Urban developments and daily travel distances

Fixed, random and hybrid effects models using a Dutch pseudo-panel over three decades

Kasraian, Dena; Maat, Kees; van Wee, Bert

DOI
10.1016/j.jtrangeo.2018.09.006

Publication date
2018

Document Version
Final published version

Published in
Journal of Transport Geography

Journal of Transport Geography 72 (2018) 228–236

Contents lists available at ScienceDirect

Journal of Transport Geography

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

Urban developments and daily travel distances: Fixed, random and hybrid effects models using a Dutch pseudo-panel over three decades

Dena Kasraian[a,*], Kees Maat[b], Bert van Wee[c]

[a] Department of Civil and Mineral Engineering, University of Toronto, 35 St. George Street, Toronto, ON M5S1A4, Canada
[b] Faculty of Architecture and the Built Environment, Delft University of Technology, Julianalaan 134, 2628, BL, Delft, the Netherlands
[c] Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628, BX, Delft, the Netherlands

ARTICLE INFO

Keywords:
Land use
Daily travel distance
Long term
Pseudo panel analysis
Hybrid fixed- and random-effects
Randstad

ABSTRACT

As people require time to adjust their travel behaviour to changes in residential location and transport infrastructure, there is a need for long-term empirical studies quantifying the relationships between locations, individuals and travel behaviour. Such empirical evidence is critical for assessing previous and candidate future land use-transport policies. Existing research however, has mostly investigated travel behaviour during relatively short time periods and for a single transport mode. This paper examines the development of travel behaviour and its socio-demographic and location determinants, using Dutch National Travel Survey data from 1980 to 2010 among other sources, for the Randstad, the Netherlands. A pseudo panel analysis is conducted to investigate the effect of various indicators on average daily distance travelled by train, car and bicycle over three decades. Econometric models including pooled ordinary least squares, fixed and random effects and a hybrid model were tested to identify the best fit. The results indicate that average daily distance travelled rose until the mid-1990s before witnessing a decrease till 2010. Interestingly, half of the Randstad inhabitants have been travelling ≤ 26 km per day over the past thirty years. Furthermore, as people grow older, they increasingly travel more by train and bicycle. Finally, a rise in suburban inhabitants decreases the average distance travelled by train and increases that of bicycle, while a rise in rural inhabitants encourages higher distances travelled by car.

1. Introduction

Measuring and modelling individual travel behaviour is highly relevant for infrastructure decisions and policies for achieving sustainable environmental and societal development. Travel is the result of decisions by which individuals try to meet their needs and preferences. They aim to achieve their goals by allocating and prioritising their activities, thereby taking into account the relative position of locations. It is assumed that distances between residential, employment and service locations directly affect individual's total travel distances, as nearby destinations will be chosen rather than more distant ones (Maat and Timmermans, 2009). Consequently, it is assumed in this paper that travel behaviour is determined by the structure of the built environment, including the location of urban cores in relation to suburbs, the rural area, and other urban cores, as well as their accessibility by transport infrastructure connections. This is a dynamic process in which travel influences the demand for infrastructure investments, leading to improvements in accessibility and thus the attractiveness of locations, which in turn encourages adjustments to the built environment (Giuliano, 2004; Wegener and Fürst, 1999). This market-driven process is also subject to exogenous influences, such as the demand for housing, developments in transport technology and changing views on environmental sustainability (Bertolini, 2012; Kasraian et al., 2016b). Policy makers aim to adjust this market process by inventing policy concepts such as 'compact urbanisation'. However, policy concepts also change over time.

Changes in transport infrastructure and the built environment, as well as changing policy responses, are assumed to have varying effects on travel behaviour. Moreover, we assume that there is a certain degree of delay in the system. Travel behaviour requires time to adjust to changes in the spatial context, new transport infrastructure and changing policies. It takes time for households to relocate to a new residential or work location, or to relocate their other activities, such as shopping to new locations. Furthermore, all market-driven developments and policy responses have their own time horizons. It is therefore the aim of this study to understand the effects of the built environment on travel behaviour, taking into account the adaptation of spatial policy concepts, over a longer period of time.

* Corresponding author.
E-mail address: D.Kasraian@utoronto.ca (D. Kasraian).

https://doi.org/10.1016/j.jtrangeo.2018.09.006
Received 22 January 2018; Received in revised form 6 September 2018; Accepted 14 September 2018
Available online 27 September 2018

Cross-sectional studies on the determinants of travel behaviour cannot provide answers to whether and to what extent changes in various built-environment and socio-demographic determinants affect the demand for travel. Studies quantifying the relationships between locations, individuals, and travel behaviour development over time are scarce (Ellder, 2014). This is mainly due to the unavailability of longitudinal travel surveys over a long time period. Of the studies that investigated this relationship over the long term, some analysed changes in travel patterns and its determinants at an aggregated level such as municipalities or tracts. Examples are induced travel demand studies (Noland and Lem, 2002), 'before-and-after' studies which test for instance the effect of new infrastructure (Baum-Snow and Kahn, 2000). Another strand of research investigated changes in travel behaviour at the individual level over time, using genuine panel data, such as the Dutch and the German mobility panels. Examples are studies which examine the concept of self-selection with longitudinal designs (e.g. Van de Coevering et al., 2016).

However, genuine panel data, where the same individuals are traced over time, are often costly, small-scale and suffer from sample attrition problems. In the absence of genuine panels, the most long-term and accessible travel behaviour data are repeated cross-sectional data. A major advantage of independent cross-sectional samples is that they are available over longer periods of time, such as the Dutch National Travel Surveys, a data set we used in this paper. Another advantage is that independent samples are not affected by dropout. A major disadvantage, however, is that they do not provide information for the same respondents across time, making it impossible to analyse intra-personal dynamics. A limited number of studies have used repeated cross sectional data from several survey waves to measure the change in factors influencing travel behaviour over time (Feng et al., 2017; Guerra, 2014; McDonald, 2015; Susilo and Maat, 2007; Zegras and Hannan, 2012). The majority of these studies has investigated changes in travel behaviour over two or three time points, by pooling the data sets on households and using interaction with time dummies or the test of preference stability to identify the effect of variables for specific time points. This method retains the individual characteristics and is easy to interpret. However, as each time point contains different observations, the data do not have a longitudinal structure where the same panel units are observed over time.

