Journal of Transport Geography 87 (2020) 102786

Contents lists available at ScienceDirect

Journal of Transport Geography

journal homepage: www.elsevier.com/locate/jtrangeo

Addressing transit mode location bias in built environment-transit mode use research

Laura Aston[a,*], Graham Currie[a], Md. Kamruzzaman[b], Alexa Delbosc[a], Nicholas Fournier[c], David Teller[d]

[a] *Department of Civil Engineering, Monash University, Melbourne, Australia*
[b] *Monash Art Design and Architecture, Monash University, Melbourne, Australia*
[c] *University of California, Berkeley, CA, United States*
[d] *Department of Transport, State of Victoria, Australia*

ARTICLE INFO

Keywords:
Public transport
Transit
Built environment
Forecasting
Matching
Multimodal

ABSTRACT

Many studies have identified links between the built environment (BE) and transit use. However, little is known about whether the BE predictors of bus, train, tram and other transit modes are different. Studies to date typically analyze modes in combination; or analyze one mode at a time. A major barrier to comparing BE impacts on modes is the difference in the types of locations that tend to be serviced by each mode. A method is needed to account for this 'mode location bias' in order to draw robust comparison of the predictors of each mode.

This study addresses this gap using data from Melbourne, Australia where three types of public transport modes (train, tram, bus) operate in tandem. Two approaches are applied to mitigate mode location bias: a) Co-located sampling – estimating ridership of different modes that are located in the same place; and b) Stratified BE sampling – observations are sampled from subcategories with similar BE characteristics.

Regression analyses using both methods show that the BE variables impacting ridership vary by mode. Results from both samples suggest there are two common BE factors between tram and train, and between tram and bus; and three common BE factors between train and bus. The remaining BE predictors – three for train and tram and one for bus - are unique to each mode. The study's design makes it possible to confirm this finding is valid irrespective of the type of locations serviced by modes. This suggests planning and forecasting should consider the specific associations of different modes to their surrounding land use to accurately predict and match transit supply and demand. The Stratified sampling approach is recommended for treating location bias in future mode comparison, because it explains more ridership variability and offers a transferrable approach to generating representative samples.

1. Introduction

The built environment (BE) influences transit demand by determining the locations to which individuals travel and the ease with which they can be accessed by transit (Cervero et al., 2002; Litman and Steele, 2017; Mitchell and Rapkin, 1954). However little research explores how the BE affects ridership for different transit modes (e.g. heavy/light rail and bus transit). Research in this area faces a major methodological problem: 'mode location bias'. Transit modes have cost and capacity characteristics and a development history which means some modes always tend to operate in areas of cities with specific BE characteristics.This represents a sample bias when trying to understand how ridership of each transit modes relates to all BE features. A new methodology is needed to objectively explore how each BE feature might affect ridership for each mode, whatever spatial areas they service.

This study develops and applies two new methods to determine how BE attributes link to ridership of individual transit modes accounting for mode location bias. In doing this, the research seeks to establish if different BE factors are associated with ridership for train, tram and bus after mode location bias is addressed.

The remainder of this paper is structured as follows; previous research literature on BE and transit use is outlined including research on mode location bias and how bias is generallyaddressed in the literature.

* Corresponding author: Public Transport Research Group, Monash Institute of Transport Studies, Department of Civil Engineering, Monash University, Clayton, Victoria 3800, Australia.
E-mail address: laura.aston@monash.edu (L. Aston).

https://doi.org/10.1016/j.jtrangeo.2020.102786
Received 9 March 2020; Received in revised form 22 June 2020; Accepted 24 June 2020
Available online 17 July 2020

Methodology is then described including the two proposed methods addressing bias. Results are described and then discussed. Conclusions outline findings and implications for research and practice.

2. Literature review

2.1. Built environment, transit use and transit modes

Empirical research finds an important role for the BE in predicting transit usage (Boulange et al., 2017; Ewing and Cervero, 2010). However almost all studies in the field explore how BE influences transit as a whole. Yet transit modes are distinctive with respect to speed, reliability and space. The priority afforded to different transit modes is a defining characteristic (Kittelson and Associates Inc. et al., 2013). Buses typically have no fixed right of way infrastructure and flexible routing on roads which renders them more cost effective and flexible than other modes that are fully segregated from traffic, such as light rail and heavy rail. The latter generally have a speed and reliability advantage over buses. These differences impact perceptions about transit but it is unclear whether they are influenced by the BE associated with them. This is a consideration worth exploring because it would have implications for efficient public transport planning.

2.2. The transit mode location bias problem

A major barrier to researching individual transit modes is the difficulty of obtaining an unbiased sample when modes of a specific type tend to always locate in specific parts of a city. Modes serve different functions and gain cost competitiveness in different urban environments (Kittelson & Associates Inc. et al. 2013). Buses operate throughout urban areas, providing a link to low-density environments in which it is not feasible to operate fixed-rail modes. For example, in Melbourne, the research setting for this study, the walkable catchment area for trams tend to be the densest, followed by trains and then buses. This coincides with an average proximity (based on Euclidian distance) to the city's strong central core that is much closer for trams (6.26 km) than for train (16.9 km) and bus (22.9 km) These patterns reflect deliberate policy decisions about where to supply each mode, and the zoning controls applied in proximity to each (Department of Infrastructure, 2002) (Fig. 1).

A few studies have differentiated findings for more than one transit mode. A study which developed separate ridership models for light rail, commuter rail and metro in Maryland found different results for each mode (Liu et al., 2016). Numerous BE factors, including employment density and whether or not the station was located in the CBD, were strong predictors of light rail ridership, whereas feeder bus frequency was the only significant predictor of commuter rail. Interestingly, the study found that combining predictions for light rail and metro increased the explanatory power of the ridership model. Another study develops models for bus, metro and combined transit ridership in Shenzhen, China (Tu et al., 2018). This model also has greater explanatory power for combined modes, with subtle differences in the predictors of bus and metro. Land use mix is negatively associated with bus ridership, whereas it is positively associated with metro ridership. Do the BE-ridership associations in these studies result from the BE or the cost, competitiveness, capacity or historical context? The concentration of modes in certain locations represents a ‘transit mode location bias’, which means it is difficult to understand the specific effects of all BE types on individual modes. A method is needed to obtain an unbiased sample with respect to latent demand for each mode.

2.3. Bias mitigation approaches

Imbalance in BE patterns associated with different transit modes resembles the sample-selection issue that affects treatment and non-treatment groups in experimental research. When an imbalance exists between groups, estimates are highly sensitive to the characteristics of the sample and independent variables; and may not be reproducible in other settings (Stuart, 2010). To overcome this bias, the sample must comprise sites for which the BE is observationally similar.

Bias caused by interaction or endogeneity of variables in travel demand models is often treated using instrumental variables or interaction terms (Estupiñán and Rodríguez, 2008; Taylor et al., 2009). However, because the present study seeks to examine the levels of the biased variable (transit modes) separately, an approach is needed which ‘balances’ the samples of the three groups.

Imbalance between subgroups is commonly addressed using ‘matching’, to improve the balance of independent variables of interest (‘covariates’) (Ho et al., 2007; Stuart, 2010). Matching techniques include one-to-one matching of the most similar members of each group; and subclassification methods that group similar types of observations together (Stuart, 2010).

Several conditions must be satisfied to ensure matching does not introduce new bias into the sample. The sample can be chosen on observations related to the independent variables themselves, but not the outcome (Ho et al., 2007). All imbalanced covariates should ideally be included as inputs to the matching solution, or evaluated for improved balance in the solution even if not used as an input (Stuart, 2010). The objective of matching is to reduce disparities in the independent variables between groups for comparison. Successful matching is judged based on the extent to which ‘balance’ is reduced, rather than on statistical measures of variance (Ho et al., 2007).

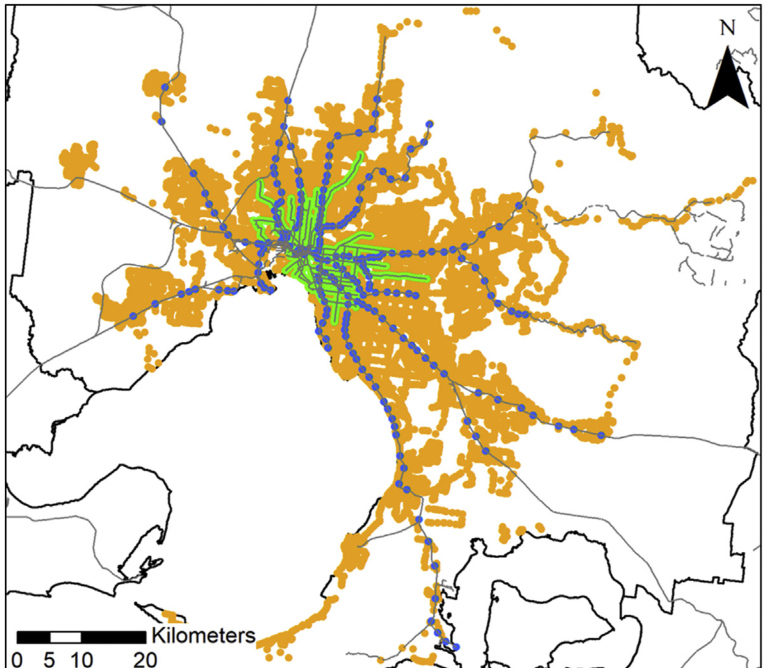
Recruitment bias is a well-established problem in individual-level studies, with relevance to travel behaviour where individual travel preferences impact residential location choices (Cao et al., 2009; De Gruyter, 2017). Self-selection of individuals preferring non-auto travel into transit-oriented neighborhoods can confound findings regarding the impact of BE interventions on travel behaviour (Cao et al., 2009; De Vos and Ettema, 2020). To account for this, a one-to-one method called Propensity Score Matching (PSM) matches residents in ‘intervention’ and ‘non-intervention’ sites based on attitudes and sociodemographic characteristics associated with joint residential-travel decision making (Cao and Fan, 2012; De Gruyter, 2017; Deng and Yan, 2019). The probability of living in a case site (receiving the intervention) is thus equalized among the samples (Heinrich et al., 2010).

