Transportation Research Part A 165 (2022) 225–245

Contents lists available at ScienceDirect

Transportation Research Part A

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

Deriving transport appraisal values from emerging revealed preference data

Panagiotis Tsoleridis *, Charisma F. Choudhury, Stephane Hess

University of Leeds, University road, Leeds, LS2 9JT, UK

A R T I C L E I N F O

Keywords:
Transport appraisal
Values of travel time estimates
Emerging data sources
GPS
Mixed Logit models

A B S T R A C T

Transport demand models are widely used to inform policy making and produce forecasts of future demand. A core output derived from demand models is the Value of Travel Time (VTT), which provides insights on the trade-offs that travellers are willing to make in terms of travel time and travel cost. VTT estimates are a critical input to cost-benefit analyses and feasibility assessments of potential projects and thus play a crucial role in transport planning and policy decisions. While much of the early work on VTT made use of revealed preference (RP) data, their use decreased due to growing concerns about reporting errors that may result in omitted observations and measurement errors in the model inputs. As a consequence, VTT measures have, for the last two decades, primarily been estimated using state-preference (SP) surveys. While SP methods can assess the individual trade-offs in a controlled manner, they are prone to behavioural incongruence. More recently, RP data from passively-collected data sources have raised the promise of accounting for some of the limitations of traditional RP surveys due to the minimal (or even no) active input from the respondent. The present study utilises such a dataset that combined a 2-week trip diary captured through smartphone GPS tracking with a household survey containing individual socio-demographic information. A mixed Logit model for mode choice was specified and the estimated parameters were then applied on the National Travel Survey to calculate the VTT estimates. Those estimates were further adjusted based on trip distances to get more representative national VTT values. This process resulted in estimates similar to the official UK guidelines used in transport appraisal that were obtained from SP data, where our results are not affected by concerns about response quality or survey artefacts. The findings hence strengthen the case for shifting towards passively generated RP data sources and are important for transport practitioners.

1. Introduction

Transport projects and schemes can substantially impact our day-to-day lives, as well as mid-term decisions like whether or not a buy a car or long-term decisions like where to live. They also have a profound impact on the economic growth of the country, its productivity and people's well-being. Cost-benefit analyses (CBA) and feasibility assessment of potential transport projects are primarily based on the monetary savings from travel time reductions. It is estimated that savings on travel time are responsible for around 80% of the predicted benefits of a new transport project in the UK (Mackie et al., 2001; Fosgerau and Jensen, 2003; Daly et al., 2014). The Values of Travel Time (VTT) estimates, which are used to quantify the trade-offs that decision-makers are willing to make in terms of travel time and travel cost, are hence critical components of CBA. An accurate estimation of VTTs is

* Corresponding author.
E-mail addresses: ts17pt@leeds.ac.uk (P. Tsoleridis), C.F.Choudhury@leeds.ac.uk (C.F. Choudhury), S.Hess@leeds.ac.uk (S. Hess).

https://doi.org/10.1016/j.tra.2022.08.016
Received 25 October 2021; Received in revised form 8 June 2022; Accepted 21 August 2022
Available online 28 September 2022
0965-8564/

thus important in order to properly evaluate the costs and the benefits of a new transport project and sufficiently forecast future demand for specific services, e.g. a new public transport route, leading to better informed decisions during the planning phase.

Estimates of the trade-offs that travellers would be willing to make in terms of travel time and cost were first produced in the 1960s. For a long period of time, VTT estimates were derived as relative values to the average wage cost or as a percentage of it (wage cost method or cost savings approach — CSA) and that method is still in use in several countries (Daly et al., 2014). Another approach involved Contingent Valuation, where VTTs were derived from direct questions about how much a participant would be willing to spend for a particular service or an improvement of a current one. In recent decades, most types of VTT analysis are based on the work of Daly and Zachary (1975), who first estimated VTT values from behavioural models based on theoretical frameworks of time allocation (Becker, 1965; DeSerpa, 1971) and the Random Utility framework (Marschak, 1960; McFadden, 1973; Domencich and McFadden, 1975; Williams, 1977; Ben-Akiva and Lerman, 1985; Train, 2009).

