

Empirical built environment and transit use literature since 2000: Comprehensive literature review protocol and results

Version 2: 5-Aug-2019

Investigators

Laura Aston ^{1*}	laura.aston@monash.edu	https://orcid.org/0000-0002-3495-4094
Professor Graham Currie FTSE ¹	graham.currie@monash.edu	https://orcid.org/0000-0001-7190-7548
Dr Alexa Delbosc ¹	alexa.delbosc@monash.edu	https://orcid.org/0000-0001-8744-3469
Tyler O'Hare ²	tylerlohare@gmail.com	
A/P MD Kamruzzaman ³	md.kamruzzaman@monash.edu	https://orcid.org/0000-0001-7113-942X
David Teller ⁴	david.teller@ecodev.vic.gov.au	

1 – Public Transport Research Group, Institute of Transport Studies, Monash University

2 – Monash Art Design and Architecture, Monash University

3 – Department of Civil Engineering, Monash University

4 – Department of Transport, State of Victoria

Summary: This protocol sets out the strategy used to identify records, screen for relevance and extract data points pertaining to empirical estimates of the relationship between the built environment and transit use. The present document is intended as a methodological reference for subsequent analyses that draw on the data obtained through the search. The purpose of conducting this comprehensive review was to obtain a representative sample of studies since 2000 from which robust meta-averages could be calculated. The secondary aim was to detect patterns of variation in effect sizes related to study invariant characteristics such as attributes of the sample, location of data collection and study design. Quantities of records identified or screened at each stage of the review are also documented to give a thorough and transparent record of results.

Version	Date	Log of changes
1	8-Jan-19	Original publication
2	14-Aug-19	Changes to coding framework 2.1 Introduction of dichotomous variable to identify control for 'regional accessibility' and 'self selection' 2.2 Removed criteria for screening by country → introduced 39 new studies 2.3 Realigned built environment indicators to include accessibility indicators, with less emphasis on 'D-variables'

Contents

Table index.....	3
Checklist of preferred reporting items for reviews (The Prisma-P Group 2015).....	4
1. Overview	6
1.1 Review Question	6
1.2 Search details (search dates, restrictions)	6
1.3 Conditions or domain being studies	6
1.4 Comparator(s)/controls	6
1.5 Types of study to be included	6
1.6 Primary outcomes.....	6
1.7 Secondary outcomes.....	6
1.8 Strategy for data synthesis	7
1.9 Analysis of subgroups or subsets	7
1.10 Risk of bias	7
2. Identification strategy.....	7
2.1 Concepts and terms of the search strategy	7
2.2 Identify and select the electronic databases	8
2.3 Pilot the search strategy	9
2.4 Define the search phrase	10
2.5 Run the search and record results	10
2.6 Collate results	11
3. Screening.....	11
3.1 Automatic filtering of results	11
3.2 Title/abstract screen	11
3.3 Eligibility based on full text screen	13
3.4 Validation	13
3.5 Screening summary.....	14
4. Extraction.....	14
4.1 Data point identification	14
4.2 Data for extraction.....	14
4.3 Attribute coding definitions	16
4.4 Data management	22
References	23

Table index

Table 1 - Key words and alternatives used in database search	8
Table 2 - Databases used in prior studies or considered based on descriptions.....	8
Table 3 - Keyword and alternative search results (Pilot conducted 7 December 2017)	9
Table 4 - Search strings and records retrieved by database (date: 7/12/17).....	10
Table 5 - Results of duplicate removal and filtering by source type and publication year.....	11
Table 6 – Criteria applied to screen records based on information in title and abstract.....	12
Table 7 - Results of title and abstract screening.....	12
Table 8 - Study and sample characteristics for determining eligibility of studies	13
Table 9 - Validation of comprehensive search based on precedent studies	13
Table 10 - Summary of characteristics extracted from priors for each data point (refer to metadata in database for full description and attribute levels coded)	15
Table 11 - Definition of significance thresholds.....	16
Table 12- Standardized transit travel behavior variables	16
Table 13 - Travel mode definitions	17
Table 14 - Population subset definitions	17
Table 15 - Trip purpose definitions	18
Table 16 – Covariance testing definitions.....	18
Table 17 - External variables and indicators in built environment and travel behavior research (Source: author’s synthesis).....	18
Table 18 – Methodological strategies to account for attitude-induced self-selection.....	19
Table 19 - Standardized built environment indicators	20
Table 20 - Assumptions pertaining to interpretation and extraction of built environment indicators	21
Table 21 - Location of built environment measurement definitions.....	21
Table 24 - Unit of analysis of independent variables - definitions and assumptions	21
Table 25 – Relationship estimation methods and associated additional information	22

Checklist of preferred reporting items for reviews (The Prisma-P Group 2015)