The present example is concerned with a continuous outcome variable (ridership) and a ‘treatment’ (land use) that is also measured on a continuous scale. Subclassification methods are more suited to addressing imbalance for continuous interventions (Stuart, 2010). Subclassification encompasses matching techniques that group similar observations, rather than drawing one-to-one matches. Cluster analysis is one method that can be used to group observations to maximize similarity with respect to variables of interest (Hair et al., 2014). It has been applied to characterize the BE associated with transit (Voulgaris et al., 2017) and to develop transit-area typologies to aid site selection in transport planning (Jeffrey et al., 2019; Kamruzzaman et al., 2014).

3. Method

3.1. Research setting

This study aims to determine whether the BE attributes that predict ridership differ by mode after adjusting for mode location bias. The sample for this study was derived from metropolitan Melbourne's network of almost 20,000 transit stops (Fig. 1) (Public Transport Victoria, 2018b). Melbourne, Australia, has been the subject of studies that explore the relationship between the BE and travel behaviour (Boulange et al., 2017; Jeffrey et al., 2019). Melbourne is also a city with a diverse modal mix, featuring extensive heavy rail, light rail (tram) and bus networks. It is therefore an appropriate location to investigate the association between the BE and demand for different public transport modes.



	Bus	Tram	Train
Number of facilities*	9,483	925	220
Radius (m)	400	600	800
	Ave	**Ave**	**Ave**
Ridership	28.1	697	3,650
Population density	2,450	4,980	3,020
Commercial density	0.09	0.20	0.15
Jobs-housing balance	0.08	0.15	0.10
Land use diversity	0.32	0.44	0.47
Housing diversity	5.11	7.56	6.88
Ped. connectivity (density)	81.2	115	89.7
Cycle connectivity	0.08	0.21	0.15
Destination score	2.62	5.39	5.44
Distance. to CBD	22.9	6.26	16.9
Dist. to Activity Center	2.39	1.12	1.42
Count of Activity Centers	0.04	0.25	0.34
Access to employment ($\times 10^{-3}$)	23.9	230	177
Proportion 'urban'	0.78	0.80	0.80
Level of service	23.0	67.4	108
Prop. FTE (censored 95%)	0.57	0.59	0.58
Means size (censored 95%)	2.88	2.52	2.66
Proportion overseas born	0.04	0.04	0.04
Proportion tertiary educated	0.40	0.61	0.50

* stop pairs and interchanges grouped as single 'facility'

Fig. 1. Distribution of transit facilities and BE characteristics by mode in Melbourne.

3.2. Unit of analysis

Each observation in the analysis pertains to a 'facility'. Each facility has at its centre a transit stop, bidirectional stop pair, or series of stops at an interchange, grouped to eliminate significant overlap effects. This was achieved by iteratively grouping bus stops and tram stops within a certain distance of each other until the number of stops was approximately halved (bi-directional stop-pairing). This distance was found to be 25 m for trams and 50 m for bus. Following consolidation of proximate stops and removal of those with incomplete data, 18,000 bus stops were reduced to 9483 'facilities'; and 1700 tram stops were reduced to 925.

Variables are measured at different units of analysis depending on the relevant proximity at which a variable impacts ridership. Transit-oriented design uses a convention of 'zones' of different radii around transit stop to prioritse different design features for optimal station area planning (Monzón et al., 2016). These zones form the relevant unit for examining each variable and its impact on ridership. Four different units of analysis are used in this study: the facility, intermodal transfer zone, walkable neighborhood Catchment; (also known as access/egress zone), and the network or Region. Nasri and Zhang (2019) adopt a similar approach, collecting variables at three levels: the half-mile station area, census block groups, and metropolitan scales.

3.3. Variables

The present study draws on 20 BE and seven external variables (Table 1), chosen following a comprehensive review of prior literature, to represent distinctive aspects of the built environment for which robust and reproducible indicators could be defined.

The sections that follow describe BE variables used in this study as well as seven variables to represent demographic and service level factors which are important explanatory factors affecting ridership (Ewing and Cervero, 2010). Each variable is explained in relation to the most relevant unit of analysis.

Table 1
Variable descriptions and data sources.

Variables	Description	Data source
Facility level		
Ridership	average normal weekday boardings	(Department of Transport, 2019a, 2019b)
CBD (Distance)	km; Euclidian (Melbourne city center)	
Activity Center (Distance)	km, Euclidian (Plan Melbourne 'Major Activity Centers')	(Department of Environment, Land, Water and Planning, 2017)
Level of service	average weekday departures (6 am – 7 pm)	(Public Transport Victoria, 2018a)
Transfer zone level		
Cycle connectivity	length of principal bicycle network paths within 3 km from?	(VicRoads, 2017)
Bicycle facilities	dummy: 1 = presence of at least 1 bicycle cage ('Parkiteer')	(Bicycle Network Victoria, 2019)
Car parking	km^2 (formal station car parking facilities only)	(Public Transport Victoria, 2019)
Overlapping level of service	average weekday departures (6 am – 7 pm) of overlapping services	(Public Transport Victoria, 2018a)
Neighborhood (access/egress catchment) level		
Employment density	number of workers/ km^2	(Australian Bureau of Statistics, 2017g)
Population density	number of residents/ km^2	(Australian Bureau of Statistics, 2017b)
Dwelling density	number of dwellings/ km^2	(Australian Bureau of Statistics, 2017c)
Activity density	(employment + population) / km^2	
Commercial density	proportion of land zoned as commercial	(Department of Environment Land Water and Planning, 2018)
Retail worker density	number of retail workers/ km^2	(Australian Bureau of Statistics, 2017h)
Jobs-housing balance	$1 - \frac{ABS(retail\ workers - residents)}{retail\ workers + residents}$	(Australian Bureau of Statistics, 2017b, 2017h)
Land use diversity	$-\frac{\sum_k [(p_i)(ln\, p_i)]}{ln_k}$ (k = 7 different land use types)	(Department of Environment Land Water and Planning, 2018)
Housing diversity	score out of 8 for housing types present[1]	(Australian Bureau of Statistics, 2017d)
Pedestrian connectivity	density of 3-or-more-way intersections	(Do Infrastructure, 2017a)
Destination score	score out of 8 for destination types present[2]	(GeoFabrik downloads, 2019; PSMA Australia Limited, 2018)
Destination count	count of destinations present[2]	
Proportion 'urban'	proportion of catchment occupied by 'urban' land uses[3]	(Department of Environment Land Water and Planning, 2018)
Activity Centers (Count)	count of 'Major Activity Centers' within catchment	(Department of Environment, Land, Water and Planning, 2017)
Proportion FTE	proportion of residents full time employed	(Australian Bureau of Statistics, 2017e)
Mean household size	residents/dwelling	(Australian Bureau of Statistics, 2017b, 2017c)
Proportion overseas born	Proportion of residents born overseas	(Australian Bureau of Statistics, 2017a)
Proportion tertiary educated	proportion residents with bachelor's degree or higher	(Australian Bureau of Statistics, 2017f)
Regional (network) Level		
Access to employment	sum of jobs (employed persons) accessible by PT in 30 min	(Australian Bureau of Statistics, 2017g; Public Transport Victoria, 2018a)

1- Appendix 2 of the variable aggregation procedure (see note 2) contains the classification of Points of Interest and Open Street Map Features of Interest into destination types.
2- Appendix 3 of the variable aggregation procedure (see note 2) contains the classification of land use zones into seven types relevant to entropy calculation, and urban land use types.
3- Housing types: Separate house, semi- detached (1 story), semi-detached (2+ story), flat or apartment (1–2 story block), flat or apartment (3 story block), flat or apartment (4+ story block), flat or apartment attached to a house, other.

3.3.1. Outcome variable

The outcome variable, transit ridership, was quantified as the average weekday boardings at facilities, according to smartcard 'touch-ons'. Average daily boardings for 2018, estimated from myki smart-card data and adjusted for non touch-ons, were used for tram and train, while average daily touch-ons (unadjusted) from August–November 2018 were used for bus (Department of Transport, 2019a, 2019b). Due to the on-board, unsupervised nature of touch-ons for trams, data was assigned to clusters of several tram stops. For the purpose of this study, the ridership for individual stops was estimated as the average ridership at each stop within respective groupings. Preliminary statistical analysis revealed a non-linear relationship between transit ridership and the covariates, which was resolved by taking the natural logarithm of ridership estimates.

3.3.2. Facility-level variables

Service frequency, or level of service (LOS) is a property of the facility. The level of service ('LOS') was measured by the frequency of service, extracted from GTFS. The LOS is represented by the total daily weekday departures between 6 am and 7 pm. Where sample locations comprised more than one stop (bidirectional stop pairs and stops at interchanges), ridership and LOS for individual stop IDs was summed to give the total. Distance to CBD and Distance to Activity Centers are also measured from the transit stop.

3.3.3. Transfer zone variables

The transfer zone represents a small radial catchment in which transfers to other services are feasible. A transfer zone threshold was set to 160 m, representing approximately half the distance of the minimum bus stop spacing in Melbourne and 20% of the commonly accepted walking distance to heavy rail of 800 m (Cervero et al., 2004; Department of Transport, 2013). Railway station car parking, bicycle facilities and other transit facilities within a 160-m (Euclidian) radius of the transit facility were identified. The length of principal bicycle network paths within 3 km of the facility were also measured, the midpoint of the range of accepted cycle distances to transit in the literature (Kager et al., 2016; Rijsman et al., 2019).

