. Ardeshiri, A. Amiri, A. Nasri, K. R. Zamir, and B. Baghaei. *Development of a Framework for Transit-Oriented Development (TOD)*. No. MD-13-SP209B4N. Maryland State Highway Administration Office of Policy and Research, Baltimore, 2013.
18. Canepa, B. Bursting the Bubble: Determining the Transit-Oriented Development's Walkable Limits. *Transportation Research Record: Journal of the Transportation Research Board*, 2007. 1992: 28–34.
19. Cervero, R., and K. Kockelman. Travel Demand and the 3Ds: Density, Diversity, and Design. *Transportation Research Part D: Transport and Environment*, Vol. 2, No. 3, 1997, pp. 199–219.
20. Cervero, R., and R. Gorham, R. Commuting in Transit versus Automobile Neighborhoods. *Journal of the American Planning Association*, Vol. 61, No. 2, 1995, pp. 210–225.
21. Vale, D. S. Transit-Oriented Development, Integration of Land Use and Transport, and Pedestrian Accessibility: Combining Node-Place Model with Pedestrian Shed Ratio to Evaluate and Classify Station Areas in Lisbon. *Journal of Transport Geography*, No. 45, 2015, pp. 70–80.
22. Ngo, V. Identifying Areas for Transit-Oriented Development in Vancouver. *Trail Six: An Undergraduate Journal of Geography*, Vol. 6, 2012, pp. 91–102.
23. Kamruzzaman, M., D. Baker, S. Washington, and G. Turrell. Advance Transit Oriented Development Typology: Case Study in Brisbane, Australia. *Journal of Transport Geography*, Vol. 34, 2014, pp. 54–70.
24. Brown, B. B., and C. M. Werner. The Residents' Benefits and Concerns Before and After a New Rail Stop: Do Residents Get What They Expect? *Environment and Behavior*, Vol. 43, 2011, pp. 789–806.
25. Werner, C. M., B. B. Brown, and J. Gallimore. Light Rail Use is More Likely on ''Walkable'' Blocks: Further Support for Using Micro-Level Environmental Audit Measures. *Journal of Environmental Psychology*, Vol. 30, No. 2, 2010, pp. 206–214.
26. Center for Transit-Oriented Development. *TOD 204 Planning for TOD at the Regional Scale: The Big Picture*. Report TOD 204. Reconnecting America, Washington, D.C., 2011.
27. Noland, R. B., K. Ozbay, S. Dipetrillo, and S. Iyer. *Measuring Benefits of Transit Oriented Development*. Mineta National Transit Research Consortium Report 12-18. Mineta National Transit Research Cosortium, Mineta Transportation Institute, College of Business, San Jose State University, 2014.
28. Center for Transit-Oriented Development. *Performance-Based Transit-Oriented Development Typology Guidebook*. Report TOD 204. Center for Transit-Oriented Development, Austin, 2010.
29. Cervero, R., and G. B. Arrington. Vehicle Trip Reduction Impacts of Transit-Oriented Housing. *Journal of Public Transportation*, Vol. 11, No. 3, 2008, pp. 1–17.
30. Cao, X. J., P. L. Mokhtarian, and S. L. Handy. Examining the Impacts of Residential Self-Selection on Travel Behaviour: A Focus on Empirical Findings. *Transport Reviews*, Vol. 29, No. 3, 2009, pp. 359–395.
31. Mokhtarian, P. L., and C. Xinyu. Examining the Impacts of Residential Self-Selection on Travel Behavior: A Focus on Methodologies. *Transportation Research Part B: Methodological*, Vol. 42, No. 3, 2008, pp. 204–228.
32. Handy, S. *Critical Assessment of the Literature of the Relationships among Transportation, Land Use and Physical Activity*. Resource paper for: Does the Built Environment Influence Physical Activity? Examining the Evidence—Special Report 282. National Academies Press, Washington, D.C., 2005.
33. Ewing, R., S. Hamidi, and J. B. Grace. Compact Development and VMT—Environmental Determinism, Self-Selection, or Some of Both? *Environment and Planning B: Planning and Design*, Vol. 43, No. 4, 2016, pp. 737–755.