Section	Topic	Item	Elaboration or reference	
Administrative information	Title: Identification	1a	Built Environment and Transit Use: A Review of the Empirical Literature since 2000	
	Title update	1b	n/a	
	Registration	2	n/a (not accepted to PROPSERO)	
	Authors: Contact	3a	Refer to front matter	
	Authors: Contributions (guarantor n/a)	3b	Name	Contribution
			Laura Aston	Primary author
			Professor Graham Currie FTSE	Conception and interpretation of analysis to be performed with data
			Dr Alexa Delbosc	Conception and interpretation of analysis to be performed with data
			Tyler O'Hare	Screening, extraction and coding of data
			A/P MD Kamruzzaman	Review and consolidation of data contained in database, interpretation
			David Teller	Conception and interpretation of analysis to be performed with data
	Amendments	4	n/a	
	Support: Sources	5a	Monash University	
	Support: Sponsor	5b	Department of Transport (formerly Transport for Victoria)	
	Support: Role of sponsor or funder	5c	Co-authorship (refer to author contributions)	
Introduction	Rationale	6	Section 1.3: Conditions or domain being studies	
	Objectives	7	Section 1.6: Primary outcomes Section 1.7: Secondary outcomes	
Methods	Eligibility criteria	8	Section 3.1: Automatic filtering of results Section 3.2: Title/abstract screen	
	Information sources	9	Table 5 - Results of duplicate removal and filtering by source type and publication year	
	Search strategy	10	Section 2 (all): Identification strategy	
	Study records: data management	11a	Section 2.6: Collate results Section 5: Data management	
	Study records: selection process	11b	Section 3.3: Eligibility based on full text screen	
	Study records data collection	11c	Section 5: Extraction	
	Data items	12	Section 5.2: Data for extraction Section 5.3: Attribute coding definitions	
	Outcomes and prioritization	13		
	Risk of bias in individual studies	14	Section 1.10: Risk of bias	

Section	Topic	Item	Elaboration or reference
	Data synthesis	15	Section 1.8 (High-level): Strategy for data synthesis
	Meta-bias(es)	16	Section 1.10: Risk of bias
	Confidence in cumulative evidence	17	Not included in this protocol

1. Overview

1.1 Review Question

Which features of the built environment influence transit ridership, by how much, and under what conditions?

1.2 Search details (search dates, restrictions)

Test databases used (7 December 2017): Scopus, Compendex, Transport. Full search databases used (7 December 2017): Scopus, Compendex, Transport, INSPEC, ATRI, TRID, Proquest, Web of Science (Core Collection). #

Detailed in Section 1: Identification strategy

1.3 Conditions or domain being studies

Researchers have extensively studied the relationship between the built environment and travel behavior, with a view to understand how the built environment can be leveraged to achieve a higher share of transportation by sustainable and active (non-private motorized) modes. This review focuses on the subset of travel and land use studies that analyse transit ridership or mode share as the dependent variable; and seeks to further segment studies by characteristics of the sample and results to identify if certain trends exist between the sample characteristics and the quantitative correlations between the built environment and transit ridership.

1.4 Comparator(s)/controls

Parameter estimates expressing a statistical relationship between an indicators of the built environment (independent variable) and transit use (dependent variable) is the focus of subsequent reviews. Comparison will be made of the magnitude of these estimates for different variables, as well as indicators of the variables.

Further segmentation will explore how these parameter estimates vary with study invariant characteristics.

1.5 Types of study to be included

Excluded: qualitative, grounded theory studies

Included: quantitative observational and quasi experimental studies with longitudinal, cross-sectional or case study sampling, simulations on a case-by-case basis

Detailed in Section 2: Screening

1.6 Primary outcomes

- (1) Updated meta-elasticities for relationships of different indicators of the built environment and transit use.
- (2) Significant patterns of association between study invariant characteristics and effect sizes.

1.7 Secondary outcomes

- (3) Qualitatively describe the body of existing literature by segmenting studies based on study type, sample characteristics, indicators measured, statistical or other model used, and conditions for causality met in the design of the study.

- (4) Comparison of effect sizes, and in particular identifying the range of variances for a given independent-dependent variable relationship, before accounting for study characteristics, will also be a secondary outcome of this review.

1.8 Strategy for data synthesis

Meta analytic methods will be used to synthesise data in subsequent analyses. Methods may include, but are not limited to: statistical meta-analysis to obtain meta-elasticities, bivariate analysis including contingency tables, meta-regression analysis.

1.9 Analysis of subgroups or subsets

This is the primary aim of the study. Studies will be segmented based on trip purpose, case study location and population demographic subgroups. Depending on the patterning in indicators studied, there may be an opportunity for common groups of variables to form segmentation groups.

1.10 Risk of bias

Screening and extraction bias is a risk inherent in this review, that are not remedied. No adjudicator has been used due to resourcing constraints. Bias is minimized through careful selection and definition of the attribute levels relevant to the study, guided by theories of meta-analysis and built environment – travel behavior study design (Boarnet, M & Crane 2001; Boarnet, MG 2011; Handy 1996; Stanley & Doucouliagos 2012). A validation step was taken to expand the sample and compare the records identified to those that formed the sample of other recent meta-analyses of the built environment and travel behavior.

1.10.1 Within-study bias

Analyses will be conducted using error-weighted values to give greater emphasis to relationships that are explained to large extents by the variables of interest.

1.10.2 Meta-bias

Test for reporting bias will be employed in subsequent analyses. Tests such as Egger's asymmetry test will be used; however such test do not enable the source of bias to be detected. These may include selection bias, delayed or positive reporting bias and sample bias (Egger, Schneider & Minder 1997).

2. Identification strategy

2.1 Concepts and terms of the search strategy

The question motivating the review was "Which features of the built environment influence transit ridership, by how much, and under what conditions?" The database search was designed around a simplification of this question: "What is the relationship between the built environment and transit use?" Four key words and their alternative phrases were entered into eight databases on 7 December 2017. The keywords were "public transport", "ridership", "built environment" and "relationship" (Table 1).