3.3.4. Neighborhood-level variables

Indicators with influence at the neighborhood level were measured within the bounds of walkable street network catchments. Walking catchments were used in preference to Euclidian catchments to account for the actual routing available to pedestrians (Boulange et al., 2017). Walking catchments were defined by centerlines corresponding to Melbourne's road network. Highways and freeways were removed, as these are considered barriers to pedestrian movement (Mavoa et al., 2015). Sensitivity tests were conducted on different catchment radii for each mode to test for significant differences in results. These tests showed little impact of catchment size on associations or predictive strength of the models, a finding that is supported by published research (Guerra et al., 2012). To define the BE 'Strata' using cluster

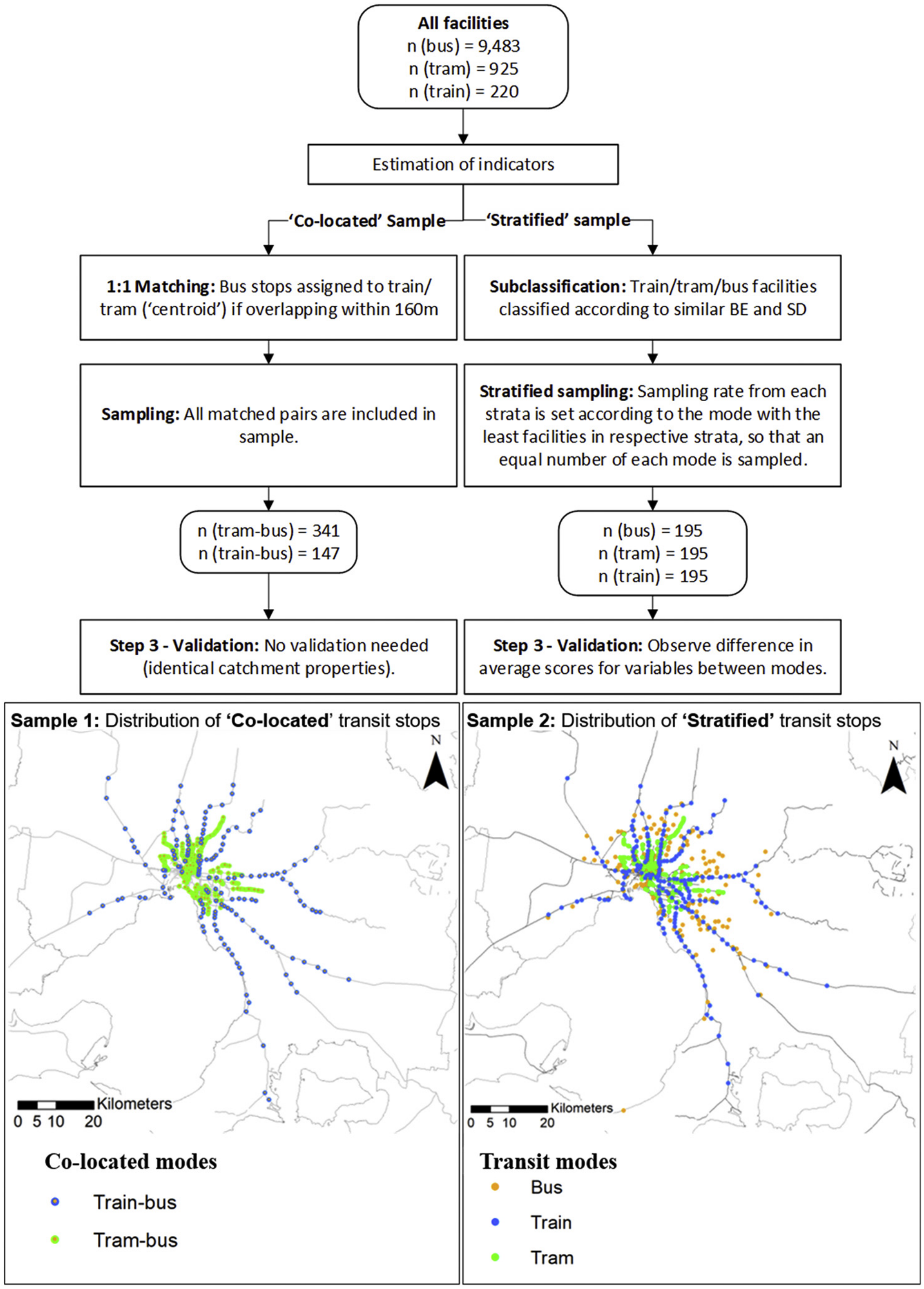


Fig. 2. Sampling procedure and sample distributions for the mode location bias mitigation methods.

analysis, a consistent walk catchment of radius 800 m was used for all facilities. For analysis, policy-relevant walking distances of 400 m for bus, 600 m for tram and 800 m for train were used as the catchment raddi for the Biased and Stratified samples (Higgs et al., 2019). For the Co-located sample, radii of 600 m for tram-bus pairs, and 800 m for train-bus were used.

Land use diversity was estimated using Shannon's entropy formula with seven land use types (Cervero, 2002; Shannon, 1948). Housing diversity was estimated as a count of the number of eight housing types present in the catchment. Local access measures were estimated according to the number of places of interest, or local services and amenities, there were in a catchment. The first indicator was a score out of eight according to how many destination types were present in the catchment. The second was the sum of places, representing "opportunity density". The eight categories were community services, convenience, child and maternal care, education, food, gealth and social services, sport and public transport. These categories were based on the local living scores used in recent empirical investigations for Melbourne (Boulange et al., 2017; Higgs et al., 2019), but adapted to the land use types specified in the two sources from which point source data was derived (GeoFabrik downloads, 2019; PSMA Australia Limited 2018).

Pedestrian connectivity can be apprixmated as the number of three-or-more way intersections, representing pedestrian crossing oppotuinties, within the access and egress catchment (Badland et al., 2017). This was measured by identifying junctions of road centerline shapefiles.

Sociodemographic variables relevant to modal preferences were also derived from Census data, and aggregated to the catchment level. Extreme outlying values were present in the estimates for the proportion of full time employed residents, and the average household size.

These indicators were censored at 95th percentile value.

3.3.5. Regional-level variables

The cumulative access to employment across the entire metropolitan area was estimated to account for the network connectivity of each transit stop. The ArcGIS network analyst feature was used to identify the spatial extent of coverage provided by Melbourne's transit network (ESRI, 2019). Service areas were generated corresponding to destinations accessible by transit or walking within a 30-min peak period on a normal weekday, using GTFS transit schedules and the walkable street network. The proportional overlap of this service area with SA2 shapefiles containing the number of working persons was used to approximate the number of accessible jobs.

3.4. Data treatment

Many BE indicators quantify inter-related aspects of demand (Voulgaris et al., 2017). It is important to eliminate this collinearity before undertaking cluster analysis and linear regression (Hair et al., 2014). The correlation matrix for all facilities and individual modes were examined to identify strong interdependencies. Patterns of inter-correlation were reasonably consistent across modes, with strong correlations ($r > 0.6$) between density variables and some socio-demographic indicators.

High correlations imply redundancy in some variables, and may distort results. Hence three variables were selected to represent theoretical dimensions of density. Commercial density was selected over retail density to represent commercial activity, since it showed stronger direct correlations with ridership. Access to employment was chosen over employment density, despite a lower direct correlation, because it was the only variable in the study that accounted for the relative connectivity of different modes. Population density was selected over dwelling density due to stronger direct correlations and theoretical basis for association with demand. The proportion of full time employed persons (FTE) was chosen over median income, since the latter measured discrete income brackets and therefore provides a less precise measurement than FTE. After removing correlated variables, a total of 19 variables were eligible for analysis.

3.5. Mode location bias mitigation methods

To eliminate mode location bias, sample matching techniques were used to identify train, tram and bus facilities with consistent ('balanced', or unbiased) BE characteristics.The two approaches are:

1. 'Co-located' sample **-** One-to-one matching of at least two different modes based on co-location.
2. 'Stratified' sample- Subclassification of transit catchments with similar BE and sociodemographic (SD) characteristics.

Fig. 2 summarizes the sampling procedure and illustrates the distribution of sites for each mode in the two methods. A more detailed explanation of the matching, sampling and validation steps for each sample are outlined in the following sections.

3.5.1. Co-located sample

Mode pairs were matched based on co-location (2). 'Co-location' was defined as being located within the a 160-m buffer around transit facilities (acceptable intermodal transfer distance) Bus and tram within 160 m of train stations were identified first, since train stations were fewest in number. 147 co-located bus and train facilities were identified, as well as 26 tram and train; and six containing all three modes. The train-tram sites and bus-train-tram sites were not included in the analysis due to small sample sizes. 361 bus facilities intersecting the 160-m transfer zone of trams were then identified. Fig. 3 (bottom left) illustrates the location of the bus-train and bus-tram sites. The bus-train sample is distributed radially, reflecting the configuration of Melbourne's train network. The bus-tram sample is more centralized and grid-like, in the form of Melbourne's light rail network.

Overlapping modes represent a small subset of transit facilities in Melbourne's network. It is not known whether there are synergies or competition between overlapping modes that only impact transit at these locations. To mitigate this risk, a second approach to bias mitigation was tested.

3.5.2. Stratified sample

The second matching approach involved classifying transit facilities with similar catchment-area properties into clusters ('strata') from which equal numbers of each mode were sampled. Bottom-up (non-hierarchical) clustering was used to classify transit catchments according to BE and SD characteristics, using the k-means clustering algorithm. Observations were assigned to one of k (pre-specified) centroids and iterated to minimize the dissimilarity between observations within each cluster (Boehmke, 2019). The R 'cluster' package was used to classify transit catchments (Maechler, 2019). Seventeen continuous SD and BE variables that were not affected by collinearity were included.