Alternatively, existing repeated cross-section samples can be restructured to behave as genuine panel data with temporal ordering, where changes within a panel unit and their determinants can be measured over time. This requires the construction of pseudo panels where individuals are grouped into homogenous groups of observations over time. Under certain conditions these groups can be treated as genuine panel units (Van de Coevering et al., 2016). Thus a trade-off is made between keeping individual characteristics and obtaining a panel structure where the same units (here groups of people) are traced over time. While this method entails a loss of individual characteristics, the resulted panel structure satisfies an important condition for causal inference, namely the temporal precedence of cause and effect. In the transport field, this method has been applied to repeated cross-sectional data to mainly model car ownership (Dargay, 2002) and public transport demand (Tsai et al., 2014).

By using such a pseudo panel analysis, this paper investigates the dynamics of daily distances travelled, related to characteristics of the built environment. The study is guided by the research question how travel behaviour has developed from 1980 to 2010 in the Dutch Randstad, in terms of distances travelled by train, car and bicycle, and which factors of the built environment and related policies, consistently or through their change, have influenced this, while controlling for the role of socio-demographic factors. In doing so, we also tested the pseudo panel approach and the best estimation model to answer this question. The study area is the Randstad, the core region of the Netherlands. This region is interesting, as its developments are partly consistent with that of other urban regions in the world, but at the same time it was subject to increasingly stringent policy objectives (see, e.g., Kasraian et al., 2017, for a long-term overview of the land use, transport infrastructure and spatial policy developments in the Randstad). This study is unique as it models a relatively long time period, namely three decades, from 1980 to 2010. In addition to most previous studies, the focus is not only on a single mode, but on train, car and bicycle travel. For this, descriptive and pseudo panel analyses are applied to a series of repeated annual surveys over 30 years, including socio-demographic, residential location and travel behaviour indicators. This study evaluates three frequently used pseudo panel estimation techniques, i.e., the pooled ordinary least squares, fixed effects and random effects. Furthermore, a hybrid estimation is applied and its performance as the best-fitting model is investigated. The results of this study provide policymakers with understanding of the long-term effects of infrastructure investments, urban growth and mitigating policies on travel behaviour.

The next section provides a brief overview of the investigated data and its preparation. We then summarise and compare the estimation techniques applied to pseudo panels and elaborate on the new hybrid method. Subsequently we compare the results of various estimation techniques and the difference between the three modes. The paper ends with reflections on the findings and recommendations for future policy and research.

2. Data

A long-term geo-referenced database was constructed by bringing together various sources. The surveys reported the origin and destination of trips at the municipal level until 2004 and afterwards at the much more detailed level of 4-digit postal codes. Spatial, socio-demographic and travel behaviour data with varying measurements were recoded and converted to the municipal borders of year 2004 (proportional to the area of each year's spatial unit existing within the 2004 municipal border) to generate a consistent dataset for seven time points: 1980, 1985, 1990, 1995, 2000, 2005 and 2010.

Travel behaviour variables were extracted from the Dutch National Travel Surveys (OVG, MON and OViN) which provide reliable travel diary data since 1979 on an annual basis (Statistics Netherlands (1979-2004); Ministry of Infrastructure and Water Management [Rijkswaterstaat, Dienst Verkeer en Scheepvaart] (2011); Statistics Netherlands (2010)). The sample was limited to the Randstad population. As sample sizes varied, they were made comparable between the time points by adding respondents from the previous and proceeding year (e.g. 1984 and 1986 were added to 1985). Respondents younger than 20 years were excluded because of their constrained travel behaviour). We estimated single-mode models for travel by train, car and bicycle. Only respondents who reported at least one trip by train, car or bicycle during the survey day, were included. The sample size per year resulted in 11,066 (1980), 13,348 (1985), 15,107 (1990), 32,596 (1995), 29,007 (2000), 32,858 (2005) and 12,690 (2010) cases.

Table 1 provides an overview of variables used in the analysis, their definitions and sources. Average daily travel kilometres were split into train, car passenger or car driver, and bicycle; multi-modal trips were recoded to the transport mode with the longest leg (in kilometres) of the trip; other modes (e.g. motorcycles, tram, bus, metro) were excluded regarding their smaller share in total distance travelled. Walking trips were left out as they artificially increased over time due to improvements in their measurement in the more recent survey waves.

Socio-demographics are age, gender, educational level, personal income and household car ownership. Residential municipalities were categorised by "daily urban systems", according to Van der Laan (1998) used in several other studies (Schwanen et al., 2001; Van Eck and Snellen, 2006). Though the Randstad and its borders have evolved, its daily urban systems have been relatively stable over time. The three categories are "urban centres", "suburbs" (including medium sized cities in the vicinity of the urban centres) and "rural". Accessibility was

Table 1
Overview of variables used in the analysis, their definition and sources.

Variable	Definition	Source
Travel behaviour	Average daily distance travelled by train [km] Average daily distance travelled by car [km] Average daily distance travelled by bicycle [km]	National Travel Survey: OVG for 1979–2000, MON for 2004–2009, OViN for 2010
Socio-demographic		
Gender		For all socio-demographic variables, National Travel Survey: OVG for 1980–2000, MON for 2005, OViN for 2010
Age		
Household size	Number of household members	
Income	Lower, around or higher than modal income	
Education	Low, medium or highly educated	
Car availability	The individual has a driving licence and there is at least one car in the household	
Employment	Employed (full-time or part-time) or not	
Land use		
Population density of the built-up area	Number of inhabitants in the individual's home municipality per square kilometre built-up area. Built-up area is defined as the physical space used for urban functions, including real estate for housing, services, companies, infrastructure and parks.	Historical land use in the Netherlands (Historisch Grondgebruik Nederland, HGN) 1980–1990 from Alterra, Wageningen University; Adjusted Land use (BBG Mutatiereeks) for 2000–2010 from Statistics Netherlands (CBS) and Kadaster; Population at each time point from CBS
Location within the Randstad	Belonging to "urban centres", "suburbs" or "rural"	Based on Van der Laan (1998)
Accessibility		
Distance from rail	Distance from the municipality's mean centre to the closest rail station [km]	Own calculation using ArcGIS
Distance from motorway	Distance from the municipality's mean centre to the closest motorway exit [km]	Own calculation using ArcGIS

measured as the Euclidian distance from the municipality's mean centre, i.e., its centroid based on the dispersion of built-up area across the municipality, to the closest railway station and motorway exit at each time point.

3. Methodology

3.1. Pseudo panel analysis

Pseudo panel analysis was initially introduced for analysing repeated cross-sectional data in consumer economics (Deaton, 1985). Its few applications in travel behaviour studies have been to model car ownership, public transport and induced travel demand. Dargay and Vythoulkas (1999) investigated car ownership in the UK from 1970 to 1994. Tsai and Mulley (2014) modelled short-run and long-run public transport demand in Sydney, Australia over 1982–1995. Weis and Axhausen (2009) investigated the induced travel demand during 1974–2005 for Switzerland.