Table 1 - Key words and alternatives used in database search

Concept 1: Public transport	Concept 2: Ridership	Concept 3: Built environment	Concept 4: Relationship
transit	“mode choice”	“built environment”	Significan*
“public transport”	ridership	“land use”	Relationship
Bus	“mode share”	“urban design”	Statistic*
Tram	modeshare	Distance	Predict*
train		Access*	Model
		density	

2.2 Identify and select the electronic databases

The eight databases used were selected based on their content descriptions and validated based on recently published transit reviews. Table 2 lists the candidate databases. The databases chosen for the comprehensive search were Scopus, TRID, Web of Science (Core Collection), Transport, ATRI, Proquest, Compendex and INSPEC.

Table 2 - Databases used in prior studies or considered based on descriptions

	Stevens 2017 (Stevens 2017)	(McCarthy et al. 2017)	Proposed (2017)
Academic search premier	✓		
Google	✓		
Google Scholar	✓		
Medline	✓		
PAIS International	✓		
PUBMED	✓		
Scopus	✓	✓	✓
TRIS online	✓	✓	✓
TRANweb	✓		
Web of Science	✓	✓	✓ (Core collection)
Web of Knowledge	✓		
Transport			✓
ATRI			✓
Proquest			✓
Compendex		✓	✓
INSPEC			✓
World transit research		✓	

2.3 Pilot the search strategy

Table 3 - Keyword and alternative search results (Pilot conducted 7 December 2017)

Concept 1: Public transport				Concept 2: Built environment				Concept 3: Ridership				Concept 4: relationship			
Keywords	Scopus	Comp.	Transp.	Keywords	Scopus	Comp.	Transp.	Keywords	Scopus	Comp.	Transp.	Keywords	Scopus	Comp.	Transp.
Transit	91,551	863,414	56,452	“built environment”	13,862	25,651	1185	“mode choice”	3,083	1,647	5,922	Significan*	8,909,121	2,048,912	75,568
“public transport”	13,857	8,005	27,722	“land use”	124,442	46,946	14,326	Ridership	1,692	1164	5,516	Relationship	3,265,432	697,622	22,099
Bus	109,307	120,659	34,907	“urban design”	7,000	3,738	1,062	“Mode share”	456	213	308	Statistic*	2,927,107	870365	46580
Train	16,4020	430,402	18,837	distance	10,19767	512,684	2,1533	modeshare		0	0	Predict*	775,510	1,411,469	5,4396
tram	5,418	3,140	2,492	Access*	137,5690	574,166	33,490					model	10,133,462	5,114,535	116,764
				density	2,495,181	1,353,646	17,745								
Total	Transit or “public transport” or bus or train or tram			Total	Density or access* or distance or “urban design” or “land use” or “built environment			Total	"mode choice" or ridership or “mode share”			Total	significan* OR relationship OR statistic* OR predict* OR model		
	365,954	1,400,148	105,236		4,852,790	2,424,254	80,961		5,017	2,895	11,293		20,314,169	7,855,668	24,3857
1 AND 2 AND 3 AND 4	(Transit or “public transport” or bus or train or tram) AND (Density or access* or distance or “urban design” or “land use” or “built environment) AND ("mode choice" or ridership or “mode share”) AND (significan* OR relationship OR statistic* OR predict* OR model)												689	453	1151
1 AND 2 AND 3	(Transit or “public transport” or bus or train or tram) AND (Density or access* or distance or “urban design” or “land use” or “built environment) AND ("mode choice" or ridership or “mode share”)												947	601	2182

2.4 Define the search phrase

The base search string formulated using the key words and alternatives in Table 1 is shown below.

Adaptations for each database are shown in Table 4.

(transit OR "public transport" OR bus OR train OR tram) AND (density OR access OR distance OR "urban design" OR "land use" OR "built environment") AND ("mode choice" OR ridership OR "mode share") AND (significan* OR relationship OR statistic* OR predict* OR model)*

2.5 Run the search and record results

The results returned for each of the eight databases selected are summarized in Table 4 along with the specific search string and any filters applied for each database. A total of 7,289 records were retrieved.

Table 4 - Search strings and records retrieved by database (date: 7/12/17)

ATRI : Australian transport index	(Transit or "public transport" or bus or train or tram) AND (<i>Density or access* or distance or "urban design" or "land use" or "built environment"</i>) AND ("mode choice" or ridership or "mode share") AND (significan* OR relationship OR statistic* OR predict* OR model)	166
TRID: the TRIS and ITRD database	(transit OR "public transport" OR bus OR train OR tram) AND (density OR access* OR distance OR "urban design" OR "land use" OR "built environment") AND ("mode choice" OR ridership OR "mode share") AND (significan* OR relationship OR statistic* OR predict* OR model)	2193
TRANSPORT	(transit OR "public transport" OR bus OR train OR tram) AND (density OR access* OR distance OR "urban design" OR "land use" OR "built environment") AND ("mode choice" OR ridership OR "mode share") AND (significan* OR relationship OR statistic* OR predict* OR model)	1151
Compendex	(Transit or "public transport" or bus or train or tram) AND (<i>Density or access* or distance or "urban design" or "land use" or "built environment"</i>) AND ("mode choice" or ridership or "mode share") AND (significan* OR relationship OR statistic* OR predict* OR model)	453
INSPEC	(((((Transit OR "public transport" OR bus OR train OR tram) WN All fields) AND ((density OR access* OR distance OR "urban design" OR "land use" OR "built environment") WN All fields)) AND (("mode choice" OR ridership OR "mode share") WN All fields)) AND ((significan* OR relationship OR statistic* OR predict* OR model) WN All fields))	137
Scopus	(Transit or "public transport" or bus or train or tram) AND (<i>Density or access* or distance or "urban design" or "land use" or "built environment"</i>) AND ("mode choice" or ridership or "mode share") AND (significan* OR relationship OR statistic* OR predict* OR model)	689
Web of science Core Collection	TOPIC: (transit OR "public transport" OR bus OR train OR tram) AND TOPIC: (Transit OR "public transport" OR bus OR train OR tram) AND TOPIC: ("mode choice" OR ridership OR "mode share") AND TOPIC: ("mode choice" OR ridership OR "mode share") Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC.	1743
ProQuest	all(Transit OR "public transport" OR bus OR train OR tram) AND all(density OR access* OR distance OR "urban design" OR "land use" OR "built environment") AND all("mode choice" OR ridership OR "mode share") AND all(significan* OR relationship OR statistic* OR predict* OR model)	757
Total exported		7,289