34. Boer, R., Y. Zheng, A. Overton, G. K. Ridgeway, and D. A. Cohen. Neighborhood Design and Walking Trips in ten US Metropolitan Areas. *American Journal of Preventive Medicine*, Vol. 32, No. 4, 2007, pp. 298–304.
35. Cao, X. J. Exploring Causal Effects of Neighborhood Type on Walking Behavior Using Stratification on the Propensity Score. *Environment and Planning A*, Vol. 42, No. 2, 2010, pp. 487–504.
36. Cao, X. J., Z. Xu, and Y. Fan. Exploring the Connections among Residential Location, Self-Selection, and Driving: Propensity Score Matching with Multiple Treatments. *Transportation Research Part A: Policy and Practice*, Vol. 44, No. 10, 2010, pp. 797–805.
37. Cao, X. J., and Y. Fan. Exploring the Influences of Density on Travel Behavior Using Propensity Score Matching. *Environment and Planning B: Planning and Design*, Vol. 39, No. 3, 2012, pp. 459–470.
38. Cao, X. J., and J. Schoner. The Influence of Light Rail Transit on Transit Use: An Exploration of Station Area Residents along the Hiawatha Line in Minneapolis. *Transportation Research Part A: Policy and Practice*, Vol. 59, 2014, pp. 134–143.
39. Baghaei, B. *Impact of Transit Oriented Development on Mode Choice with Consideration for Multimodal Accessibility and Self-Selection*. Master of Science thesis, University of Maryland, College Park, 2014.
40. Ewing, R., G. Tian, J. P. Goates, M. Zhang, M. J. Greenwald, A. Joyce, J. Kircher, and W. Greene. Varying Influences of the Built Environment on Household Travel in 15 Diverse Regions of the United States. *Urban Studies*, Vol. 52, No. 13. 2015, pp. 2330–2348.
41. Atkinson-Palombo, C., and M. J. Kuby. The Geography of Advance Transit-Oriented Development in Metropolitan Phoenix, Arizona, 2000–2007. *Journal of Transport Geography*, Vol. 19, No. 2, 2011, pp. 189–199.

42. Charrad, M., N. Ghazzali, V. Boiteau, A. Niknafs, and M. M. Charrad. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software*, Vol. 61, No. 6, 2014, pp. 1–36.
43. Oakes, J. M., and P. J. Johnson. Propensity Score Matching for Social Epidemiology. *Methods in Social Epidemiology*, No. 1, 2006, pp. 370–393.
44. Rosenbaum, P. R., and D. B. Rubin. Reducing Bias in Observational Studies Using Subclassification on the Propensity Score. *Journal of the American Statistical Association*, Vol. 79, No. 387, 1984, pp. 516–524.
45. Austin, P. C. Optimal Caliper Widths for Propensity-Score Matching When Estimating Differences in Means and Differences in Proportions in Observational Studies. *Pharmaceutical Statistics*, Vol. 10, No. 2, 2011, pp. 150–161.
46. Lechner, M. Program Heterogeneity and Propensity Score Matching: An Application to the Evaluation of Active Labor Market Policies. *Review of Economics and Statistics*, Vol. 84, No. 2, 2002, pp. 205–220.
47. Renne, J. *Transit Oriented Development: Measuring Benefits, Analyzing Trends, and Evaluating Policy*. PhD Dissertation. Rutgers University, New Brunswick, NJ, 2005.
48. Kamruzzaman, M., D. Baker, S. Washington, and G. Turrell. Residential Dissonance and Mode Choice. *Journal of Transport Geography*, Vol. 33, 2013, pp. 12–28.
49. Renne, J. L., T. Tolford, S. Hamidi, and R. Ewing. The Cost and Affordability Paradox of Transit-Oriented Development: A Comparison of Housing and Transportation Costs Across Transit-Oriented Development, Hybrid and Transit-Adjacent Development Station Typologies. *Housing Policy Debate*, Vol. 26, No. 4–5, 2016, pp. 819–834.

The Standing Committee on Transportation and Land Development (ADD30) peer-reviewed this paper (18-00768).