The pseudo panel approach relies on the construction of homogenous groups of respondents based on shared characteristics which are associated with the dependent variable and which are time-invariant (do not change over time). Each group has a number of cohorts, which equals the number of survey waves for that group. Younger groups participated in less survey waves, thus have less cohorts. Defined groups are treated as panel units and cohort means are treated as observations in a panel. Cohort means are then traced over time just like observations of an individual in a genuine panel data. The point of departure of pseudo panel analysis is a simple genuine panel data equation:

$$y_{it} = \beta_0 + \beta_1 x_{it} + \mu_i + \varepsilon_{it} \quad (1)$$

where i denotes the panel units, namely individuals, t denotes the time period, μ_i is the individual-specific error term or unobserved individual effect, and $\varepsilon_{it}$is the independent error term. The dependent variable y represents the average daily kilometres travelled by train, car or bicycle. The predictors include a number of socio-demographic and location variables (Table 1). When applied to a pseudo panel, the equation transforms into:

$$\bar{y}_{gt} = \beta_0 + \beta_1 \bar{x}_{gt} + \bar{\mu}_{gt} + \bar{\varepsilon}_{gt} \quad (2)$$

Here the units of analysis are no longer individuals, but homogenous groups of individuals, denoted by subscript g instead of i. Respectively, dependent and independent variables are transformed into $\bar{y}_{gt}$ and $\bar{x}_{gt}$ which are the mean values within each group g at time t and calculated as: $\bar{y}_{gt} = \frac{1}{k}\sum_{i=1}^{k} y_{it}$ and $\bar{x}_{gt} = \frac{1}{k}\sum_{i=1}^{k} x_{it}$, where k is the number of individuals in a group's cohort. The term $\bar{\mu}_{gt}$ specifies the average group-specific error or the average unobserved group effect which, unlike unobserved individual effect μ_i, varies over time. The reason is that the cohorts of a certain group include different members at various time points. However, the unobserved group effect is considered time-invariant and groups are treated as panel units as cohorts are large enough (including at least 100 members as a rule of thumb), and there is sufficient between-group variation and within-group homogeneity (Deaton, 1985; Verbeek and Nijman, 1992).

The chosen grouping criteria are birth year and level of education. There is a trade-off between cohort size and the number of groups as an increase in one results in a decrease in the other (Tsai et al., 2013). Thresholds for subcategories within each criterion and respectively the number of groups were chosen regarding consistency over time (the common denominators of varying measures at different survey waves) and to maximise the number of members per all cohorts. Table 2 shows the constructed groups, their average cohort size and the number of cohorts. All dependent variables and the majority of the independent ones have higher between group than within group standard deviations. Thus the constructed pseudo panel satisfies the sufficient between-group variation condition.

Birth year was not registered as a continuous variable for several time points and respondents' age was collected within varying age bands over the survey waves. Due to different age classifications, respondents with the age of 20 or more had to be chosen to achieve consistent groups over time. This threshold plus broad classifications for education groups reduce the likelihood of change within education categories over the study period.

The final pseudo panel dataset consists of twenty-one groups, based on a 7 by 3 classification of birth year and education. The number of cohorts for every group equals the number of survey waves where that group is present. Eleven of the twenty-one groups have observations for

Table 2
Overview of the constructed groups based on birth year and education, their average cohort size and number of cohorts.

Group	Grouping Criteria		Average cohort size	No. of cohorts
	Birth year	Education		
1	≤1925	Low	585	7
2	1926–1940	Low	633	7
3	1941–1950	Low	352	7
4	1951–1960	Low	221	7
5	1961–1970	Low	125	5
6	1971–1980	Low	78	3
7	1981–1990	Low	13	1
8	≤1925	Medium	1059	7
9	1926–1940	Medium	2308	7
10	1941–1950	Medium	2846	7
11	1951–1960	Medium	3227	7
12	1961–1970	Medium	3592	5
13	1971–1980	Medium	2396	3
14	1981–1990	Medium	519	1
15	≤1925	High	273	7
16	1926–1940	High	646	7
17	1941–1950	High	1108	7
18	1951–1960	High	1482	7
19	1961–1970	High	1829	5
20	1971–1980	High	1788	3
21	1981–1990	High	345	1
			Total	111

Final number of cohorts: 103, 8 cohorts < 100 members.

the whole study period. Thus, each of these groups has seven cohorts corresponding to 1980, 1985, 1990, 1995, 2000, 2005 and 2010 survey waves. However, groups based on the more recent birth years have increasingly fewer observations (cohorts), as their members were not yet born or were too young to qualify for the threshold age of 20 years at earlier survey waves. Eight cohorts containing < 100 members were excluded from the panel dataset to minimise the measurement error, leaving a total of 103 cohorts, which equals the total number of observations in the pseudo panel dataset.

3.2. Estimation techniques

Pseudo panel data can be investigated with the same variety of estimation techniques as genuine panel data including pooled ordinary least squares, fixed and random effects and instrumental variable estimators (Tsai and Mulley, 2014). The choice of the estimation technique is not straight forward as panel data estimators perform differently based on the characteristics of the dataset (Tsai et al., 2014). For instance, depending on the ratio of within- to between-group variation or the presence of unobserved group effect, some estimation techniques might produce biased results (see below).Furthermore, there is a hierarchical structure of the pseudo panel data constructed from repeated cross-sectional observations. That is, the inherent temporal hierarchy is due to every group having several observations (cohort means) over the study period. In other words, observations are nested within groups. When temporal hierarchies exist, observations within the same group are very likely related to each other over time (Bell and Jones, 2015). The remainder of this section summarises widely used pseudo panel estimation techniques for continuous response variables, and introduces a hybrid model which to the best of our knowledge is not applied to pseudo panel studies so far. These techniques and their characteristics are summarised in Table 3.

3.2.1. Pooled ordinary least squares

Pooled ordinary least squares technique (POLS) has been recommended for pseudo panel datasets with larger between-group than within-group variation (Tsai et al., 2013). This technique *pools* data from all survey waves and uses all the available variation in the data, including variation between waves and groups. In the presence of unobserved group effect, this estimator will be biased and inconsistent (Tsai et al., 2013). As all observations are treated the same, the geographic and/or temporal hierarchy between the higher levels (e.g. countries, groups) and lower levels (e.g. states, observations) are overlooked.