Revealed Preference (RP) data, usually coming from travel diaries, would at face value provide the natural data source for estimating VTTs, and indeed were used in early studies (Beesley, 1965; Daly and Zachary, 1975). Nonetheless, while RP data provide the ability to capture real-world choices, most of the parameters influencing them are outside of the analyst's control. Furthermore, traditional RP data sources include recalled/reported data that are prone to issues like omitted trips (particularly short ones), perception and rounding errors, etc. — often leading to large measurement errors. During the 1980s, there was also an increasing desire to capture VTTs for non-work trips – largely ignored up to that point – in addition to commuting VTTs. RP data, however, were proved to be unsuitable for providing useful real-world observed choices on non-work trips with the available data collection methods of that era. The aforementioned limitations of RP data led to the growing popularity, over the last two to three decades, of Stated Preference (SP) data as the main input to models, with RP data often being used only in limited scale for verification of the SP results (Mackie et al., 2003). SP surveys present respondents with a number of hypothetical scenarios, where they are asked to choose among a set of alternatives. This approach has a long tradition for example in the United Kingdom (UK) with the first major SP survey conducted in 1984 (MVA et al., 1987) and follow-up studies in 1994 (Accent and Hague Consulting Group, 1996) with the same data then re-analysed by Mackie et al. (2003) before the most recent study involving primary data collection taking place in 2014–2015 (Batley et al., 2019; Department for Transport, 2015; Hess et al., 2017).

SP surveys are generally seen to have the advantage of providing the analysts with an environment where they have control over a large number of parameters that could influence VTT estimates, such as the attributes of the alternatives. On the other hand, SP surveys are prone to behavioural incongruence and hypothetical bias and are often criticised for being too sensitive to the experimental design and the representation of the SP scenarios (Brownstone and Small, 2005; Daly et al., 2014; Haghani et al., 2021). Concern in a VTT context has also been raised in relation to the use of overly simplistic settings in some countries (Hess et al., 2020).

In a recent study examining the impact of hypothetical bias in SP surveys within several domains including transport (Haghani et al., 2021), the authors concluded that although it is more sensible to assume that individuals would likely overstate their Willingness-to-pay (WTP) in a hypothetical scenario (Li et al., 2020), there are a number of transport studies showing the opposite (Nielsen, 2004; Brownstone and Small, 2005; Shires and de Jong, 2009; Krcal et al., 2019). That downward bias of SP has also been proven in two meta-analyses on VTT values across countries and time of Shires and de Jong (2009) and Wardman et al. (2016).

Evidence from neuro-imaging studies also suggests that individuals would often react differently in a stressful situation compared to the lab setting of an SP survey, e.g. be willing to pay more in order to avoid an unpleasant outcome (Loewenstein, 2005; Kang and Camerer, 2013; Haghani et al., 2021). In addition to that, Kang and Camerer (2013) showed that a certain part of the brain was more strongly activated when participants had to make a real choice compared to a hypothetical one. Other psychological effects can also come into play during an SP survey influencing participants' responses, such as the desirability to appear more socially acceptable to the analyst (*social desirability bias*) (Champ and Welsh, 2006), the feeling that their choices will lack of any real-world consequences (*lack of consequentiality*) (Krcal et al., 2019), or the opposite with respondents deliberately giving misleading answers to avoid a potentially harmful outcome resulting from that study, e.g. a road pricing scheme (*strategic bias*) (Lu et al., 2008; Meginnis et al., 2018).

With that evidence in mind and considering that VTT estimates are to be used for the purpose of project evaluation during a CBA, it is only sensible to assume that policy makers would be mostly interested in the trade-offs individuals are willing to make under real-life conditions, sometimes stressful, while taking into account real distributions of travel time and cost and not the ones imposed by the analyst (Louviere and Hensher, 2001; Brownstone and Small, 2005). This thus motivates an increased interest in revealed preference (RP) data for VTT studies. For example, in a review of transport appraisal studies performed in various countries, Daly et al. (2014) concluded that despite SP data being the standard approach so far, researchers and practitioners should reconsider the use of RP data due to the benefits they can provide, while also taking advantage of the new emerging and more robust data collection methods. In addition to that, several studies using traditional RP data sources, such as national travel surveys, have showcased that VTT estimates can still be derived, which are consistent with the official SP-based values (Lorenzo Varela et al., 2018). Nonetheless, there is still a lack of similar studies utilising emerging data sources for VTT estimation purposes.