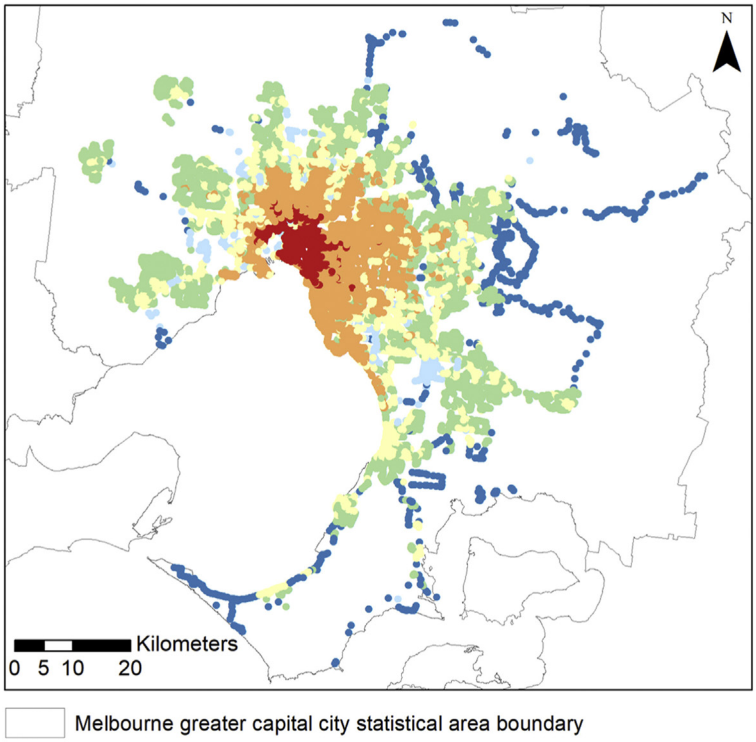
Unlike other statistical methods, cluster analysis does not offer objective measures of fit (Hair et al., 2014). The fit of cluster solutions with different numbers of centroids was compared using the ratio of distance between cluster centroids, 'betweenness sum-of-squares' (BSS), to total dissimilarity of the sample, or 'total sum-of-squares' (TSS). Stuart (2010) notes that 5 to 10 subclassifications are typical of matching solutions. When applied to group spatial units according to land use, previous empirical research has identified three to seven groupings (Jeffrey et al., 2019; Kamruzzaman et al., 2014; Voulgaris et al., 2017). As such, cluster solutions were tested for 2 to 10 clusters. The TSS of the sample was 85,000. A 2-cluster solution explained 30% of the variance between clusters, while a 10-cluster solution explained 52%. As more centroids are added to the solution, more of the variance is explained. However, the ability to interpret and delineate clusters is also an important consideration. Thus, scatterplots were also visualized to check for distinctiveness of clusters. Scatterplots suggested five or six strata solutions were optimal. These groupings were checked for intepretability. The six-strata solution was chosen because it distinguished between Industrial and Mixed-Use Suburban clusters. Fig. 3 illustrates the distribution of strata and their composition by mode.

Traditional stratified sampling involves delineating groups according to important segments that are to be equally represented in the sample (Ortúzar and Willumsen, 2011). However, the objective of stratifying the sample in the present case is to generate a sample of each mode that exhibit similar characteristics. Therefore, the modes were sampled at the same rate, but that rate differed for each strata. The rate of sampling from each cluster depended on the mode with the fewest members in each cluster. In the Urban core, Inner urban and Industrial clusters, train was the constraining mode, with 54, 84 and 1 facility respectively. Tram was the constraining mode in the Mixed use suburban, Residential suburban and Fringe suburban clusters, with 53, three and zero facilities. Since no tram facilities were located in Fringe suburban clusters, no sites from this cluster were included in the sample. The non-constraining modes in each cluster were sampled randomly using the R base function *sample()* (RCore team, 2019).

3.5.3. Validating sample matching

The purpose of matching was to ensure that the observations for each mode had similar BE characteristics. Validation of the Co-located sample was not necessary because the samples comprised matching mode pairs with identical location characteristics by virtue of being co-located. However, validation was necessary to determine whether the Stratified sample exhibited similar properties across modes.

To determine the success of the match, the relative 'balance' for different independent variables was examined. The difference between



Strata description	n (bus)	n (tram)	n (train)	n (total)	n (sample per mode)
Urban core	319	438	54	811	54
Inner urban	2,844	426	84	3,354	84
Mixed use suburban	1,905	53	58	2,016	53
Residential suburban	3,548	3	19	3,579	3
Industrial	236	5	1	242	1
Fringe suburban	632	0	4	636	0
				Total sample size per mode	195

Fig. 3. Size and distribution of transit stop BE and SD classifications ('strata') – stratified sampling method.

average scores for two groups, or mean difference, is useful for interpreting the relative balance of BE variables across modes. Standardized mean difference (SMD) is the absolute difference between mean scores, divided by the standard deviation of the population (all modes) for a variable. A rule of thumb for assessing whether imbalance remains in the sample is to compare the mean differences to the natural constant of variability, which is equal to 0.25 times the original standard deviation (Cochran, 1968; Ho et al., 2007). If SMD between combinations of modes ('bus: train', 'bus: tram', 'tram: train') is less than 0.25, then the samples are considered to be appropriately 'balanced' (Ho et al., 2007). Table 2 compares the SMD for all facilities, and then for the Stratified sample.

Table 2 indicates that the covariate balance was improved in the Stratified Sample for the majority of indicators, with most becoming 'balanced', as represented in bold with values less than 0.25. The clustering was most effective at improving the balance between bus and tram; the most imbalanced subgroups to begin with.Ten indicators that were imbalanced across all tram and bus facilities became balanced in the Stratified sample. After balancing, thirteen tram and bus indicators were balanced, while fourteen were balanced for train and bus (net improvement of six) and fifteen for train and tram (net improvement of eight). Persistent imbalance occurred for employment access across all samples, as well as for land use diversity and housing diversity in the tram: bus and train: bus samples. Observations pertaining to variables with persistent imbalance will be less reliable than for balanced covariates.

Mode pairs are also biased with respect to most SD variables. The Stratified sampling approach also reduced bias for all but two mode pairs with respect to SD variables.

3.6. Analysis approach

The analysis procedure for all samples across the Biased, Co-located and Stratified approaches proceeded according to six steps.

1. First, the variance inflation factor (VIF), was observed to check for collinearity. Variables with VIF score greater than five were omitted.
2. Direct (bivariate) correlations between independent variables and transit use were then examined. The indicators whose bivariate correlation with transit ridership were less than 0.25 were excluded.
3. A maximally adjusted regression model, containing all variables, was specified. The regression methods differed for the two bias mitigation approaches and are detailed below.
4. A parsimonious model was defined according to the optimal combination of variables to explain ridership.
5. After identifying a parsimonious regression model for each sample,

Table 2
Covariate balance statistics (Standardized mean difference) for Biased and Stratified sample.

	All facilities (n = 10,629)				Stratified sample (n = 585)		
	All	Tram: Bus	Train: Bus	Train: Tram	Tram: Bus	Train: Bus	Train: Tram
	σ	$\frac{\lvert d\rvert}{\sigma}$			$\frac{\lvert d\rvert}{\sigma}$		
Variables included in cluster solution							
ln(population density)	1.05	0.82	0.29	0.52	**0.13**	**0.07**	**0.20**
Commercial density	0.18	0.60	0.33	0.27	**0.03**	**0.01**	**0.03**
Jobs-housing balance	0.12	0.57	**0.16**	0.41	**0.21**	**0.03**	**0.24**
Land use diversity	0.17	0.70	0.88	**0.18**	0.38	0.49	**0.11**
Housing diversity	1.70	1.44	1.04	0.40	0.55	0.47	**0.08**
Intersection density	37.3	0.91	**0.23**	0.68	**0.00**	**0.19**	**0.20**
Cycle connectivity	0.16	0.80	0.43	0.37	**0.10**	**0.11**	**0.00**
Destination Score	1.75	1.58	1.61	**0.03**	0.81	1.21	**0.40**
Dist. to CBD	13.2	1.26	0.45	0.80	0.39	**0.09**	0.49
Dist. to Activity Center	2.62	0.49	0.37	**0.11**	**0.03**	**0.08**	**0.05**
Proportion 'urban'	0.23	**0.09**	**0.09**	**0.00**	**0.08**	**0.00**	**0.09**
Access to employment ($\times 10^{-3}$)	98.5	2.02	1.55	0.54	0.74	1.11	1.85
Proportion FTE[1]	0.04	0.69	**0.23**	0.46	**0.14**	**0.00**	**0.14**
Mean size[2]	0.36	1.01	0.61	0.39	**0.10**	**0.15**	**0.06**
Proportion overseas born	0.02	**0.00**	**0.00**	**0.00**	**0.11**	**0.02**	**0.14**
Proportion tertiary educated	0.15	1.36	0.65	0.71	0.50	**0.12**	0.38
Variables not included in cluster solution							
Car parking	0.12	**0.09**	**0.09**	**0.00**	**0.10**	**0.09**	**0.01**
Bicycle facilities	0.24	0.531	**0.18**	0.35	**0.13**	0.29	**0.17**
Count of Activity Centers	996	0.422	**0.10**	0.32	**0.11**	**0.09**	**0.02**
Balanced		3	8	7	13	14	15
Balanced (net)					10	6	8
Improved but imbalanced					6	5	4
Less balanced					0	2	4

Note: Variables for which the SMD suggests naturally occurring variability ($< 0.25*\sigma$) are shown in bold.

diagnostic plots were generated using the *R plot()* function to check adherence of the model to statistical assumptions (R Core Team, 2013). Variables exceeding Cook's distance were considered outliers and removed (Bruce and Bruce, 2017). Several observations that did not fit the linearity, normality and variance plots were also removed.

6. After removing outliers, the maximally adjusted and parsimonious models were iterated until no influential outliers remained.