3.2.2. Fixed effects

Fixed effects (FE) models have been generally used for estimating static pseudo panels (Tsai et al., 2014). This specification involves demeaning Eq. 2, i.e., subtracting the time-mean of each unit $\bar{y}_g$ (where a unit is a group g in pseudo panel analysis) from each observation of that unit (y_{gt}), which results in the omission of unobserved group effect. FE is a *within* estimator as it measures the deviation from the mean *within* each group over time. There are two caveats regarding its usage in pseudo panel analysis. First, the unobserved group effect varies over time – as opposed to unobserved individual effect which is fixed. So the rational of FE which controls for group effect by “cancelling it out” would be problematic in pseudo panel analysis (Tsai et al., 2014). That is why this model is used with large cohort sizes with at least 100 members in order to ignore the measurement error. The second problem is that even if the measurement error is minimised by using large cohort sizes, FE focuses on the relation between outcome and predictors within an entity (group in this case) and disregards the differences between entities. Thus FE will be inefficient for panel data with much larger between-entity variation than within-entity variation (Allison, 2009). This condition applies to pseudo panel groups which are constructed with the aim to maximise between-group variance as opposed to within-group variance.

3.2.3. Random effects

Random effects models (RE) have also been applied to pseudo panel analysis (Dargay and Vythoulkas, 1999). This model assumes exogeneity between the regressors and the unobserved group effect. It is estimated by generalised least squares (GLS) or maximum likelihood (ML) and takes into account both between-group and within-group variations. RE takes account of the data's hierarchical structure with its multi-level structure, as it partitions the unexplained residual variance into higher and lower levels. In this case, the lower level consists of observations which are nested in the higher level of pseudo panel groups. However, if the exogeneity assumption fails to hold, RE estimates will be biased. A conventional criteria for deciding between RE and FE is the result of the Hausman test. It examines the orthogonality of unobserved group effect from other regressors and compares the efficiency of FE and RE.

3.2.4. Hybrid method

Another method – which to the best of our knowledge has not yet been applied to pseudo panel analysis in travel studies – is a hybrid estimation based on a reformulation of Mundlak (1978). This specification is especially of importance for repeated cross-sectional data with temporal hierarchies (Bell and Jones, 2015). It includes, like FE, the independent variables with a time-demeaned transformation, i.e., in the form of deviations from the group's time-mean ($x_{gt} - \bar{x}_g$). Thus it measures the within-group, or *time-series* variation. Furthermore, the means of higher entities over time are also included as explanatory variables. In this case, these are group-specific means ($\bar{x}_g$), which are time invariant components of the independent variables as they are averages over time. The coefficients of these time-mean variables will show the between-group, or *cross-section* variation. Finally, the model with independent variables transformed as above is estimated with an RE model which takes the hierarchical structure into account.

The hybrid technique specifically models the unobserved group effect, making Hausman test obsolete (Bell and Jones, 2015). Furthermore, the within and between effects are both measured and clearly separated, making interpretations easy. Finally the temporal hierarchy

Table 3
Most used pseudo panel estimation techniques and their characteristics.

Model name and specification	Controls for unobserved group effect	Multi- level structure	Characteristics	Estimation technique
Pooled ordinary least squares (POLS) $y_g = \beta_0 + \beta_1 x_g + \varepsilon_g$	No	No	Pools all observations; is biased in the presence of unobserved group effect	OLS
Fixed effects (FE) $y_{gt} - \bar{y}_g = \beta_1(x_{gt} - \bar{x}_g) + (\varepsilon_{gt} - \bar{\varepsilon}_g)$	By cancelling it out	No	Unable to model time-invariant variables or between-unit effects	OLS
Standard random effects (RE) i. random intercept: $y_{gt} = \beta_0 + \beta_1 x_{gt} + \mu_g + \varepsilon_{gt}$ ii. random intercept and slope: $y_{gt} = \beta_0 + \beta_{1j} x_{gt} + \mu_g + \varepsilon_{gt}$	By including it in the model (assumed exogenous)	Yes	Can model time invariant variables and between-unit effects; biased if assumption of exogeneity is not met; assumes a random intercept which can be estimated with a fixed or random slope	GLS; ML
Hybrid $y_{gt} = \beta_0 + \beta_1(x_{gt} - \bar{x}_g) + \beta_2 \bar{x}_g + \mu_g + \varepsilon_{gt}$	By including it in the model (assumed exogenous) plus a higher level mean for time-varying covariates	Yes	As above, the addition of higher level mean ensures that the effect estimates of higher level variables are corrected for between cluster differences of x_{it}; β1 and β2 are easy to interpret as they explicitly show within and between group variations	GLS; ML

Note: To indicate that our units of analysis are groups rather than individuals, we use g in the notations rather than an i.

of the data is taken into account as it is a multi-level model. This specification has been suggested by others as a "hybrid" method, i.e., a compromise between FE and RE (Allison, 2009; Schmidt, 2012). Bell and Jones (2015) correctly contend that this model is in fact an RE model. However, in order to distinguish between this method and the standard RE and to use the terminology of pervious works, this paper also adopts the term "hybrid".

4. Results

4.1. Descriptive analysis at the traveller level

Fig. 1 summarises long-term trends in daily distances *per traveller* (i.e., a respondent who has reported at least one trip by either train, car or bicycle), by separate transport modes and the three together, from 1980 to 2010. As expected, the majority of the kilometres are travelled by car, as the share of car trips is significantly higher than that of the train and bicycle. The *average* daily distance travelled by all three modes (the purple line) has increased since 1980, reaching a peak at 1995, and has been more or less decreasing ever since. However, there is a caveat: the average is influenced by outliers (e.g. a limited number of respondents who travel very long distances). The *median* daily kilometres travelled (the dashed purple line) is much lower than average, revealing that half of the travellers living in the Randstad have been travelling below 30 km per day over three decades. As many respondents did not report a bicycle or a train trip on the survey day, their medians are zero. A remarkable trend is the decline in kilometres travelled by car after mid-1990s, which is more significant in the Randstad in comparison with the rest of the Netherlands. We believe this could be attributed to the 'Randstad effect', which is a combination of a number of interrelated factors: spatial and transport policies which have encouraged compact urban developments and improved the competitiveness of public transport, the saturation of accessibility by car and the growing disutility of using cars in the Randstad regarding the high costs of parking in the big cities. The following sections will summarise the above trends at the group level and model them.