Emerging data sources, primarily from sensors, such as GPS and mobile phone data, have provided new breakthroughs and challenges to researchers. Travel diaries captured through GPS tracking are able to produce large panels of RP data per participant at a very high spatial and temporal resolution. Compared to traditional pen-and-paper diaries, GPS-based surveys offer the advantage of capturing an increased number of daily trips giving a more representative depiction of individual mobility behaviour without resulting in user fatigue. Though there have been limited efforts to infer VTTs from anonymous RP data sources (e.g Bwambale et al., 2019), the absence of socio-demographic information of the travellers and trip characteristics (e.g. trip purpose) have meant that it is not possible to capture the heterogeneity in the VTTs among different socio-demographic groups of users or due to the differences in trip purpose (e.g. commute, business, leisure) from such data.

A passively collected GPS trip diary without any additional mode or trip purpose information could require significant pre-processing efforts (Schuessler and Axhausen, 2009), which could still might not be sufficient enough to avoid biased estimates (Vij and Shankari, 2015). Contrary to that, a semi-passive GPS travel diary with minimum input from the participants and linked to a background household survey can help to account for those limitations. Several studies have used GPS datasets complimented with a background survey, but most of them have limited their analysis on descriptive statistics of individual mobility behaviour based on the observed choices (Arifin and Axhausen, 2012) or estimated models of mode choice, but without reporting VTT estimates (Montini et al., 2017; Huang et al., 2021). An exception is the study of Calastri et al. (2018), who estimated mode choice models based on GPS data for the purpose of uncovering latent mode availability and consideration constraints of the individuals during their decision-making process. Their study also reported VTTs based on the estimated parameters, however, this was purely as a means of validating their proposed approach, with no emphasis on extrapolating the findings to a representative sample, as required for official VTT values.

The focus of GPS studies so far in the literature has thus not been on the estimation of behaviourally accurate VTTs, representative of the country's population, which are derived from GPS tracking, and more importantly they have not been compared with national official estimated SP-based VTTs before. That limitation in the current literature and the lack of empirical evidence could partly explain the reluctance of policy makers to accept the use of new emerging GPS data for VTT estimation for appraisal purposes, a task that is still heavily reliant on SP surveys. Aiming to address that limitation, the current study utilises such an emerging data source, namely a 2-week GPS trip diary including 540 participants and 12 524 trips, collected as part of the European Research Council funded "DECISIONS" project, for the purpose of estimating a behavioural model of mode choice. The estimated parameters are then applied to the National Travel Survey (NTS) data and VTT estimates are derived, which are further adjusted by distance band to ensure proper representativeness of the UK's population. The main aim of the study is to compare the final distance-weighted VTT estimates with the latest official SP-based VTTs currently used in appraisal in the UK.

The remainder of the paper is as follows. In the following section, a review of the literature concerning previous VTT studies and their findings and the use of GPS data for transport-related research is performed. In the third section, the datasets used in the current study are described, while in the fourth section, the modelling framework is outlined. Following that, the modelling outputs and the derived VTT estimates are analysed in the fifth and sixth sections, respectively. A discussion regarding the policy implications of the study is performed in the following section, while in the final one the conclusions and limitations of the current study are summarised and the scope for future studies is outlined.

2. Literature review

2.1. Studies on values of travel time estimates

Originating from the studies of Becker (1965) and DeSerpa (1971), the idea of optimal time allocation and the monetisation of non-work activity participation led to the first formulations of the Value of Time. Besides the importance of VTT estimates for CBA and transport project appraisal, they also provide important insights on individual transport behaviour that can lead to better informed policy measures. The willingness-to-pay for a reduction of travel time is closely related to the overall scheduling and time allocation of the individual during the day. As a result, individuals will tend to value higher their time in contexts that would lead to more significant time restrictions (or more potential time savings) in their overall daily schedules. Such contexts can be longer distance trips, which will leave the individuals with significantly less time to accomplish other activities. That is largely empirically proven in the literature, with VTTs increasing by distance (Small, 2012). Another context can be types of activities for which individuals are required to arrive in time, such as commuting trips, or times of day which provide the individuals with greater time restrictions, such as trips during the am peak period. The aforementioned rationale is closely related to the prospect theory (Tversky and Kahneman, 1991) postulating that individuals will put a higher value to avoid a negative outcome than achieving a positive one.