3.6.1. Linear regression formulation

For the Biased samples and the Stratified sample, multiple linear regression models of BE and ridership were estimated in RStudio (RCore team, 2019). The parsimonious model was estimated by removing predictors with the largest significance level one at a time, until all predictors had an alpha level less than 0.25. from the model in a stepwise using the R base function, *step()*, from the *stats* package, to select the optimal combination of variables based on the Akaike Information Criteria (AIC).

The Co-located sample required an adjusted model. Ridership of Co-located modes was expected to be interdependent due to the increased network accessibility achieved by providing intermodal transfer opportunities. Table 3 summarizes the inputs and results of covariance testing, which confirmed a strong positive correlation between train and bus ridership (r = 0.67), and a weakly positive association between ridership for tram and bus (r = 0.29).

Table 3
Descriptive statistics for Co-located transit.

	Bus-tram	Bus-train
Sample size (clusters)	341	147
Average ln(ridership) (Train)		7.55
Average ln(ridership) (Tram)	5.47	
Average ln(ridership (Bus)	3.18	4.93
Pearson's product-moment correlation	0.29	0.67
95% confidence interval	0.19, 0.39	0.58, 0.75

Covariance was accounted for by simultaneously analyzing the two modes using multivariate multiple linear regression (MMLR). MMLR tests a modified hypothesis that indicators jointly predicted the combined patronage, using multiple analysis of variance (MANOVA) (Ford, 2017).

Step 1 and 2 in the analysis of the Co-located sample proceeded as for the Stratified sample. However, only the variables that did not meet the significance threshold of $p < .25$ for both modes in each sample were excluded from multivariate testing.

At Steps 3 and 4, linear regression models were generated for individual modes. The Anova function from the "car" package in R was used to test the joint significance of each predictor variable considering all other predictors in the model (Type II sum-of-square errors) (Fox and Weisberg, 2019). To produce a parsimonious model (Step 4), variables with the largest F-tests statistics produced from the MANOVA test were removed stepwise until only variables with F-statistics less than 0.1 remained. A relaxed threshold of Pr(F) < 0.1 was used for retaining variables to reflect the subjectivity of stepwise regression.

3.6.2. Treatment of outliers

Over-sampling of sites within Melbourne's free tram fare zone ('Free Tram Zone') was identified to be causing violation of these assumptions in the Co-located sample. The 20 sites in the bus-tram sample located in the Free Tram Zone were excluded from analysis, reducing the sample to 341 sites. Four outliers were removed from the bus-tram sample and six from bus-train. Six outliers were removed from the Stratified train sample, and one from the Stratified tram sample.

4. Results

Descriptive statistics for all facilities in Melbourne are shown in Fig. 1. Descriptives for each sample are available in the online data associated with this paper (3).

Table 4
Parsimonious MMLR model for bus-tram Co-located sample.

					Bus			Tram	
Outliers removed				4					
Sample size				337					
Outcome variable				ln(bus ridership)			ln(tram ridership)		
	MANOVA				Multivariate multiple regression				
	VIF	Pillai test stat.	p	β	SE	p	β	SE	p
ln(population density)	2.33	**0.042**	0.001	0.207	0.204	0.123	**0.145**	0.104	0.016
Commercial density	1.45	**0.027**	0.011	**0.120**	0.480	0.013	**0.107**	0.246	0.024
Land use diversity	1.38	**0.024**	0.021	−0.006	0.505	0.892	**0.124**	0.259	0.008
Housing diversity	1.87	**0.055**	< 0.001	**−0.211**	0.102	< 0.001	**−0.156**	0.053	0.004
Destination Score	1.67	0.018	0.052	0.081	0.056	0.119	**0.112**	0.029	0.028
Access to employment	2.34	**0.041**	0.001	−0.096	< 0.001	0.122	**0.175**	< 0.001	0.004
Level of service (bus)	1.22	**0.393**	< 0.001	**0.637**	0.002	< 0.001	0.068	0.001	0.116
Level of service (tram)	1.10	**0.358**	< 0.001	0.032	0.002	0.454	**0.543**	0.001	< 0.001
Overlapping LOS (train)	1.12	**0.021**	0.031	**0.110**	0.002	0.010	0.004	0.001	0.918
Proportion FTE[1]	1.32	**0.029**	0.009	**−0.113**	1.004	0.015	**−0.110**	0.515	0.016
Proportion overseas born	1.40	**0.019**	0.043	−0.062	3.859	0.191	**−0.112**	1.980	0.016
Proportion tertiary educated	1.83	**0.064**	< 0.001	**0.243**	1.058	< 0.001	−0.017	0.543	0.754
	Intercept			0.000	1.54	0.123	0.000	0.791	< 0.001
	Residual standard error			1.18			0.60		
	df			324			324		
	R^2			0.48			0.50		
	Adjusted R^2			0.46			0.48		

1 - Bus *Proportion Full Time Employed* censored at 95th percentile.

4.1. Regression results for co-located sample

Multiple linear regression was conducted to identify significant associations between transit ridership by mode and BE and socio-demographic catchment characteristics, as well as supply-side variables. To account for covariance of ridership of tram and bus and train and bus in the Co-located samples, an additional MANOVA test was performed to identify jointly significant predictor variables. The parsimonious models resulting from linear regression and MMLR are included in Tables 4 and 5.

The bus-tram model (Table 4) explains approximately half the variance in ridership for both modes, but is a slightly better fit for tram (Adj. $R^2 = 0.48$) compared to bus (Adj. $R^2 = 0.46$).

Results of MMLR for co-located tram and bus (Table 4) suggest six BE characteristics are significant predictors of tram ridership, while only two are associated with bus ridership. The two modes share three common predictors: commercial density, housing diversity and proportion full time employed. Housing diversity, measured as the number of eight housing types present in a catchment, was negatively associated with ridership for both modes. Tram and bus service frequency are the most important predictors of ridership. However, the remaining predictors differ. Bus is related to the frequency of overlapping train services and an increasing proportion of the population that is tertiary educated. Tram ridership is associated with a decreasing proportion of overseas born population. While the association of population density with bus ridership does not meet the significance threshold for rejecting the null hypothesis ($p < .05$), the effect size is modest.Furthermore the model is presumed to have a low power due to the large variance of population density at bus stops, ranging from zero to 22,000 persons/km^2 and the relatively small sample size. This means there is potential to overlook a significant results when relying on significance threshold (Aberson, 2010). Although the effect size falls within the range of small (0.1) and medium (0.3) according to Cohen's power law thresholds (Cohen, 1988), research consistently finds small effect sizes for BE and travel behaviour (Stevens, 2017). This suggests lower bounds may be more appropriate; for characterizing medium and large effects to

Table 5
Parsimonious MMLR model for bus-train Co-located sample.

				Bus			Train		
Outliers removed				6					
Sample size				141					
Outcome variable				ln(bus ridership)			ln(train ridership)		
		MANOVA		Multivariate multiple regression					
	VIF	Pillai test stat.	p	β	SE	p	β	SE	p
ln(population density)	1.75	0.041	0.068	0.148	0.265	0.103	**0.186**	0.101	0.031
Jobs-housing balance	1.52	**0.074**	0.007	**0.207**	1.372	0.002	0.121	0.524	0.058
Pedestrian connectivity	1.48	**0.047**	0.045	**−0.142**	0.004	0.033	**−0.128**	0.002	0.043
Cycle connectivity	1.17	**0.051**	0.034	−0.003	0.519	0.953	**−0.137**	0.198	0.015
Bicycle facilities	1.44	**0.139**	< 0.001	**0.132**	0.213	0.044	**0.028**	0.081	< 0.001
Dist. to Activity Center	2.69	**0.046**	0.048	0.171	0.123	0.056	−0.063	0.047	0.456
Count of Activity Centers	2.38	**0.085**	0.027	**0.029**	0.277	0.001	0.070	0.106	0.379
Access to employment	2.05	**0.078**	0.005	0.007	< 0.001	0.933	**0.227**	< 0.001	0.002
Car parking	1.16	**0.055**	0.027	0.093	< 0.001	0.112	**0.147**	< 0.001	0.009
Level of service (bus)	1.48	**0.396**	< 0.001	**0.567**	0.002	< 0.001	**0.391**	0.001	< 0.001
Level of service (train)	1.30	**0.137**	< 0.001	**0.126**	0.002	0.043	**0.264**	0.001	< 0.001
Intercept			0.000		2.067	0.853	0.000	0.790	< 0.001
Residual standard error			1.05				0.40		
df			129				129		
R^2			0.62				0.66		
Adjusted R^2			0.59				0.63		

capture pratical significance (Khalilzadeh and Tasci, 2017).

In the bus-train sample (Table 5), bus ridership is predicted by four BE variables, and train by six. The availability of bicycle facilities and decreasing pedestrian connectivity are common to both. The decreasing association of ridership with pedestrian connectivity, measured as the density of intersections which represent crossing opportunities for pedestrians, was an unexpected result. This can be explained by the segregation of rail reservations from the road network. Junction density within the rail reservation is therefore non-existent, even when the reservation is highly permeable for pedestrians (Batty, 2013). The strongest predictor of bus ridership was again bus service frequency. Interestingly, it is also the strongest predictor of train ridership, while train frequency was the second strongest predictor. Neither mode shows strong associations with any sociodemographic variables. The explanatory power of the model for train was stronger than that for bus, with adjusted fit indices of 0.63 and 0.59 respectively.

As in the previous model, results for population density (bus), distance to activity center (bus) and jobs-housing balance (train) have modest effects sizes that may have practical significance for planning, but this effect shows variability across the sample.

4.2. Regression results for stratified sample

The results of linear regression for the stratified sample are summarized in Table 6. Continuing the pattern, service frequency of bus and tram are the strongest predictors of ridership for these modes, while train LOS was only the sixth most important predictor of ridership at train stations. No other predictors were shared by all modes. However, bus shared four of its six significant predictors with train. Housing diversity is again negatively associated with bus ridership, as it was in the bus-tram sample, but is not significantly associated with either tram or train ridership. Without probing associations of particular housing types with ridership, it is difficult to discern whether this result is logical or not. For example, it may be that neighborhoods with an absence of single family and semi-detached houses, characterized by low diversity but high density, are driving this correlation.