2.6 Collate results

Results saved and exported as RIS files to Endnote Library. Exported as text file to excel, with journal article fields: title, author, year of publication, database used, name of journal. Comparable fields were specified for non-journal article literature.

3. Screening

3.1 Automatic filtering of results

Screening of records was undertaken in three stages. The first stage, automatic filtering, removed 3,580 records which were either duplicates, ineligible source types (books and conference proceedings), or were published prior to 2000. This step was undertaken in excel by filtering based on the basic citation information for each record. Eligible sources included journal articles, serials, reports, conference papers and theses.

Table 5 - Results of duplicate removal and filtering by source type and publication year

Filter applied	Description of filter	Excluded	Remaining
Duplicates	Manual screen for matching author/title/abstract	760	6,529
Source type	Filter by "source type". Include only: Journal article, serial, report, conference paper, TRID, thesis (Exclude: book, book chapter, conference proceedings, generic)	1,491	5,038
Pre 2000	Include only texts published dated 2000 or later	1,331	3,707

A total of 3,707 studies remained after automatic removal of duplicates and filtering by study type and year of publication.

3.2 Title/abstract screen

The second stage of screening involved reviewing the title and abstract of remaining studies. Abstracts and titles were screened to manual check for duplicates, assess functional elements of records, and check for relevant dependent and independent variables. Stanley and Doucouliagos (2012) state that unpublished studies or grey literature can be important for identifying differences in research methods that lead to certain results; or can include more recent results, and introduced new publication bias. Thus reports, conference papers and theses were included as eligible source types.

Studies were required to contain a quantitative effect size expressing an observed relationship between one of the transit travel behavior variables and built environment predictor variables. Table 6 provides a complete list of the eligible modes and variables, as well as the other criteria applied during abstract and title screening.

Table 6 – Criteria applied to screen records based on information in title and abstract

Parameter	Categories	Assessment
(1.1) Duplicates		Exclude
(1.2) Language	English/French	Include
	Other	Exclude
(1.3) False result	Y/N	Exclude
(1.4) Analysis framework	Clear delineation of independent and dependent variables	Include
	Narrative	Exclude
	Methodological/conceptual only	Exclude
	Forecast	Include (TBC)
(1.5) Source type	Specified in scope (journal article, serial, thesis, conference paper, report)	Exclude
	Other (e.g. table of contents, conference proceedings, book, workshop paper)	Exclude
(2.1) Transit travel behavior indicator	Mode choice or mode share, ridership, willingness to pay	include
	other (VMT, VKT, trip frequency, trip length)	Exclude
	none	exclude
(2.2) Modes considered	Private automobile, walk, bicycle, taxi, motorcycle/scooter	Exclude if no transit modes
	Transit (general), Bus, BRT, light rail/tram/streetcar, train/subway/heavy rail/metro, train/subway/heavy rail/metro	Include
	High speed rail (exclude)	Exclude
	Ferry (exclude)	Exclude
	Other mode [specify]	Exclude if no transit modes
(3.1) Independent variables	Built environment: Density, accessibility, land use, urban design, spatial configuration, distance	Include
	External and system (non-built environment) : psychosocial, sociodemographic, demand management, service quality	Exclude if no built environment indicators
	Other suspected built environment variable not captured in this list [specify]	Include

From the 3,707 studies screened, a total of 2,271 studies were excluded based on the functional screening criteria, including a further 1,523 duplicates. 591 studies were screened out due to irrelevance of the dependent variable, while 300 studies of the remaining studies were screened out on the basis of irrelevant independent variables.

Table 7 - Results of title and abstract screening

	Criteria	Excluded	Remaining
(1) Functional screen	Duplicate (exclude)	1,523	2,184
	Language other than English/French	9	2,175
	False result/ freight	30	2,145
	Not quantitative	316	1,829
	Source type out of scope	30	1,799
	Subtotal	1,908	1,799
(2) Relevance: dependent variable		646	1,153
(3) Relevance: independent variable		434	719

3.3 Eligibility based on full text screen

The full text of the remaining studies was skimmed to validate eligibility based on the criteria listed in Table 8. The purpose of this stage of screening was to include only those studies that yielded a quantified relationship between a transit mode choice or ridership variable, and one or more built environment variables. Due to the size of the sample, only indicators of density, diversity and design were included after this stage.