According to empirical evidence from the literature, SP-based VTTs tend to be lower compared to their RP-based counterparts and that can be attributed to several biases arising in the hypothetical setting of an SP design (Shires and de Jong, 2009; Wardman et al., 2016). As an example, we refer to the context of deriving WTP for toll pricing and specifically to the studies of Vrtic et al. (2010) and Brownstone and Small (2005). Vrtic et al. (2010) using SP data found that individuals have lower VTTs in the presence of tolls, compared to untolled roads, as they are willing to take longer routes in order to bypass the tolls. That finding, however, might be subject to *strategic bias* (Lu et al., 2008; Meginnis et al., 2018) from individuals who purposefully overstate their cost sensitivities to dissuade policy makers from such a measure, which will in turn decrease their WTP. Their behaviour in reality might in fact be significantly different as stated in Brownstone and Small (2005), who examined and compared SP and RP WTP for toll pricing among a range of road corridors in the United States. The authors of that second study found that the estimates obtained from hypothetical SP surveys systematically underestimated the VTTs compared to RP-based estimates. A potential cause could be that VTT estimates will depend to a large extent on the travel time and cost range values of the SP scenarios, which could differ substantially in a real-life scenario, such as the case of severe congestion during the morning peak, forcing individuals to place higher valuations of time so as to avoid arriving late at work. That could be the case in a real-life context, in which individuals might not leave enough time to take the longer route to avoid the tolls due to other constraints (a case of *lack of consequentiality* Krcal et al., 2019). That demonstrates that the hypothetical setting of an SP survey might not be sufficient to capture real-life behaviour, which would adapt according to the time restrictions arising in the daily activity schedules, resulting in biased estimates.

Another reason for the higher RP-based VTTs, however, could also be attributed to choice set mispecification for the case of choice tasks including only one alternative actually under consideration (i.e. captive users), as mentioned in Shires and de Jong (2009). It is generally acknowledged that lack of any explicit information on availability/consideration of alternatives is an important limitation of RP datasets that has hindered their wider adoption for VTT estimation. The analyst observes only the chosen alternatives and has no control over the alternatives actually considered by the individual during her decision making process and thus included in her choice set. Li et al. (2015) have stated that this problem is not exclusive to RP, however, since choices on SP experiments can also be subject to latent choice set formation mechanisms. The presence of those types of behavioural mechanisms on SP data has been proved empirically, as well, in the study of Thiene et al. (2017) examining destination choices for recreational activities.

It is clear then that the derivation of unbiased VTT estimates relies to a large extent on the SP survey design (Bliemer and Rose, 2009). Several techniques have been developed to accomplish that, such as the inclusion of an opt-out alternative, which guarantees that respondents are not forced to choose an alternative. That, however, might also have the adverse effect of choosing the opt-out more often if the attributes of the SP scenario are not reasonable or the SP design is not meaningful to the respondents. As a result, significant effort has been put for making the SP survey more comprehensive and relatable to the respondents by including images and graphics to better describe the attributes of the choice setting. Several sources of the aforementioned biases could also be influenced by the type of SP survey, with higher chances of *social desirability bias*, for example, occurring from face-to-face interviews (Champ and Welsh, 2006).

RP data of limited scale have also been used in VTT studies to provide realistic attribute levels for the SP design — to pivot the SP values around RP values in particular. That has been established as the usual approach in SP survey design for VTT estimation, where respondents are asked to provide information of a small number of recently completed trips (Li et al., 2020). Those reference trips are then being used as anchor points to pivot the attributes of travel time and cost for the SP choice scenario around them using reasonable variations of the observed times and costs. It may be noted that even if the combination of RP and SP data is the recommended approach to get the best out of both kinds of data, RP data is still largely avoided as the main source of data input for model estimation due to reporting/measurement errors and difficulties in getting a wide range of time-cost trade offs in the real world.

In addition to the above, there is a constant debate about SP survey complexity and whether to include more attributes per choice task or to have more choice tasks with a limited number of attributes per alternative. The reason for that is to avoid causing too much cognitive fatigue to the respondent, whose choices would become more random as she becomes more fatigued (Bradley and Daly, 1994). The opposite might also be true, however, with choices becoming more deterministic as respondents learn how to respond to choice tasks as they go along (Hess et al., 2012). Instances of both fatigue and learning can occur for the same individual with evidence suggesting the presence of better quality responses at first (learning) followed by a quality decrease in further choice tasks (fatigue) (Hess et al., 2012). Furthermore, empirical evidence suggests that respondents, after a while, will tend to put more emphasis on certain attributes, such as cost, and neglect others, a behavioural process known as attribute-non-attendance (Hensher and Greene, 2010).