4.2. Comparison of estimation techniques

To select the optimal model, the four estimation techniques are applied to average daily train kilometres travelled. Table 4 compares the standardised coefficient estimates achieved by pooled ordinary least squares (POLS), fixed and random effects (FE and RE) and hybrid techniques. The dependent variable is logarithm transformed to minimise the effect of its positive skewness and heteroscedasticity. Controlling for multicollinearity determined the final choice of independent variables. These variables include the groups' share of (1) different age bands, (2) females, (3) medium and high-income earners, and (4) people living in the suburbs and rural areas, as well as the groups' average of (5) members' household size, and (6) living municipalities' distance to the closest motorway exit. Time-demeaned variables are indicated with (*d*). For instance, *% Female* (*d*), is the deviation of the

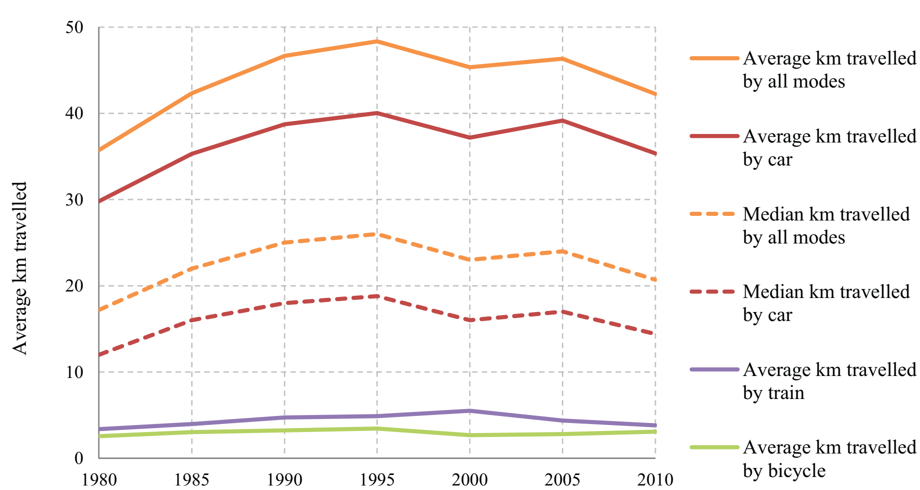

Fig. 1. Daily distance travelled by all and separate modes for travellers living in the Randstad from 1980 to 2010.

Table 4
Model estimates based on pooled ordinary least squares, fixed effects, random effects and hybrid techniques for the natural logarithm of average daily distance travelled by train [km].

Independent variables	(1) Pooled OLS	(2) Fixed effects	(3) Random effects	(4) Hybrid
Age				
% 30–39	−0.293*		−0.286**	
% 40–49	−0.193		−0.176	
% 50–64	−0.277**		−0.213**	
% 65 or older	−0.175		−0.157	
% Female	0.238*		0.149	
Household size	−0.175		−0.105	
% Medium income	0.191		0.120	
% High income	0.756***		0.631***	
% Living in suburb	−0.221**		−0.282***	
% Living in rural area	0.057		0.050	
Dist. to motorway exit	−0.320**		−0.261	
Age (d)				
% 30–39		−0.071		0.069
% 40–49		0.150		0.338*
% 50–64		0.396		0.566*
% 65 or older		0.548		0.574*
% Female (d)		−0.005		0.004
Household size (d)		0.133		−0.009
% Medium income (d)		−0.198		−0.193
% High income (d)		−0.166		−0.086
% Living in suburb (d)		−0.329**		−0.165**
% Living in rural area (d)		0.141		0.080
Dist. to motorway exit		−0.263		−0.338
Age (m)				
% 30–39				−0.112
% 40–49				−0.049
% 50–64				−0.436**
% 65 or older				−0.802
% Female (m)				0.051
Household size (m)				−0.794
% Medium income (m)				−0.105
% High income (m)				0.236
% Living in suburb (m)				0.116
% Living in rural area (m)				0.217
Dist. to motorway exit				−0.460**
Year				
1985	−0.645	−0.313	−0.548	−0.471
1990	−0.832**	−0.623	−0.660	−1.231*
1995	−0.853**	−0.654	−0.661	−1.326*
2000	−0.957*	−1.141	−0.811	−2.027**
2005	−1.264**	−1.399	−1.069	−2.420**
2010	−1.815***	−1.602*	−1.454**	−2.797**
Constant	2.132***	1.015	1.971***	2.757***
Group fixed effects	No	yes	no	no
R^2				
Within		0.466	0.373	0.453
Between		0.113	0.722	0.891
Overall	0.602	0.808	0.590	0.739

$N = 103$; Standardised coefficients (robust standard errors in parentheses).
All independent variables except for year dummies are standardised.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.001$.

proportion of females in the group from the group-specific mean over time ($x_{gt} - \bar{x}_g$). The group-specific mean over time ($\bar{x}_g$) itself is indicated with the same variable name marked with (*m*). The results of various tests such as Breusch and Pagan Lagrangian multiplier test for random effects, Ramsey's RESET test for omitted variables and Hausman test were compared to identify model misspecification.

The first column shows the result of POLS. The signs and significance of the variables are as expected, however, the Breusch and Pagan Lagrangian multiplier indicates the presence of random effects, which means that POLS estimates are biased. The FE model in the second column includes group fixed effects, i.e., dummies for all groups (minus one) which absorb the unobserved group effect. The model fit in terms of R^2 is the highest, however FE fails to explain the pseudo panel's high between variance.

The last column shows the result of the hybrid specification. The *within* estimator part of this model $\beta_1(x_{gt} - \bar{x}_g)$ explains the variance between waves, i.e. the development of average daily train kilometres travelled over time. In other words, it measures the effect of a change in explanatory variables (marked with *d*), on a change in average daily train kilometres travelled, within each group over the study period. A rise in the share of people above the age of 40 is shown to cause a rise in the amount of average daily train kilometres travelled over time. Another significant coefficient is the percentage of members living in the suburbs within a group, implying that an increase in the share of suburban inhabitants over time will result in lower train kilometres travelled. In theory, the coefficients achieved with the within estimator of the hybrid model and those of the FE model should be identical. However, this is not the case as the time-mean of every variable – e.g. year – is *not* included in the model (Schmidt, 2012) and there is variation in the number of cohorts per group (Allison, 2009).

The *between* estimator part of the model, $\beta_2\overline{x}_i$, explains the variance at the group level. In other words, it measures differences between groups at the cross section level. The coefficients measure the explanatory power of time-invariant variables (marked with an *m*), in this case the time-means of higher-entities which are the groups. Here, two variables show a negative effect: the groups' share of 50–64 year olds and the average distance to the motorway exit. This means that, considering the whole study period as a cross section, groups with a higher share of people in their fifties and early sixties or a high average distance to motorway exits are associated with lower daily train kilometres travelled.