Table 8 - Study and sample characteristics for determining eligibility of studies

Parameter	Categories	Assessment
2.2 Modes considered	Transit (general) or a transit mode is included as a standalone variable	Include
	Transit (general) or a transit mode only features as part of an aggregate mode indicator with non-transit	Exclude
3.1 Independent variables	Density, diversity, design, destination accessibility	Include
	Other built environment indicators	Exclude
3.2 Regression/Discrete Choice	Yes (must include regression coefficients, odds ratios or elasticities)	Include
	No	Exclude

In addition, studies were eliminated if they did not employ a statistical regression technique to produce the effect size between travel behavior (outcome) and built environment (predictor) variables. 585 studies were eliminated through full text screening, leaving 135 studies.

3.4 Validation

In the absence of an adjudicator, the search process was validated by comparing the studies identified to those captured in the 2010 meta-analysis as well as the meta-analysis of the built environment and vehicle miles travelled, conducted in 2017 (Ewing & Cervero 2010; Stevens 2017). In addition, a synthesis of studies exploring residential location choice and travel behaviour was consulted to identify additional relevant sources {Cao, 2009 #23737}. Of the 107 unique studies included in these syntheses, 20 met the eligibility criteria for the present study. Of these, nine had been identified through the search strategy. A summary of eligible studies at each stage of screening is shown in Table 11. The additional eleven studies that met the eligibility criteria for the present study were added to the sample. This brought the total number of studies eligible for extraction to 146.

Table 9 - Validation of comprehensive search based on precedent studies

	Overall (overlap)	(Ewing & Cervero 2010)	(Stevens 2017)	{Cao, 2009 #23737}
Total	107 (29)	61	37	38
Meeting title abstract screen criteria	40			
Meeting full text screen criteria	20			
Of which retained after title abstract screen	9			
Added to sample	11			

3.5 Screening summary

Error! Reference source not found. summarizes the comprehensive search methodology.

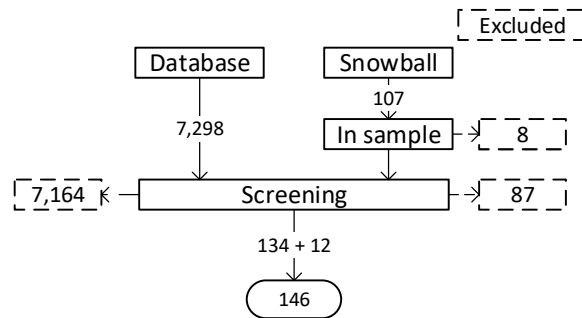


Figure 1 - Summary of priors screening and final sample

4. Extraction

Data was extracted manually from each record. Each of the 146 records included in the final sample were read thoroughly. The magnitude of a relationship between a transit use (dependent or outcome variable) and the built environment (independent or predictor variable) was the primary datum of interest for this study. Therefore data was extracted for each unique relationship. All data items were sought for every data point. Data were entered into an excel spreadsheet, with each row entry representing a unique data point.

4.1 Data point identification

Each unique relationship was assigned a ‘data point ID’, made up of the study ID, model ID and data point number. Therefore each data item is identified by an ID taking a form such as:

16.2.7

Where:

- ‘16’ is the Study identifier
- ‘2’ is the model identifier for that particular study
- ‘7’ is the data point identifier for the particular model

4.2 Data for extraction

Table 12 provides a summary of the data collected for each variable. A complete list of data extracted for each data point is available as a supplement to this protocol, as a metadata file. Where a detailed explanation of attribute levels is warranted, it is provided in the section that follows. In these cases, a table reference is provided in the right-hand column of Table 12.

Table 10 - Summary of characteristics extracted from priors for each data point (refer to metadata in database for full description and attribute levels coded)

Data type	Description and attributes		Detail
Data point identification	Identification numbers, abbreviated citation information and table/page reference		
	Analysis groupings	Notation used to segment data points into “like” groupings that measure comparable relationships	Refer to metadata
Data point typology	Significance	Whether the data point is significant or not	Table 11
	Direction	positive, negative, or ‘omitted’ if no relationship is reported	
	Theoretical consistency	Adherence to hypothesis (yes/no) or insignificant)	Table 19
Dependent (travel behavior) variable	Transit travel behavior variable	E.g. mode choice (probability), mode share and ridership	Table 12
	Mode of dependent variable	Any combination of: transit (general), park’n’ride connecting to transit (general), bus, bus rapid transit, school bus, light rail transit, heavy rail, shuttle or demand responsive transit	Table 13
	Aggregation level, unit of analysis for aggregate dependent variables, data type, data source		
Sample attributes	Population subset	E.g. workers, located in a TOD	Table 14
	Trip subset	E.g. morning peak, inter-peak	Table 15
	Country, city and year of data collection		Categorical
Study Design	Sample size		Continuous
	Covariance testing		Table 16
	Experiment type		
	External variables		Table 17
	Variables to account for specification issues in BE-TR (self-selection and regional accessibility)		Table 18
Independent (built environment) variables	Independent built environment variable	Density, Diversity or design	Table 19 (definitions) Error! Not a valid result for table. (assumptions)
	Indicator used to measure the D-variable		
	Standardized indicators; grouped into the measurement of ‘like’ attributes		
	Measurement of the built environment	Journey leg	Table 21
		Unit of analysis	Table 22
Model specification	Estimation method	Logistic, ordered probit, ordinary least squares, hierarchical regression	Table 23
	Reference case	For discrete choice models, the reference case is recorded.	
	Functional form	Form of the dependent and independent variables (i.e. log or linear)	
Quantitative data	Magnitude of the relationship between the travel behavior (dependent) variable and the built environment (independent) variable		
	p-value and t-statistic		
	Standard error of the regression coefficient		
	Standard error of the elasticity, calculated from the source or derived using data supplied		
	Standard error of the elasticity, imputed using the average of analysis subgroups		
Data for imputation	Measure of the variance explained by the model (e.g. R^2)		
	Mean and standard deviation of the dependent and independent variables where available		