All of the aforementioned sources of hypothetical biases can have adverse impacts to the overall survey. It is true to say that the need to account for them has provided a strong motivation for developing state-of-the-art methodological frameworks, most of which having been implemented in the latest official UK VTT study and in other similar studies across Europe. Nonetheless, their mere existence poses significant limitations considering the importance of deriving national VTT estimates to be used in project appraisal, which would be driving future investments for at least the following decade, given their slow update rate documented so far. Even if the analyst is able to successfully account for the majority of those biases in the estimated VTTs, it requires significant effort to do so, which increases the cost and the time required to design an appropriate survey that would minimise any source of hypothetical bias, which is still never guaranteed.

Despite those limitations and potential pitfalls, SP surveys are currently the state-of-the-art approach for VTT estimation, with RP data being used only as auxiliary data to inform the attributes in the SP survey (Small, 2012; Ehreke et al., 2015). Studies have also proposed a combination of RP-SP choices during estimation for forecasting purposes acknowledging the hypothetical natural of SP data and the limitations that could arise in forecasting future demand (Cherchi and Ortuzar, 2006). Nonetheless, even in those cases, the VTT estimation primarily relies on SP data to avoid the limitations of RP data to capture non-linearities in the sensitivities, a notion that is still relevant among the research community due to the data limitations of the past. The dominance of SP data on VTT estimation can be clearly seen by examining the latest reported studies on national VTT values. The Danish (Fosgerau, 2006), Swiss (Axhausen et al., 2006), Norwegian (Halse et al., 2022), Swedish (Börjesson and Eliasson, 2014), Dutch (Kouwenhoven et al., 2014), German (Ehreke et al., 2015) and the UK (Batley et al., 2019) national VTT studies have based their analysis on SP data using only a limited number of RP trips as reference points for the SP attributes.

2.2. GPS data for transport research

Various forms of new emerging data sources are increasingly being used for transport-related research during the past decade (Grant-Muller et al., 2021). Mobile phone data have been one of the first emerging data sources, which gained popularity among researchers and practitioners due to their ubiquitous nature and their ability to capture a wide spectrum of daily urban mobility patterns (Bwambale et al., 2019; Essadeq and Janik, 2021). Similar to mobile phone data, social media data (e.g. Twitter, Foursquare,

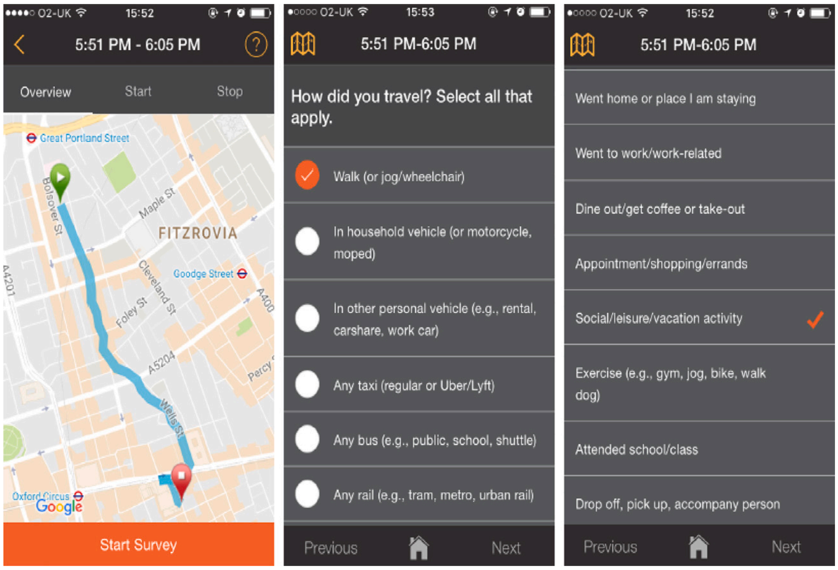

Fig. 1. User interface of smartphone application used for the GPS trip diary (Calastri et al., 2020).

Weibo etc.) have also been used for the purpose of understanding individual mobility behaviour and are capable of providing interesting insights of aggregate mobility patterns (Yan and Zhou, 2019; Ebrahimpour et al., 2020).

Contrary to the aforementioned emerging passively collected data sources, GPS data offer the advantage of providing large panels of real-world observed behaviour at a high spatio-temporal resolution. GPS data have been extensively used for transport research during the past twenty years (Wolf, 2004) for studies of understanding daily travel time budgets (Gallotti et al., 2015), destination (Huang and Levinson, 2015, 2017), mode and destination (Tsoleridis et al., 2021), mode (Montini et al., 2017; Calastri et al., 2018), trip chaining and mode (Huang et al., 2021) and route choice behaviour (Li et al., 2005; Hess et al., 2015).