Bus ridership shares only one significant predictor with tram, population density, although the modest effects size of commercial density suggests it may also be important for bus as well as tram, for which it is clearly significant. Rail way stations car parking is significant for train and bus. It is possible that this result be attributed to the higher ridership among bus stops located at train station interchanges. Access to employment is a strong predictor of both tram and train, while the remaining four predictors of tram ridership are unique. Tram ridership is associated with pedestrian and cycle connectivity, destination score and decreasing proportion of the population that is full time employed. The negative impact of overlapping tram services with tram ridership is an unexpected result, and is at odds with the result for the 'biased' sample of tram which found a significant positive (complementaryt) relationship between overlapping tram and ridership. This suggests the stratified model may not be representative of the entire population of tram stops in the network.

Population density for bus and jobs-housing balance for train were not significant in the co-located model but had modest effect sizes. These two associations were found to be significant in the stratified model. This suggests these two variables have practical significance. Train ridership was associated with seven strong BE preedictors, and had the highest explanatory power of the modes (Adj. $R^2 = 0.68$).

Self-selection into transit-rich neighborhoods was also accounted for using sociodemographic variables, matched in the same way as the BE variables. They were included in the statistical models as control variables; with results suggesting that sociodemographic variables also differ by mode. This finding warrants further investigation to determine whether self-selection is relevant for individual modes.

5. Discussion

Table 7 provides an overview of the key results by bias mitigation method and also compares the results with the analysis undertaken without the use of any bias mitigation. Results with alpha levels meeting a critical threshold of 0.05 are considered to have statistical significance and are highlighted bold. Results which do not meet the

Table 6
Parsimonious linear regression model for Stratified sample.

	Bus				Tram				Train			
Outliers removed	0				3				7			
Sample size	195				192				188			
Outcome variable	ln(bus ridership)				ln(bus ridership)				ln(bus ridership)			
	VIF	**β**	**SE**	**p**	**VIF**	**β**	**SE**	**p**	**VIF**	**β**	**SE**	**p**
ln(population density)	2.62	**0.283**	0.190	< 0.001	1.87	**0.170**	**0.077**	0.010	2.42	0.111	0.093	0.088
Commercial density	1.56	0.122	0.498	0.056	1.52	**0.157**	0.249	0.009				
Jobs-housing balance	2.06	**0.192**	1.011	0.009					1.60	**0.105**	0.471	0.046
Housing diversity	1.90	**−0.245**	0.078	0.001								
Pedestrian connectivity					2.55	0.116	0.001	0.131				
Cycle connectivity					1.25	**0.145**	0.211	0.007				
Bicycle facilities (dummy)									1.51	**0.233**	< 0.001	< 0.001
Distance to CBD					2.96	**0.286**	0.015	0.001				
Count of Activity Centers	1.32	0.069	0.417	0.237					1.27	**0.198**	0.075	< 0.001
Destination score					1.63	**0.136**	0.029	0.028				
Proportion 'urban'	1.46	0.099	0.562	0.109								
Access to employment					3.75	**0.258**	< 0.001	0.006	2.64	**0.324**	< 0.001	< 0.001
Car parking	1.24	**0.121**	< 0.001	0.034					1.18	**0.154**	< 0.001	< 0.001
Level of service (LOS)	1.23	**0.536**	0.004	< 0.001	1.27	**0.536**	0.002	< 0.001	1.26	**0.277**	0.001	< 0.001
Overlapping LOS (bus)									1.68	**0.289**	0.001	< 0.001
Overlapping LOS (tram)					1.14	**−0.152**	0.001	0.003	1.64	**0.276**	0.001	< 0.001
Proportion FTE[1]	1.12	−0.083	1.995	0.125	1.30	**−0.266**	0.529	< 0.001		0.078	0.001	0.144
Proportion overseas born												
Proportion tertiary educated	1.73	**0.229**	0.786	0.001					2.01	−0.073	0.339	0.218
Intercept (p)	0.000	1.23	1.23	0.003	0.000		0.667	< 0.001	0.000		0.704	< 0.001
Residual standard error		1.12				0.50				0.44		
df		182				181				176		
R^2		0.53				0.58				0.70		
Adjusted R^2		0.50				0.56				0.68		

1 - Bus *Proportion Full Time Employed* censored at 95th percentile.

Table 7
Comparison of built environment predictors of mode ridership for different methods.

	All Combined	Bus Biased	bus-tram	bus-train	Stratified	Tram Biased	bus-tram	Stratified	Train Biased	bus-train	Stratified
Sample size (facilities)	10,629	9483	337	141	195	925	337	192	206	141	188
Adj. R^2	0.60	0.42	0.46	0.59	0.50	0.66	0.48	0.56	0.72	0.63	0.68
	β	β				β			β		
ln(population density)	**0.020**	**0.158**	0.207	0.148	**0.283**		**0.145**	**0.170**	**0.177**	**0.186**	0.108
Commercial density	0.088	**0.069**	**0.120**		0.121	**0.242**	**0.107**	**0.157**			
Jobs-housing balance	**0.122**	**0.070**		**0.207**	**0.192**	**−0.079**				0.121	**0.105**
Land use diversity	**0.090**	**0.042**				**0.063**	**0.124**				
Housing diversity	**0.011**		**−0.211**		**−0.245**		**−0.156**				
Pedestrian connectivity		**0.049**		**−0.142**				0.116		**−0.128**	
Cycle connectivity						**0.064**		**0.145**		**−0.137**	
Bicycle facilities	**0.067**			**0.132**					**0.234**	**0.028**	**0.233**
Distance to CBD								**0.286**			
Distance to Activity Center				0.171							
Count of Activity Centers	**0.043**	**0.017**		**0.029**					**0.136**		**0.198**
Destination Score	**0.151**	**0.117**				**0.068**	**0.112**	**0.136**			
Free Tram Zone									**0.444**		**0.297**
Proportion 'urban'	**0.037**	**0.027**				**0.067**					
Access to employment	**0.097**	**−0.052**					**0.175**	**0.258**	**0.306**	**0.227**	**0.324**
Car Parking	**0.032**				**0.121**				**0.148**	**0.147**	**0.154**
Level of service (LOS)	**0.464**	**0.476**	**0.637**	**0.567**	**0.536**	**0.563**	**0.543**	**0.536**	**0.288**	**0.264**	**0.277**
Overlapping LOS (bus)	**0.037**	**−0.025**				**0.041**			**0.291**	**0.391**	**0.289**
Overlapping LOS (tram)	0.011					**0.101**		**−0.152**			
Overlapping LOS (train)		**0.066**	**0.110**	**0.126**							
Proportion FTE[1]	0.009		**−0.113**				**−0.110**	**−0.266**			
Mean size[1]	**0.056**	**0.057**				**0.059**					
Proportion overseas born	**0.031**	**0.039**					**−0.112**				
Proportion tertiary educated	**0.067**	**0.080**	**0.243**		**0.229**						

1. - Bus *Proportion Full Time Employed* and *Mean Size* censored at 95th percentile.

significance threshold but which have a standardized effects size exceeding 0.1 are included in the table, to show effects that may have practical significance.

5.1. Comparison of results for combined and separate modes

In the Combined model, all predictors have limited practical significance (effect sizes less than 0.1). When disaggregated by mode ('Biased' samples) the explanatory power of the model for tram and train are improved. This contrasts with the findings from prior studies that find the predictive strength of models for transit modes combined is stronger than for individual modes (Liu et al., 2016; Tu et al., 2018). The relative number and distribution of modes in a network, as well as the number of transfer opportunities between modes, may determine whether variability is better explained by combined or separate models.

In the case of Melbourne, where bus transit dominates the physical supply of transit, the predcictors of train and tram in the combined model are diluted. When disaggregated by mode, the nuance of modes is clearer. Effect sizes are larger, because there is less variability among modes of the same type. Significant effects and effect directions differ in the mode specific models. However, this comparison is fraught due to the possibility that differences in the model for 'all modes'; compared to individual modes is attributable purely to the different locations in which each mode is found. The following sections present findings for samples that do not suffer from location bias.

5.2. Comparison of results for biased and unbiased samples

In almost all cases, the direction of variable affects (+/−) are consistent between biased and unbiased results; suggesting some degree of theoretical consistency in all models tested. However, the order of the effect size is quite different and the significant variables selected are not exactly the same.

- In almost all cases, Level of Service (LOS) is the variable with the biggest effect size, whether the sample is biased or not, or what bias mitigation method is used.However, there is some variation in which LOS indicator is important in each case.
- For tram and train, the biased data have slightly better statistical fit than the unbiased data.
- The Stratified model produced an unexpected result for the association of overlapping tram frequency with tram ridership. In the Biased sample the results is positive, while in the Stratified model the result is negative. This suggests the Stratified model is not representative of tram ridership across Melbourne.
- For bus, all bias mitigation methods have a stronger fit (R^2) than the biased data. This suggests bus locations are highly variable across Melbourne.

5.3. Comparison of results for modes

There is a consistent picture that different variables, and different BE variables in particular, affect transit modes in different ways. The following BE attributes affect modes across the majority of samples:

- Train, tram, bus: Population density
- Train, bus: Jobs-housing balance, Activity centers
- Tram, bus: Commercial density
- Tram, train: Access to employment
- Train only: Bicycle facilities, Car parking, Located in Free Tram Zone
- Tram only: Land use diversity, Cycle connectivity, Destination score
- Bus only: Housing diversity (−)

Results across the biased and unbiased approaches all suggest the make-up and rank of BE predictors of ridership differ by mode. Testing unbiased samples makes this conclusion generalizable: the BE affects ridership of modes differently, irrespective of the type of location each mode services.