4.3 Attribute coding definitions

4.3.1 Data point typology definitions

‘Omitted’ variables: Data point corresponds to a relationship that is not reported because it was found to be insignificant

- In some cases, the variable was displayed with a blank entry in the summary of model outputs.
- In some cases the author explicitly stated that the variable was omitted due to either insignificance or collinearity; however in most cases the reason for omission from the final model was not explicitly stated.

Table 11 - Definition of significance thresholds

Code	Definition
Y	Significant: <ul style="list-style-type: none"> • $P < 0.1$ • Author’s assessment (no statistics given to evaluate) • $t > 1.5$
N	Insignificant: does not meet above criteria
Not stated	

4.3.2 Dependent (transit travel behavior) variable definitions

Table 12- Standardized transit travel behavior variables

Indicators of the dependent variable differentiated between mode choice (probability), ridership and mode share. Any probability estimates were also distinguished based on the reference case of the dependent variable.

Dependent Variable (DV)	Reference Case of DV.	DV_abbrev.
Discrete methods		
Transit use (probability), Preference (probability)	No transit trip	Transit use
Mode choice (probability)	Single mode, includes automobile, shared ride, walking	Mode choice (single alternative)
Mode choice (probability)	Any mode other than transit	Mode choice (multi alternatives)
Continuous methods		
Mode share	n/a	Mode share
Ridership, transit users	n/a	Ridership
Ridership frequency, ridership thresholds (probability)	n/a	Frequency

Table 13 - Travel mode definitions

Samples may comprise multiple transit modes whilst still being selective. As such, the sample was characterized as measuring (1) or not measuring (0) travel by the following modes.

	Code	Definition
Specific mode	Transit	Mode of the travel behavior (dependent) variable is transit, not otherwise distinguished as any of the other categories of transit
	PNR	Park'n'ride
	Bus	Local bus
	BRT	Bus rapid transit or commuter bus
	Sch_bus	School bus
	LRT	Light rail, tram or street car
	HR	Heavy rail, subway or metro
	Shared	Demand responsive transit or shuttle service

4.3.3 Sample attributes definitions

Table 14 - Population subset definitions

Samples may comprise multiple population segments whilst still being selective. As such, the sample was characterized as belonging (1) or not belonging (0) to each of the subsets below.

	Code	Description/assumptions
Specific	General	No population mentioned, trips not limited to work trips
	Worker	Applies to all data points where work trips are measured exclusively
	Adult	Includes samples limited to age 16+; all 'workers' assumed to be adult
	Child	If specified as school aged or under 16/18
	Aged	Described as 'older' or 'aged' in discussion of sample
	Transit user	Applied to all data points sourced from patronage data
	High transit	Sample drawn from a location with high transit access or high transit use
	Low transit	Sample drawn from a location with low transit access or low transit use
	High income	Described as 'high income' in discussion of sample
	Low income	Described as 'low income' or 'disadvantaged' in discussion of sample
	TOD/high density/mixed use	Sample drawn from a location described as being "transit-oriented", or high density or mixed use.
	Non TOD/ low density	Sample drawn from a location described as being not "transit-oriented", or low density or comprising homogeneous land use.
	Other	Categorical entry

Table 15 - Trip purpose definitions

Data points were classified as one of the following that best described the trip purpose

	Category	Inclusions
	All	Trip purpose not restricted
	Home-based	Trip originating at the home, not aligned to a more specific category
	Work	Work-trip
	Non-work	None-work trip
Specific	Peak	Trips taking place during peak hour, but not restricted to commute trips
	Off Peak	Trips taking place during off-peak periods
	Weekday	Trips taking place on a weekday
	Weekend	Trips taking place on a weekend
	Station Access	Trip is made to access a transit station
	Other	Includes trips specified as socialisation, discretionary, consumption, maintenance, shared or other specific trip purpose not covered by other categories.

4.3.4 Study Design Definitions

Table 16 – Covariance testing definitions

Where collinearity or covariance testing was not otherwise detected from a reading of each record, a keyword search for the following terms was undertaken: "cova" "Multi" "Col" "interact*" "VIF", "Moran*" (PCA, factorisation, clustering), X'X' identity matrix*

Code	Criteria
Yes	Evidence of the covariance testing E.G., factorisation, VIF, exclusion of variables (theoretical), covariance matrix, Moran's I test.
No	No evidence of covariance testing

Table 17 - External variables and indicators in built environment and travel behavior research (Source: author's synthesis)

Each of the following variables commonly associated with travel behavior and built environment interactions (Ewing & Cervero 2010) considered 'controlled' (1) if one or more indicators for each appear in the final model, or if reference made to including them in initial rounds of model calibration.