Traditional pen-and-paper trip diaries are likely to result in misreporting of trips, where shorter trips might be omitted by the individuals, while also time spend on travelling tends to be overstated compared to the actual one leading directly to biased (lower) VTT estimates (Kelly et al., 2013). GPS data, on the other hand, due to their passively collected nature, do not require individuals to recall their daily trips (Hess et al., 2015) resulting on average in a larger number of trips per day (Forrest and Pearson, 2005). The type of the GPS device can also have an impact on the quality of the reporting trips, with studies based on GPS loggers noting that many individuals tend to forget them Bohte and Maat (2009). Contrary to that, it is much less likely for individuals to forget their smartphone making it a more suitable GPS device for capturing their trips (Calastri et al., 2020).

Despite those advantages, however, there are instances of missing trips in those datasets, as well, due to signal issues or due to individuals turning off the GPS tracking from their devices either for battery or privacy preservation (Calastri et al., 2020). GPS data provide values at a very high resolution, however that characteristic is also one of their most important limitations since significant pre-processing efforts are required to make the data useful for analysis and for deriving insights on mobility behaviour (Stopher et al., 2005; Marchal et al., 2011).

Coupled with minimum input from the participants, such as mode and trip purpose, and an additional background household survey, GPS data have the potential of providing significant advantages over traditional RP data. Despite their extensive use in transport research over the past years, however, no study so far has tried to utilise such a dataset for the estimation of nation-wide VTT estimates and for their comparison with the official SP-based values.

3. Data

3.1. DECISIONS data

Several datasets are utilised in the current study. A behavioural model of mode choice is estimated using the labelled GPS dataset, which was collected between November 2016-March 2017, as part of the "DECISIONS" research project aiming to understand individual transport and energy choices. A detailed description of the dataset (referred to as DECISIONS dataset in the remainder of this paper) is presented in Calastri et al. (2020). That survey consists of several submodules including a trip diary captured through GPS tracking using a smartphone application and a household survey capturing important sociodemographic information of the participants. The GPS trip diary includes the participants' trips during a 2-week period, in which additional information on the purpose and the chosen mode had to be provided at the end of each trip (semi-passively collected) as depicted in Fig. 1 showing the interface of the smartphone application.

The GPS diaries initially included 721 unique individuals and 56,693 observed trips around the UK (5.7 daily trips per individual) with the vast majority of those being around the region of Yorkshire and the Humber, and predominantly around the city of Leeds.

Fig. 2. Spatial distribution of interzonal flows between MSOAs across the UK.

As a result, only trips within the region of Yorkshire were selected for the subsequent analysis to avoid larger estimation errors for less represented areas, such as London. The spatial distribution of trips initially included in the dataset, represented as interzonal flows between MSOA zones across the UK, is depicted in Fig. 2, while Fig. 3 shows only the trips starting and finishing within the region of Yorkshire. The observed modes of transport included car, bus, rail, taxi, cycling and walking.

A significant effort was undertaken during the cleaning phase with an emphasis on detecting inconsistencies between consecutive trips, in terms of time (following trip starting before the end of the previous trip) and space (space gaps between two consecutive trips). Furthermore, a large number of trips were left *untagged*, meaning that participants did not provide mode and purpose information, and these thus had to be removed from the analysis. No pattern was identified for the erroneous observations or untagged trips that were removed and those were considered as missing at random for the subsequent analysis. Unique activity locations were defined by clustering the observed latitude/longitude coordinate pairs. Hierarchical Agglomerative Clustering (HAC) was used for that purpose, since it does not require the analyst to predetermine the number of clusters. A distance threshold of 200 m was selected for the observed destinations to be considered in the same cluster, which resulted in the most plausible clusters for the sample after testing thresholds between 50–500 m. As a result from that process, home and work locations were identified based on the tagged purposes for those locations, the time of the day that those locations were visited and the time spent there. Trips were then assigned into tours, starting from and finishing at the home location, per individual and for each day of the survey.