The hybrid model measures the within and between variations in a manner which is easy to interpret, in line with our aim to distinguish between determinants' constant and varying influences over the study period. Unlike FE, this model captures the high between group variation existing in our dataset and outperforms the RE in terms of R^2. Thus it is the most suitable model for our dataset. Although not reported, the four models for bicycle and car kilometres yielded the same results. The comparison of the hybrid specification between the modes is reported in the next section.

4.3. Distances travelled by train, car and bicycle and their determinants over time

With the hybrid method selected as the best fit for our research questions and database, this section presents hybrid models for all modes, i.e., for the natural logarithm of average daily distances travelled by train, car and bicycle (Table 5). The train model shows that considering the within-group change over time (*d* variables), an increase in the share of people above 40 is related to higher daily train kilometres travelled. On the contrary, a rise in groups' share of suburban inhabitants –at the cost of leaving the urban core– is associated with lower train kilometres travelled. As for between-group differences (*m* variables), groups with higher share of 50–64 year olds and those with higher distance to motorway exit have lower daily train kilometres. It seems counterintuitive that living further away from motorways results in lower train kilometres, however this is most likely because many motorways were constructed parallel to railways in an attempt to bundle transport infrastructures and the two networks are highly correlated.

The different signs for the share of 50–64 year olds, measured by the within and between estimators of the hybrid model might seem contradictory. However, it should be kept in mind that the within estimator measures the variables' deviation from the average while the between estimator measures their average over time. So, while an increase in the

Table 5
Comparison of the result of the hybrid model for the natural logarithm of average daily distance travelled by train, car and bicycle [km].

Independent variables	(1) Train		(2) Car		(3) Bicycle	
Age (d)						
% 30–39	0.069	(0.118)	0.104**	(0.038)	0.095**	(0.042)
% 40–49	0.338*	(0.184)	0.163**	(0.067)	0.197**	(0.074)
% 50–64	0.566*	(0.298)	0.244**	(0.111)	0.280**	(0.105)
% 65 or older	0.574*	(0.329)	0.149	(0.121)	0.298**	(0.112)
% Female (d)	0.004	(0.047)	0.002	(0.013)	−0.021	(0.016)
Household size (d)	−0.009	(0.163)	0.010	(0.062)	−0.067	(0.051)
% Medium income (d)	−0.193	(0.177)	0.040	(0.034)	−0.035	(0.024)
% High income (d)	−0.086	(0.123)	−0.012	(0.030)	−0.031	(0.037)
% Living in suburb (d)	−0.165**	(0.054)	−0.003	(0.018)	0.038*	(0.020)
% Living in rural area (d)	0.080	(0.136)	0.037**	(0.017)	−0.011	(0.018)
Dist. to motorway exit (d)	−0.338	(0.235)	−0,011	(0.036)	−0.037	(0.037)
Age (m)						
% 30–39	−0.112	(0.140)	−0.021	(0.024)	−0.060*	(0.033)
% 40–49	−0.049	(0.126)	0.010	(0.025)	0.024	(0.046)
% 50–64	−0.436**	(0.148)	−0.042**	(0.018)	0.008	(0.029)
% 65 or older	−0.802	(0.714)	−0.303***	(0.084)	−0.267	(0.174)
% Female (m)	0.051	(0.146)	0.017	(0.020)	0.021	(0.053)
Household size (m)	−0.794	(0.589)	−0.203**	(0.069)	−0.170	(0.193)
% Medium income (m)	−0.105	(0.123)	0.020	(0.016)	−0.063**	(0.029)
% High income (m)	0.236	(0.276)	0.192***	(0.041)	−0.097	(0.084)
% Living in suburb (m)	0.116	(0.197)	0.046**	(0.020)	−0.051	(0.058)
% Living in rural area (m)	0.217	(0.230)	0.046	(0.037)	0.062	(0.059)
Dist. to motorway exit (m)	−0.460**	(0.203)	−0.021	(0.033)	−0.152**	(0.055)
Year						
1985	−0.471	(0.462)	0.130**	(0.062)	0.145**	(0.073)
1990	−1.231*	(0.693)	0.117	(0.123)	−0.184	(0.148)
1995	−1.326*	(0.771)	0.124	(0.133)	−0.108	(0.175)
2000	−2.027**	(0.918)	−0.081	(0.184)	−0.675***	(0.199)
2005	−2.420**	(1.046)	−0.080	(0.237)	−0.786**	(0.271)
2010	−2.797**	(1.155)	−0.398	(0.278)	−0.791**	(0.318)
Constant	2.757***	(0.722)	3.547***	(0.145)	1.459***	(0.184)
R^2						
Within	0.453		0.765		0.648	
Between	0.891		0.987		0.826	
Overall	0.739		0.908		0.714	

N = 103; Standardised coefficients (robust standard errors in parentheses).
All independent variables except for year dummies are standardised.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.001$.

share of the 50–64 year olds over time will result in higher train kilometres, in general higher shares of people in their fifties and early sixties is associated with lower train kilometres travelled.

The second model indicates that an increase in age until 65, as well as an increase in the share of people living in rural areas are associated with higher car kilometres travelled over the study period. While in general, groups with a higher share of people above 50 (especially above 65), and higher average household sizes, travel lower distances by car per day. As expected, groups with a higher share of high-income people and suburban inhabitants also have higher car kilometres travelled. The third model shows that an increase in age within the group results in progressively higher average daily bicycle kilometres travelled over the study period. This implies that people are increasingly biking more as they become older (especially above 50). Furthermore, a rise in suburban inhabitants is linked to higher bicycle kilometres travelled. On the other hand, groups with a higher share of people in their 30s, medium-income earners and with higher average distance to motorway exits, have lower bicycle kilometres. Similar to the train model, distance from motorway exits is a proxy for transport infrastructure availability and eventually how isolated an area is (the Randstad's outer ring and Green Heart in this case). Here opportunities are less concentrated and spread over longer distances thus are not feasible for bicycle travel.

Trends in the time fixed effects (year dummies) show that controlling for sociodemographic and location characteristics, average daily kilometres by bicycle and to a higher degree by train have been increasingly decreasing with more recent survey waves compared to 1980. There is a similar decrease in the car kilometres travelled, however, the results are not significant.