Variable		Indicators
Direct costs		Fare, fuel price/index, parking cost, relative cost
Indirect costs	Trip characteristics	Journey length, duration, pleasantness of journey
	Level of service	frequency, vehicle quality, coordination, Supply Index, PTAL, SNAMUTS
Distance to transit		Transit stop density, distance to nearest transit stop, Supply Index
Demand management		parking restraint or charging, congestion charging, fare level
Additional residential self-selection criteria		Neighbourhood preference stated
Sociodemographic		income, age, gender, employment type and status
Other		Time of day, weather, attitudes

Table 18 – Methodological strategies to account for attitude-induced self-selection

Specification issue	Description of appropriate control		
Regional accessibility	<ol style="list-style-type: none"> Distance to CBD Number of jobs accessible by transit from survey location within time period 		
Self selection (demographic)	<ol style="list-style-type: none"> Median income, household income Number of children Employment status or industry 		
Self-selection (Attitudinal) (Mokhtarian & Cao 2008)	<ol style="list-style-type: none"> Explicit measures of attitudes and preferences 	<ol style="list-style-type: none"> Non-explicit (instrumental) measures 	<ol style="list-style-type: none"> Sample selection
	<ul style="list-style-type: none"> - Direct (surveyed) measures of attitudes - Stated or revealed preference for <ul style="list-style-type: none"> o residential location o lifestyle o Personality - Attitudes about <ul style="list-style-type: none"> o Travel o Neighbourhood type o Environmental values 	<p>(Boarnet, MG & Crane, R 2001, pp. 95 - 6; Boarnet & Sarmiento 1998):</p> <ul style="list-style-type: none"> - % housing stock built before 1940 - % housing stock built before 1960 (Greenwald & Boarnet 2001): - % housing units classified as rural but not classified as farms - % housing units classified as urban dwelling units 	<p>Discrete types of residential location such as urban vs. suburban (Heckman 1976; Weisbrod, Ben-Akiva & Lerman 1980)</p>

4.3.5 Independent (built environment) variable definitions

Table 19 - Standardized built environment indicators

The independent built environment variables are characterized by the high level 'D' category into which it falls and the indicator used to operationalise this variable

Scale	'D' variable	Indicator	Definition	Hypothesised relationship to transit use
N	Density	Employment density	Gross or net jobs/employment per area (total or by type of employment)	+
N		Population density	Gross or net households, dwellings or persons per total area	+
N		Activity Density	Active floor space ratio, sum of jobs and population and/or commercial/retail opportunities per area	+
N		Commercial Density	Commercial or retail density, number of establishments, commercial or retail land use proportion	+
N	Diversity	Land use mix	mix of land use (floor area), vertical mix of land use, Bhat/Shannon/ Simpson/Heip/McIntosh/Smith-Wilson measures of evenness or diversity (normalised so that maximum score = 1)	+
N		Housing mix	Mix of housing type, mix of housing affordability, mix of tenure type, ethnic diversity of 20neighbourhood (normalised)	+
N		Attraction/Generation balance	Jobs-housing balance, ratio of trip origins to trip destinations (normalised)	+
micro	Design	Pedestrian and cycle amenities	Pedestrian or cycle amenities: canopy, pleasantness, street furniture, facilities	+
Micro		Pedestrian and cycle amenities (categorical)	Building setback, building orientation, pedestrian orientation, 20neighbourhood type, building age, presence of sidewalk, grid street network	+
N		Personal and physical safety	Lighting, perception of security, curbs, shoulder width, perception of safety from vehicle traffic	+
N		Pedestrian and cycle connectivity	Total path length, number of pedestrian crossings, average footpath width, footpath density, crossing density, link to node ratio, intersection density	+
N		Automobile connectivity	road network density, block size, path impedances, road length per capita, cul-de-sac or dead-end density, street segment length	-
Micro/N		Dis-amenity, barriers (categorical)	appearance, upkeep, homogeneity, lack of safety/security, isolation or lack of local access, Curvilinear street type, new development	-
N	Access	Local access/ walkability	WalkScore, local living score, count of services/mixed use opportunities, CBD dummy, TOD, perception of convenience of access to amenities and activities	+
N/macro		Accessible destination (categorical)	Located in CBD, TOD or close to transit, high density, above average density	+
Macro		Centrality	Proportion of population within CBD, Compactness index	+
Macro		Decentralization	Sprawl index, distribution of population or employment, city shape	-
N		Lack of local access or walkability (dummy)	Located in low transit/ low access/ low density area, Low density, below average density	-

Scale notation:

N – variable with impact at the neighbourhood level

Micro – variable with impact at line of sight

Macro – variable with regional impact

Table 20 - Assumptions pertaining to interpretation and extraction of built environment indicators

Interpretation
CBD Dummy treated like TOD – accessible destination
Categorical, factorized or clustered indicators were interpreted and classified qualitatively.
Example: Study 15– present 20 clusters. Judged on 5 variables: density, housing age, street density, intersection density, cul de sac density. Judged to be either auto dependent (des.1) or pedestrian friendly (des.2) depending on into which category the majority of scores fell.
Many studies included ‘population’ or ‘employment’ as indicators; however these were standardized to ‘population density’ and ‘employment density’ where the analysis area was uniform, otherwise excluded (if the catchment area was not uniform and therefore the measurements represented absolutes).
Exclusions
Some indicators were removed following an initial round of coding due to a determination that they were not comparable, or were not strictly indicators of density, diversity and design. Exclusions:
<ul style="list-style-type: none"> • Counts of roads for non-equivalent sized units of analysis (counts of amenities and paths were included) • Residential proportions without intensity or diversity cannot yield meaningful hypotheses from a mode choice/transit use perspective (same as above → relates to trip generation). Also excluded. • Land use areas for non-equivalent sized units of analysis, for commercial, office, retail etc. (land use trip generation is a separate field of enquiry to the more granular built environment field)