The aforementioned process allowed us to adopt a tour-based approach in terms of mode availability per individual and choice task. In that sense, if an individual chooses car for the first trip of the tour and she is the only person participating in that trip, we can safely assume that she is the driver and she will have to return the car back home. As a result, all mode alternatives will be available for the trips of the tour with the exception of the last trip, where only car is available. Therefore, in such a choice task the last trip is not included in the model, since it is not relevant in a mode choice context. Conversely, if the individual is not the only person participating in a car trip, then we cannot assume that she is the driver and as a result all modes will be available for the remainder of the tour. In the case, the individual chooses any other mode for the first trip, i.e. bus, rail, taxi, cycling or walking, then car would be available for the first trip of the tour and it would become unavailable for the remaining trips of the tour. In such cases, car would be included as an alternative only in the choice set of the first trip. Therefore, the utilised approach despite still being trip-based, it utilises some tour-based feasibility constraints in an attempt to increase the behavioural realism of the model.

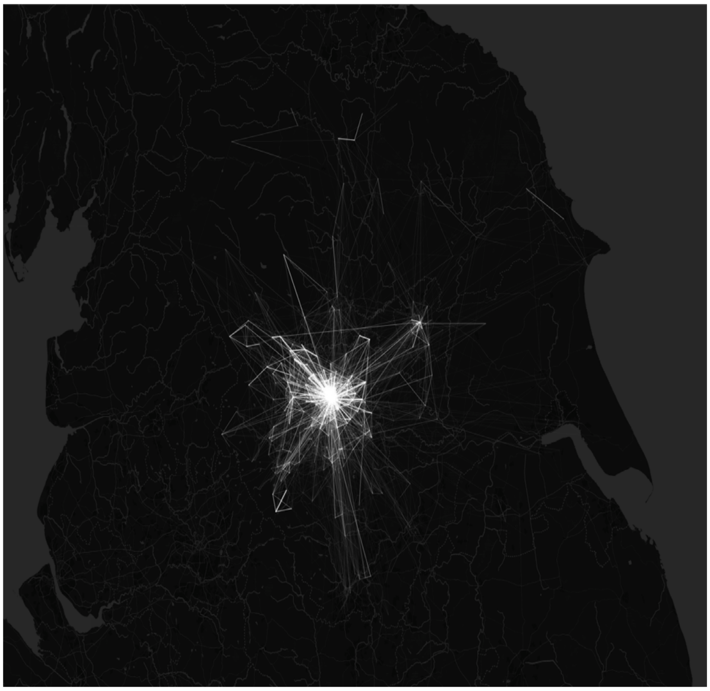

Fig. 3. Spatial distribution of interzonal flows between MSOAs across the region of Yorkshire.

A significant problem inherent to RP data and especially to those derived from new emerging data sources, such as GPS, is the lack of any information on the non-chosen alternatives. To overcome that obstacle, the "Directions" Google API was implemented providing travel times and distances for a range of modes.[1] Travel times/distances for both chosen and unchosen mode alternatives were re-calculated using the API for consistency reasons and to ensure that all values would come from the same data generation process (Calastri et al., 2018). The requests passed on to the API were for two weeks in the future from the date of analysis and for the same day of the week and time of day as in the initial dataset. Two weeks in the future were chosen, so as to avoid the short-term negative impacts of a recent traffic disruption and because the API cannot be used for past dates.

The API provided travel times for car/taxi based on traffic information for the specific time of day when the observed trip was performed. Travel times on the shortest distance routes are calculated for walking and cycling trips. For bus and rail trips, a timetable approach is used and a detailed breakdown of the whole route is provided including walking segments, waiting times and transfers between different services. That level of detail was essential in order to quantify in-vehicle and out-of-vehicle travel times and to be in line with the official VTT estimates. The travel times for the observed modes acquired from the API were compared with the stated travel times for validation purposes. The mean absolute difference across modes was very small, namely 8 mins, indicating that, on average, Google API is capable of providing travel times comparable with the actual ones.

Travel cost was also missing for all alternatives. Car costs were calculated using the official WebTAG specifications regarding fuel[2] and operating costs (Department for Transport, 2014). Parking costs were also calculated based on the location of the observed destination (central areas, local high streets etc.) and the activity duration there. Information on hourly and fixed parking costs was obtained from the local authorities in the region of Yorkshire. Fuel, operating and parking costs were added together to form the total car travel cost used for estimation. Bus and rail costs were calculated based on a distance-based fare of the most popular bus and rail operators in the region and a discount was applied for season ticket holders. Finally, taxi cost was calculated using fixed, hourly and distance-based average costs for different cities around Yorkshire.