5. Conclusions and discussion

This paper investigated the trends in average daily distance travelled and its determinants, for the three modes of train, car and bicycle, from 1980 to 2010 in the Dutch polycentric metropolitan region of the Randstad, using descriptive and pseudo panel analysis with various econometric models. Its aim was to find out how travel behaviour has developed over the long term in the Randstad and to pinpoint the impact of the changing built-environment on the development of travel behaviour, while controlling for the role of socio-demographics. Its specific contribution to the field is threefold. First, it investigated the development of travel behaviour over a very long time period. Second it compared average daily distance travelled between three transport modes (train, car and bicycle) and their determinants over time. Finally it applied a new hybrid method to the pseudo panel analysis in travel studies.

Analysis at the traveller level revealed that median daily distance travelled in the Randstad has remained under 26 km, even at its highest point in 1995, has dropped to 20 in 2010 for all modes, and for cars even to 15. In other words, half of the Randstad inhabitants have been travelling ≤ 26 km a day over the past thirty years (and this is excluding those who reported no trip during the survey day). This makes alternative transport modes with relatively limited range such as electric cars and bicycles suitable for the Dutch context and especially the Randstad. With the emerging electric bicycles, more than half of the car trips can be easily replaced by cycling (Lee et al., 2015).

Regarding the effect of change (of determinants) on change (of travel behaviour) at the group level, a rise in age is increasingly related to higher train and bicycle kilometres travelled. In other words, as people grow older, they increasingly travel more by train and bicycle, possibly as they consider it more important to remain fit (Fishman et al., 2015) and have more time compared to busy young adults and families with children. This calls for an increase in train and bicycle facilities for the elderly, in particular safer bicycle infrastructure, as older electric cyclists appear to be highly vulnerable (SWOV factsheet, 2013).

The role of the built environment is measured by the impact of residential location (urban, suburban or rural) and distance to motorway exits. The findings show that an increase in the inhabitants of suburban areas –at the cost of leaving the urban core– is associated with significantly lower distances travelled by train. This is most likely due to a lower supply and service quality of suburban rail infrastructure in comparison with the central cities (Limtanakool et al., 2006). However, a higher share of suburbanites results in higher bicycle kilometres travelled, which could be for the reason that various land uses are further apart in less dense suburban environment compared to urban cores, which necessitates travelling longer distances to perform activities. Moreover, it is easier to bike in a suburban environment than in congested big cities, and furthermore bicycles have more competition from urban public transport in the bigger cities (Bamberg et al., 2003). A rise in the share of rural inhabitants is associated with higher car kilometres. This was expected as the "rural" category includes municipalities in the Green Heart and the Randstad's outer ring which mostly lack an efficient and well-connected rail infrastructure, have less disincentives for car use and more space for parking (Kasraian et al., 2016a). Areas on higher distances from motorway exits are most likely in the Green Heart and outer rings, which also lack proper train connections, thus train and also bicycle kilometres travelled show lower values.

Pseudo panel analysis provides the opportunity to analyse travel behaviour over long time spans at a relatively disaggregated (group) level while using available travel survey datasets. There are, however, a number of caveats to be considered. First, it is not possible to control for the issue of residential self-selection with this method. Second, the extent to which the grouping criteria for the construction of pseudo panel groups and the characteristics of the constructed groups bias the results, is not clear and calls for further investigations, such as the work of Tsai et al. (2013). Third, the findings are based on groups of individuals. While we have tried to minimise the loss of information on the variation of the observations within groups by ensuring that the within-group variation is small compared to between-group variation, it is not known whether and to what degree the results differ from an individual level analysis. Despite these caveats, pseudo panel analysis contributes to investigating changes in travel behaviour and its determinants over time spans and at a level of aggregation which are otherwise not possible in the absence of genuine panel data. The hybrid technique applied in this paper is shown to be the best estimation for analysing our pseudo panel dataset, as it explicitly models unobserved group effect, takes account of the temporal hierarchy of repeated cross-sectional data, and measures both within- and between-group variation in a manner which is easy to interpret and outperforms the RE estimate. Furthermore, it can include time-invariant variables if needed.

This study used a static specification, future research could identify short- and long-term rate of change in travel behaviour in response to land use and transport infrastructure, e.g. by incorporating dynamic partially adjusted models (Tsai et al., 2014). Future research could test and compare more complicated accessibility measures such as potential accessibility indicators to gain a more in-depth understanding of the (changing) role of access to transport in travel demand. Furthermore, the model fit in terms of R^2 was shown to be highest for car, compared to train and bicycle. This suggests that the chosen variables can explain the variation in distance travelled by car more than train and bicycle. It is important to identify and estimate the role of other train and bicycle related factors such as train fare prices, quality and supply of bicycle and train infrastructure or attitudes towards their use, which could provide more insight into how their travel behaviour varies from car. Finally, this analysis measured the impact of land use on travel behaviour. Nevertheless, it is very likely that a reverse relation exists where travel behaviour affects land use over the long time span of the study. Further research should include the possibility of reverse causality and measure its impact as well.

Declaration of interests

None.

Funding statement

This work was supported by the Netherlands Organisation for Scientific Research (NWO), under grant 438-12-458 (CASUAL: Co-creating Attractive and Sustainable Urban Areas and Lifestyles –Exploring new forms of inclusive urban governance), funded under the Urban Europe Joint Programming Initiative.

Acknowledgements

We are thankful to Frans Rip from Alterra for providing us with historical land use maps (Historisch Grondgebruik Nederland, HGN), Hans van Amsterdam from the Environmental Assessment agency (PBL) for his help by sharing data and thoughts especially on the Dutch historical road network and municipal populations, Kees van Goeverden for providing us with the reworked Dutch National Travel Survey database and detailed insights into it, and Jaap Nieuwenhuis for his assistance with the hybrid methodology, the latter two both from Delft University of Technology.