Table 21 - Location of built environment measurement definitions

Code/name	Definition/assumptions
Origin	Built environment of the trip origin or home (aggregate) only is measured
Destination (Dest)	Built environment of the trip destination or workplace (aggregate) only is measured
Both	Built environment of the trip origin and destination are measured for the same data point
Any	Built environment maximums are measured; and may come from trip origin, destination or tour destinations along the way
Not distinguished (ND)	Study does not identify whether built environment is measured at origin or destination
n/a	Not applicable

Table 22 - Unit of analysis of independent variables - definitions and assumptions

Boundary, or unit of analysis for built environment, recorded as stated in description of data. Assumptions and groupings were made to make data more comparable.

Assumptions
<ul style="list-style-type: none"> • ¼ mile converted to 400m • ½ mile converted to 800m • ¾ mile and greater left in imperial units • Assumed radial Euclidean catchment unless otherwise indicated (e.g. network) • If grid, linear or radial nature unspecified, assumed radial

4.3.6 Estimation method attributes

Table 23 – Relationship estimation methods and associated additional information

Estimation methods	Additional information
2SLS	
3SLS	
Bootstrapped regression	
Global regression	
GWR	
Hierarchical logistic regression	• weighting functions
Hierarchical regression	• binary, binomial, multinomial, fractional multinomial
Hierarchical negative binomial regression	• multivariate, multilevel
linear probability equation	• linear, linear and nonlinear
Logistic	• Bernoulli trial
Mixed multiple-discrete continuous extreme value	• fixed effects, random effects
multi-level model	• log-log (form)
Multiplicative	• long run, short run
Negative binomial regression	• pooled
OLS	• nested
Ordered logistic	• partial adjustment
Ordered probit	• stepwise
Poisson regression	• Tobit (censored)
Seemingly unrelated regression	
SEM	
Simultaneous Auto-Regression (SAR)	
Simultaneous logit model	
Tobit	

4.4 Data management

The complete database comprising all data emerging from the present review is retained by the author. Subsets of the data will be published along with the corresponding analysis, with a view to publish the complete set of data. As with this protocol, each dataset will be deposited in an online repository, identifiable by a DOI. The results of the review are documented in this paper, to ensure transparency over the quantity of records that emerged from the review.

References

- Boarnet, M & Crane, R 2001, 'The influence of land use on travel behavior: specification and estimation strategies', *Transportation Research Part a-Policy and Practice*, vol. 35, no. 9, pp. 823-45.
- Boarnet, MG 2011, 'A broader context for land use and travel behavior, and a research agenda', *Journal of the American Planning Association*, vol. 77, no. 3, pp. 197-213.
- Egger, M, Schneider, M & Minder, C 1997, 'Bias in meta-analysis detected by a simple, graphical test', *British Medical Journal*, vol. 315, no. 7109, pp. 629-34.
- Ewing, R & Cervero, R 2010, 'Travel and the built environment: A meta-analysis', *Journal of the American Planning Association*, vol. 76, no. 3, pp. 265-94.
- Handy, S 1996, 'Methodologies for exploring the link between urban form and travel behavior', *Transportation Research Part D: Transport and Environment*, vol. 1, no. 2, pp. 151-65.
- Haustein, S & Nielsen, TAS 2016, 'European mobility cultures: A survey-based cluster analysis across 28 European countries', *Journal of Transport Geography*, vol. 54, pp. 173-80.
- McCarthy, L, Delbosc, A, Currie, G & Molloy, A 2017, 'Factors influencing travel mode choice among families with young children (aged 0–4): a review of the literature', *Transport Reviews*, vol. 37, no. 6, pp. 767-81.
- Ortúzar, JD & Willumsen, LG 2011, *Modelling Transport*, Fourth edn, John Wiley & Sons Ltd, Chichester, West Sussex, United Kingdom.
- Stanley, TD & Doucouliagos, H 2012, 'Meta-regression analysis in economics and business', in Routledge, London and New York.
- Stevens, M 2017, 'Does Compact Development Make People Drive Less?', *Journal of the American Planning Association*, vol. 83, no. 1, pp. 7-18.
- The Prisma-P Group 2015, *PRISMA-P: Preferred Reporting Items for Systematic review and Meta-Analysis Protocols 2015 checklist: recommended items to address in a systematic review protocol*, viewed 8/7/17, <<http://www.prisma-statement.org/documents/PRISMA-P-checklist.pdf>>.
- The World Bank & Food and Agriculture Organisation 2017, *Population density (people per sq. km of land area)*, viewed 20/12/17 2017, <<https://data.worldbank.org/indicator/EN.POP.DNST?view=chart>>.
- The World Bank & Organisation for Economic Cooperation 2017, *GDP per capita (US\$)*, viewed 19/12/17 2017, <<https://data.worldbank.org/indicator/NY.GDP.PCAP.CD>>.