5.4. Preferred bias mitigation method

The make-up of significant predictors and their relative predictive strength is slightly different across all samples. Each sample reveals a set of predictors that is important under the chosen sampling conditions:

- LOS indicators are significant variables with the largest affect sizes by all methods, with some only minor variations in LOS indicator selected between mitigation method results.
- Apart from LOS being the 1st ranked variable in each case,other variable results are quite different between bus-tram and bus-train. This may illustrate the impacts of different sample characteristics between bus-tram, which is inner metropolitan in nature and bus-train, which is middle and outer metropolitan in nature.In this respect, the bias mitigation method includes some spatial bias.
- The two bias mitigation methods identify quite similar significant variables but the ranking of these is quite different between method (excluding LOS). Significant variables are not entirely the same between mitigation approaches, however where different significant variables are selected by a method, they tend to be lower ranked.
- Stratified sampling has slightly stronger explanatory power (R^2) than Co-located results for train and tram. For bus results are more mixed.

It is important to choose a sampling method that is most representative of the network. Not all networks have co-located transit. Catchment overlap effects are likely to impact results and have limited generalizability to non-overlapping locations. In contrast, the Stratified method is flexible. Clusters can be defined to represent the distinctive Strata of a transit network catchment; with the sample comprising transit facilities from each. Findings support this: the Stratified model shared more relationships in common with the Biased sample (which constitutes a census of the network). For example, the bus-train model shared only four of its seven predictors with the Biased sample, whereas the Stratified sample shared six.

The Stratified model also prevails as the superior approach according to explanatory power (R^2). The Stratified model explains more variability in all samples except one. The only case for which this is not the case is the bus-train model (for bus only). The strong covariance of bus with train (Train 3) suggests bus ridership is reated to its function as a feeder mode to train at these sites.

On balance, we suggest that the Stratified approach is superior to the Co-located approach because it identifies underlying patterns that are representative of an entire network and samples according to these. To account for the strong interdependence of ridership at these locations, future studies should distinguish bus-rail interchanges through experiment design or dummy variables.

To determine the necessity of removing location bias in future studies, it would be useful to first examine the standardized mean difference of variables at different mode catchments. Samples with few 'imbalanced' mode pairs do not need to be treated for location bias. Examining smaller, homogeneous study areas, such as a central city, or individual strata such as the BE and SD strata identified in this study, would not need to be treated for bias.

6. Conclusions

This is the first study that compared BE associations with demand for different transit modes while accounting for systematic bias in the location characteristics of modes. Testing unbiased samples makes the main finding of this study generalizable: the BE affects ridership of modes differently, irrespective of the type of location each mode services.

Linear regression of transit ridership for train, tram and bus in Melbourne produce different results when modes are modelled individually, rather than in combination. However, the location bias in the supply of different modes prevents interpretation of this result as evidence that the BE predictors of modes in the same location are different.

To account for important differences in the locations associated with each mode, this study developed two bias mitigation approaches. The first samples Co-located transit so that mode pairs are matched one-to-one based on co-location. The second uses Stratified sampling to draw observations from subcategories with similar BE characteristics.

Results across the biased and unbiased approaches all suggest the make-up and rank of BE predictors of ridership differ by mode. Only population density and level of service are predictors of all three modes. Results suggest train ridership is consistently related to intermodal transfer facilities. Both bus and train are related to jobs housing balance and the presence of Activity Centers in the catchment. Both tram and train are strongly related to access to employment, however this is not a factor affecting bus ridership. Bus ridership is inversely related to housing diversity, and positively related to the frequency of overlapping rail services, and to the proportion of the population that is tertiary educated. Both tram and bus are related to commercial density, while tram alone shows strong relationships to measures of diversity and access including land use diversity and destination score.

When the objective of statistical analysis is to draw robust comparisons and identify predictors with influence irrespective of location type, it is important to compare imbalanced samples. The necessity of mitigating location bias in future comparison should be assessed by examining the standardized mean difference of variables between modes. Imbalanced samples should be treated using the Stratified sampling approach, which involves defining bespoke clusters that are representative of a network as a whole.

When the objective of modelling is to make accurate predictions, it is important to use a sample that is the most representative of the network. Based on the findings of this study, the most representative sample is one that uses all available observations for individual transit modes. The findings from this study confirm that models for individual modes provide a better fit for train and tram than when these modes are combined with bus. Combined forecasting models may either undervalue or overlook the varied associations of different modes with transit use. Separate models for individual modes reveal certain predictors to have much larger associations with ridership than suggested by a combined model. These results suggest that ridership models could provide more accurate forecasts if predictions are differentiated by mode.

6.1. Limitations and further research

This study was designed to maximize the generalizability of comparisons between modes, rather than to identify specific predictors of ridership. Its observational study design means that the factors associated with ridership of each mode are sensitive to the research setting and study design. (Holz-Rau and Scheiner, 2019; Scheiner, 2018; van de Coevering et al., 2015). Despite taking measures to balance the covariates to make the modal comparison as robust as possible, validation against samples in other networks are needed to test whether ridership predictors differ by mode in different networks. A limited body of prior research that compare predictors for individual modes, including studies conducted in Maryland, USA, and Shenzhen, China, find that combined models for transit modes explain more variability in the ridership of transit, which contrasts with this study's findings (Liu et al., 2016; Tu et al., 2018). Comparison with other networks would enable testing whether location bias is a prevalent problem in other cities; and if transit mode predictors differ only in situations where the supply locations of transit modes differ significantly.

Results also suggest that sociodemographic variables differ by mode. This finding warrants further investigation to determine whether self-selection is relevant for individual modes.

Further research should seek to understand what is driving differences in the predictors for specific modes from the perspective of perceptions, attitudes and daily mobility-travel patterns that might be mediated by the BE. This will enable policies that leverage specific BE factors associated with demand for different modes.

7. Notes

1 - Results pertaining to the 'Co-located' sample were presented as a poster at the 2020 Transportation Research Board Annual Meetingtion Aston et al., 2020.

2 - The data and data aggregation procedures pertaining to this paper are available on *Bridges* at https://doi.org/10.26180/5d9994f4704ea.

3 - Supplementary results and tables are available on GitHub at https://github.com/Laura-k-a/BE-TU_Melb_Clusters and indexed in supplementary Table 8 associated with this paper.

Acknowledgments

We would like to acknowledge funding provided by the Victorian Department of Transport and the Commonwealth Government's Research and Training Program to support the research activities of Laura Aston, the lead author of this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jtrangeo.2020.102786.

References

Aberson, C.L., 2010. Applied Power Analysis for the Behavioral Sciences. Routledge, New York.

Aston G. Currie, L., 2020. Exploring Variation in Built Environment Predictors of Ridership by Transit Mode. In Transportation Research Board Annual Meeting Washington, D.C.

Australian Bureau of Statistics, 2017a. Victoria (STE) (Statistical Area Level 1) 2016 General Community Profile: G09H Table Country of Birth of Person by Age by Sex, 2016 Census of Population and Housing. https://datapacks.censusdata.abs.gov.au/datapacks/.

Australian Bureau of Statistics, 2017b. Victoria (STE) (Statistical Area Level 1) 2016 General Community Profile: Table G01: Selected Person Characteristics by Sex, 2016 Census of Population and Housing. https://datapacks.censusdata.abs.gov.au/datapacks/.

Australian Bureau of Statistics, 2017c. Victoria (STE) (Statistical Area Level 1) 2016 General Community Profile: Table G32: Dwelling Structure, 2016 Census of Population and Housing. https://datapacks.censusdata.abs.gov.au/datapacks/.

Australian Bureau of Statistics, 2017d. Victoria (STE) (Statistical Area Level 1) 2016 General Community Profile: Table G38: Dwelling Structure by Household Composition and Family Composition, 2016 Census of Population and Housing. https://datapacks.censusdata.abs.gov.au/datapacks/.

Australian Bureau of Statistics, 2017e. Victoria (STE) (Statistical Area Level 1) 2016 General Community Profile: Table G43B: Dwelling Structure by Household Composition and Family Composition, 2016 Census of Population and Housing. https://datapacks.censusdata.abs.gov.au/datapacks/.

Australian Bureau of Statistics, 2017f. Victoria (STE) (Statistical Area Level 1) 2016 General Community Profile: Table G46B Non-School Qualification: Level of Education by Age by Sex, 2016 Census of Population and Housing. https://datapacks.censusdata.abs.gov.au/datapacks/.

Australian Bureau of Statistics, 2017g. Victoria (STE) (Statistical Area Level 2) 2016 Working Person Profile: Table W01 Labour Force Status by Age by Sex, 2016 Census of Population and Housing. https://datapacks.censusdata.abs.gov.au/datapacks/.

Australian Bureau of Statistics, 2017h. Victoria (STE) (Statistical Area Level 2) 2016 Working Person Profile: Table W09 Industry of Employment by Sex, 2016 Census of Population and Housing. https://datapacks.censusdata.abs.gov.au/datapacks/.

Badland, H., Mavoa, S., Boulange, C., Eagleson, S., Gunn, L., Stewart, J., David, S., Giles-Corti, B., 2017. Identifying, creating, and testing urban planning measures for transport walking: findings from the Australian national liveability study. J. Transp. Health 5, 151–162.

Batty, M., 2013. The New Science of Cities. MIT Press, Cambridge, Massachusetts.

Bicycle Network Victoria, 2019. Parkiteer Locations. viewed 5/10/19. https://www.bicyclenetwork.com.au/our-services/parkiteer/locations/.

Boehmke, B., 2019. K-means Cluster Analysis. viewed 8/10/19. http://uc-r.github.io/kmeans_clustering.

Boulange, C., Gunn, L., Giles-Corti, B., Mavoa, S., Pettit, C., Badland, H., 2017. Examining associations between urban design attributes and transport mode choice for walking, cycling, public transport and private motor vehicles. J. Transp. Health 6, 155–166.