References

Allison, P.D., 2009. Fixed Effects Regression Models. The United States of America. Sage Publications, Inc.
Bamberg, S., Ajzen, I., Schmidt, P., 2003. Choice of travel mode in the theory of planned behavior: the roles of past behavior, habit, and reasoned action. Basic Appl. Soc. Psychol. 25 (3), 175–187.
Baum-Snow, N., Kahn, M.E., 2000. The effects of new public projects to expand urban rail transit. J. Public Econ. 77 (2), 241–263. https://doi.org/10.1016/S0047-2727(99)00085-7.
Bell, A., Jones, K., 2015. Explaining Fixed Effects: Random Effects Modeling of Time-Series Cross-Sectional and Panel Data. Political Sci. Res. Methods 3, 133–153. https://doi.org/10.1017/psrm.2014.7.
Bertolini, L., 2012. Integrating mobility and urban development agendas: a manifesto. disP-The Planning Review 48 (1), 16–26.
Dargay, J.M., 2002. Determinants of car ownership in rural and urban areas: a pseudo-panel analysis. Transport. Res. Part E-Logistics Transport. Rev. 38 (5), 351–366. https://doi.org/10.1016/S1366-5545(01)00019-9.
Dargay, J.M., Vythoulkas, P.C., 1999. Estimation of a dynamic car ownership model - a pseudo-panel approach. J. Transport Econ. Policy 33, 287–301.
Deaton, A., 1985. Panel Data from Time-Series of Cross-Sections. J. Econ. 30 (1–2), 109–126. https://doi.org/10.1016/0304-4076(85)90134-4.
Ellder, E., 2014. Commuting choices and residential built environments in Sweden, 1990-2010: a multilevel analysis. Urban Geogr. 35 (5), 715–734. https://doi.org/10.1080/02723638.2014.916906.
Feng, J., Dijst, M., Wissink, B., Prillwitz, J., 2017. Changing travel behaviour in urban China: evidence from Nanjing 2008–2011. Transp. Policy 53, 1–10.
Fishman, E., Böcker, L., Helbich, M., 2015. Adult active transport in the Netherlands: an analysis of its contribution to physical activity requirements. PLoS One 10 (4).
Giuliano, G., 2004. Land Use Impacts of Transportation Investments. The Geography of Urban Transportation. Vol. 3. pp. 237–273.
Guerra, E., 2014. The built environment and car use in Mexico City: is the relationship changing over time? J. Plan. Educ. Res. 34 (4), 394–408.
Kasraian, D., Maat, K., van Wee, B., 2016a. Development of rail infrastructure and its impact on urbanization in the Randstad, the Netherlands. J. Transport Land Use 9 (1), 151–170. https://doi.org/10.5198/jtlu.2015.665.
Kasraian, D., Maat, K., Stead, D., van Wee, B., 2016b. Long-term impacts of transport infrastructure networks on land-use change: an international review of empirical studies. Transp. Rev. 36 (6), 772–792.
Kasraian, D., Maat, K., van Wee, B., 2017. The impact of urban proximity, transport accessibility and policy on urban growth: a longitudinal analysis over five decades. Environ. Plan. B (2399808317740355).
Lee, A., Molin, E., Maat, K., Sierzchula, W., 2015. Electric bicycle use and mode choice in the Netherlands. Transp. Res. Rec. 2520, 1–7.
Limtanakool, N., Dijst, M., Schwanen, T., 2006. The influence of socioeconomic characteristics, land use and travel time considerations on mode choice for medium-and longer-distance trips. J. Transp. Geogr. 14 (5), 327–341.
Maat, K., Timmermans, H.J.P., 2009. A causal model relating urban form with daily travel distance through activity/travel decisions. Transp. Plan. Technol. 32 (2), 115–134.
McDonald, N.C., 2015. Are millennials really the "go-nowhere" generation? J. Am. Plan. Assoc. 81 (2), 90–103.
Ministry of Infrastructure and Water Management [Rijkswaterstaat, Dienst Verkeer en Scheepvaart], 2011. Mobiliteitsonderzoek Nederland (MON) 2004–2009. SPSS, Netherlands DANS EASY.
Mundlak, Y., 1978. Pooling of time-series and cross-section data. Econometrica 46 (1), 69–85 (doi: 10.2307/1913646).
Noland, R.B., Lem, L.L., 2002. A review of the evidence for induced travel and changes in transportation and environmental policy in the US and the UK. Transport. Res. Part D-Transport Environ. 7 (1), 1–26. https://doi.org/10.1016/S1361-9209(01)00009-8.
Schmidt, A., 2012. The development of public demand for redistribution. In: A Pseudo-Panel Model for Decomposing Within-and Between-Effects. GK SOCLIFE Working Paper Series.
Schwanen, T., Dieleman, F.M., Dijst, M.J., 2001. Travel behaviour in Dutch monocentric and policentric urban systems. J. Transp. Geogr. 9 (3), 173–186.
Statistics Netherlands, 1979-2004. Onderzoek Verplaatsingsgedrag (OVG). Centraal Bureau voor de Statistiek, Den Haag/Heerlen.
Statistics Netherlands, 2010. Onderzoek Verplaatsingeng in Nederland (OViN). Centraal Bureau voor de Statistiek, Den Haag/Heerlen.
Susilo, Y.O., Maat, K., 2007. The influence of built environment to the trends in commuting journeys in the Netherlands. Transportation 34 (5), 589–609. https://doi.org/10.1007/s11116-007-9129-5.
SWOV factsheet (2013). Retrieved from www.swov.nl/sites/default/files/publicaties/gearchiveerde-factsheet/nl/factsheet_fietsers_gearchiveerd.pdf (5 March 2017).
Tsai, Mulley, C., 2014. Identifying short-run and long-run public transport demand elasticities in Sydney a pseudo panel approach. J. Transport Econ. Policy 48, 241–259.
Tsai, Waiyan, L., Mulley, C., Clifton, G., 2013. Examining Estimation Bias and Efficiency for Pseudopanel Data in Travel demand Analysis. Transport. Res. Record 1–8 (doi: 10.3141/2354-01).
Tsai, Mulley, C., Clifton, G., 2014. A review of pseudo panel data approach in estimating short-run and long-run public transport demand elasticities. Transp. Rev. 34 (1), 102–121. https://doi.org/10.1080/01441647.2013.875079.
Van de Coevering, P., Maat, K., Kroesen, M., van Wee, B., 2016. Causal effects of built environment characteristics on travel behaviour: a longitudinal approach. EJTIR 16 (4), 674–697.
Van der Laan, L., 1998. Changing Urban Systems: an Empirical Analysis at two Spatial Levels. Reg. Stud. 32 (3), 235–247.
Van Eck, J.R., Snellen, D., 2006. Is the randstad a city network? evidence from commuting patterns. In: European Transport Conference, (Strasbourg, France).
Verbeek, M., Nijman, T., 1992. Can cohort data be treated as genuine panel data? Empir. Econ. 17 (1), 9–23.
Wegener, M., Fürst, F., 1999. Integration of Transport and Land Use Planning, Deliverable D2a–Land Use Transport Interaction: State of the Art. European Commission, 4ème programme cadre de recherche, Bruxelles.
Weis, C., Axhausen, K.W., 2009. Induced travel demand: evidence from a pseudo panel data based structural equations model. Res. Transp. Econ. 25, 8–18. https://doi.org/10.1016/j.retrec.2009.08.007.
Zegras, P., Hannan, Veronica, 2012. Dynamics of automobile ownership under rapid growth: case study of Santiago, Chile. Transport. Res. Record 2323, 80–89.