Bruce, P., Bruce, A., 2017. Practical Statistics for Data Scientists. O'Reilly Media.

Cao, X., Fan, Y., 2012. Exploring the influences of density on travel behavior using propensity score matching. Environ. Plan. Plan. Design 39 (3), 459–470.

Cao, X., Mokhtarian, P.L., Handy, S.L., 2009. Examining the impacts of residential self-selection on travel behaviour: a focus on empirical findings. Transp. Rev. 29 (3), 359–395.

Cervero, R., 2002. Built environment and mode choice: toward a normative framework. Transp. Res. D 7 (2002), 265–284.

Cervero, R., Ferrell, C., Murphy, S., 2002. Transit-Oriented Development and Joint Development in the United States: A Literature Review. Institute of Urban and Regional Development, University of California, Berkeley, University of California, Berkeley.

Cervero, R, Murphy, S, Ferrell, C, Goguts, N, Tsai, YH, Arrington, GB, Boroski, J, Smith-Heimer, J, Golem, R, Peninger, P, Nakajima, E, Chui, E, Dunphy, R, Myers, M, McKay, S & Transportation Research, B 2004, Transit-Oriented Development in the United States: Experiences, Challenges, Prospects, Transportation Research Board, 0309087953.

Cochran, W.G., 1968. The effectiveness of adjustment by subclassification in removing bias in observational studies. Biometrics 24, 295–313.

Cohen, J., 1988. Statistical Power Analysis for the Behavior Sciences, 2 edn. Erlbaum, Hillsdale, New Jersey.

De Gruyter, C., 2017. Travel Plans for New Residential Developments : Insights from Theory and practice', Thesis.

De Vos, J., Ettema, D., 2020. Travel and residential change: an introduction. Travel Behav. Soc. 19, 33–35.

Deng, Y., Yan, Y., 2019. Propensity score weighting with generalized boosted models to explore the effects of the built environment and residential self-selection on travel behavior. Transpor. Res. Rec. 2673 (4), 373–383. https://doi.org/10.1177/0361198119837153.

Department of Environment Land Water & Planning, 2018. Planning scheme zones - Vicmap Planning. http://services.land.vic.gov.au/SpatialDatamart/viewMetadata.html?anzlicId=ANZVI0803002909.

Department of Environment, Land, Water and Planning, 2017. Activity Centres Overview. Plan Melbourne. State of Victoria.

Department of Infrastructure, 2002. Implementation Plan 4: Activity Centres (Draft). Melbourne 2030. State of Victoria.

Department of Transport, 2013. Public Transport: Guidelines for Land Use and Development, State of Victoria, Victoria.

Department of Transport, 2019a. Data Correction: Metropolitan Bus Patronage - Updated Stop Level (2018).

Department of Transport, 2019b. Data Request Metropolitan Patronage - Stop Level (2018).

Do Infrastructure, 2017a. Vicmap Transport. data.vic.gov.au. https://www.data.vic.gov.au/data/dataset/road-network-vicmap-transport (viewed 6/11/2-17).

ESRI, 2019. ArcMap, 1 edn. New York Street Redlands, CA, USA.

Estupiñán, N., Rodríguez, D.A., 2008. The relationship between urban form and station boardings for Bogotá's BRT. Transp. Res. A 42 (2), 296–306.

Ewing, R., Cervero, R., 2010. Travel and the built environment: a meta-analysis. J. Am. Plan. Assoc. 76 (3), 265–294.

Ford, C., 2017. Getting Started with Multivariate Multiple Regression. University of Virginia Library, Virginia.

Fox, J., Weisberg, S., 2019. An {R} Companion to Applied Regression. Sage, Thousand Oaks CA.

GeoFabrik downloads, 2019. GIS OSM Pois Free. Spatial, Australia, pp. 1.

Guerra, E., Cervero, R., Tischler, D., 2012. Half-mile circle does it best represent Transit Station catchments? Transp. Res. Rec. 2276, 101–109.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 2014. Pearson new international edition. Edn. In: *Multivariate data analysis*, Seventh Edition. Pearson Education Limited, Harlow, Essex.

Heinrich, C., Maffioli, A., Vazquez, G., 2010. A Primer for Applying Propensity-Score Matching No. IDB-TN-161. Inter-American Development Bank.

Higgs, C., Badland, H., Simons, K., Knibbs, L.D., Giles-Corti, B., 2019. The urban liveability index: developing a policy-relevant urban liveability composite measure and evaluating associations with transport mode choice. Int. J. Health Geogr. 18 (1), 14.

Ho, D.E., Imai, K., King, G., Stuart, E.A., 2007. Matching as nonparametric Preprocessing for reducing model dependence in parametric causal inference. Polit. Anal. 15 (3), 199–236.

Holz-Rau, C., Scheiner, J., 2019. Land-use and transport planning – a field of complex cause-impact relationships. Thoughts on transport growth, greenhouse gas emissions and the built environment. Transp. Policy 74, 127–137.

Jeffrey, D., Boulangé, C., Giles-Corti, B., Washington, S., Gunn, L., 2019. Using walkability measures to identify train stations with the potential to become transit oriented developments located in walkable neighbourhoods. J. Transp. Geogr. 76, 221–231.

Kager, R., Bertolini, L., Te Brömmelstroet, M., 2016. Characterisation of and reflections on the synergy of bicycles and public transport. Transp. Res. A 85, 208–219.

Kamruzzaman, M., Baker, D., Washington, S., Turrell, G., 2014. Advance transit oriented development typology: case study in Brisbane, Australia. J. Transp. Geogr. 34, 54–70.

Khalilzadeh, J., Tasci, A.D.A., 2017. Large sample size, significance level, and the effect size: solutions to perils of using big data for academic research. Tour. Manag. 62, 89–96.

Kittelson & Associates Inc, Parsons Brinckerhoff Inc, KFH Group Inc, Texas a&M Transportation Institute & Arup, 2013. 2. Mode and Service Concepts. 165 Transit Cooperative Highway Research Program (TCRP).

Litman, T., Steele, R., 2017. Land Use Impacts on Transport: How Land Use Factors Affect Travel Behavior. Victoria Transport Policy Institute (viewed 24/6/2017).
Liu, C., Erdogan, S., Ma, T., Ducca, F.W., 2016. How to increase rail ridership in Maryland: direct ridership models for policy guidance. J. Urban Plann. Dev. 142 (4).
Maechler, M., 2019. Finding Groups in Data: Cluster Analysis. Extended Rousseeuw et al., 2.1.0 Edn.
Mavoa, S., Koohsari, M.J., Badland, H.M., Davern, M., Feng, X., Astell-Burt, T., Giles-Corti, B., 2015. Area-level disparities of public open space: a geographic information systems analysis in metropolitan Melbourne. Urban Policy Res. 33 (3), 1–18.
Mitchell, R.B., Rapkin, C., 1954. Urban Traffic : A Function of Land Use. Columbia University Press.
Monzón, A., Hernández, S., Di Ciommo, F., 2016. Efficient urban interchanges: the City-HUB model. Transportation Research Procedia 14, 1124–1133 no. C.
Nasri, A., Zhang, L., 2019. Multi-level urban form and commuting mode share in rail station areas across the United States; a seemingly unrelated regression approach. Transp. Policy 81, 311–319.
Ortúzar, J.D., Willumsen, L.G., 2011. Modelling Transport, Fourth edn. John Wiley & Sons Ltd, Chichester, West Sussex, United Kingdom.
PSMA Australia limited, 2018. PSMA Australia Limited, PSMA Features of Interest (Polygon) (August 2018). accessed from AURIN on 1/3/2019, Spatial.
Public Transport Victoria, 2018a. 1 March 2018, PTV GTFS. https://transitfeeds.com/p/ptv/497/20180319.
Public Transport Victoria, 2018b. Public Transport Points Public Transport: A collection of PTV datasets. http://services.land.vic.gov.au/catalogue/metadata?anzlicId=ANZVI0803005298&publicId=guest&extractionProviderId=1.
Public Transport Victoria, 2019. PTV Train Station Car Park. https://discover.data.vic.gov.au/dataset/ptv-train-station-car-park.
R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
RCore team 2019, The R Base Package, 3.6.2 Edn, R-Core R-core@R-project.org.
Rijsman, L., van Oort, N., Ton, D., Hoogendoorn, S., Molin, E., Teijl, T., 2019. Walking and Bicycle Catchment Areas of Tram Stops: Factors and Insights. TU Delft, Delft.
Scheiner, J., 2018. Why is there change in travel behaviour? In: Search of a Theoretical Framework for Mobility Biographies Erkunde Archive for Scientific geography. 72. pp. 41–62 1.
Shannon, C.E., 1948. A mathematical theory of communication. Bell Syst. Tech. J. 27, 379–423 623–56.
Stevens, M., 2017. Does compact development make people drive less? J. Am. Plan. Assoc. 83 (1), 7–18.
Stuart, E.A., 2010. Matching methods for causal inference: a review and a look forward. Stat. Sci. 25 (1), 1–21.
Taylor, B.D., Miller, D., Iseki, H., Fink, C., 2009. Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. Transp. Res. A Policy Pract. 43 (1), 60–77.
Tu, W., Cao, R., Yue, Y., Zhou, B., Li, Q., Li, Q., 2018. Spatial variations in urban public ridership derived from GPS trajectories and smart card data. J. Transp. Geogr. 69, 45–57.
van de Coevering, P., Maat, K., Van Wee, B., 2015. Multi-period research designs for identifying causal effects of built environment characteristics on travel behaviour. Transp. Rev. 35 (4), 1–21.
VicRoads, 2017. Principal Bicycle Network, Digital Shapefile.
Voulgaris, C.T., Taylor, B.D., Blumenberg, E., Brown, A., Ralph, K., 2017. Synergistic neighborhood relationships with travel behavior: an analysis of travel in 30,000 US neighborhoods. J. Transp. Land Use 10 (1), 437–461.