The final dataset used for model estimation contained 12,524 trips and 540 unique individuals, which is significantly smaller than the SP sample used for analysis in the official study consisting of 7692 individuals and 15 choice tasks per individual (Hess et al., 2017). Regarding the observed/chosen modes, 47.6% were car trips, 14.6% bus, 5.2% rail, 3.2% taxi, 3.3% cycling and 26.1% walking trips. Commuting and business trips were observed for 19.1% and 9.7% of the sample, respectively. The majority of respondents were female (59.6%) of an average of 40 years old, 75.4% had at least one car in their household and finally 20.7% and 13.3% had a bus and rail season ticket, respectively. The average trip distance across modes per choice task is 2.4 miles

[1] More details can be found here: https://developers.google.com/maps/documentation/directions/overview.

[2] Historical petrol and diesel prices for the survey period (November 2016-March 2017) were obtained from https://www.racfoundation.org/data/uk-pump-prices-over-time.

(3.9 km) with a maximum of 61.1 miles (98.4 km) depicting the urban nature of the trips captured in this survey and the absence of longer-distance interurban trips.

In addition to the aforementioned tour-based approach for defining mode availability, a further step was taken to define mode consideration in order to form the final set of alternatives actually considered during the decision making process. Several approaches have been proposed to account for consideration of alternatives, which differ in their level of complexity ranging from fully probabilistic choice set formation approaches (Manski, 1977; Swait and Ben-Akiva, 1987; Calastri et al., 2018) to captivity models (Gaudry and Dagenais, 1979; Swait and Ben-Akiva, 1986) to the incorporation of penalties in the utility function for alternatives exceeding certain attribute thresholds (Cascetta and Papola, 2001; Martinez et al., 2009) and finally to defining consideration of alternatives in a deterministic manner with exogenous thresholds. In the current study, the latter approach was utilised taking into account the observed behaviour in the sample and the results obtained from the API. In that regard, car and taxi trips were excluded for short trips below 5 mins, which was the minimum observed time for those modes in the sample. Bus and rail were excluded for trips in which the API returned only walking segments due to short distances or lack of service. Finally, regarding cycling and walking, which are modes that require physical effort, they were excluded for long distance trips, namely 20 km and 3 km respectively, which were the maximum observed cycling and walking distances in the sample. Having said that, however, we do acknowledge that there are more behaviourally accurate ways to model alternative consideration, e.g. using a probabilistic choice set formation approach, which could form the basis of future research on that topic in order to assess the impact of uncovering such behavioural mechanisms on the estimated VTTs.

3.2. NTS dataset

The estimates derived from the best-performing model estimated on the DECISIONS were applied on the NTS dataset. The NTS is the official annual survey in the UK providing invaluable long-term information on travel behaviour and mobility trends.[3] Three consecutive years of the NTS data were used, namely 2015–2016–2017, to ensure a representative sample, while also providing an overlap with the period of the DECISIONS survey. A non-sensitive version of the NTS dataset was acquired,[4] where information on personal income was missing. Furthermore, household income was not reported for more than 65% of participants for each year rendering it practically unusable. In addition, since the parameters were estimated on a dataset containing trips mostly around the region of Yorkshire (DECISIONS), it was decided to exclude trips in London from the NTS data, due to the individuals there generally having a different set of available modes, e.g. underground. Contrary to London, the remaining areas around the UK have similar mode availability with regard to public transport, i.e. bus and rail. Furthermore, an important assumption made at this point is that individuals in Yorkshire exhibit a mobility behaviour, which is transferable to the rest of the UK.

The NTS dataset did not provide any information on travel cost for car trips and for a large number of bus and rail trips. For the former, only information on parking cost was provided and fuel and operating costs were imputed using the same approach as in the DECISIONS dataset (Department for Transport, 2014). Fare cost for bus and rail was missing or reported as zero for 49% and 13.4% of bus and rail trips, respectively, which were performed by season ticket holders. Since it is not reasonable to assume a zero VTT for season ticket holders, an average daily cost of a season ticket was applied for bus and rail, based on the cost calculations performed for the DECISIONS data.

The final NTS dataset used for the analysis, excluding London-based trips, included 453,438 trips performed by 29,127 unique individuals (3.1 daily trips per individual), with 52.6% of those being female and with an average age of 50 years old. The average trip distance is 7.9 miles (12.7 km) with a maximum of 719 miles (1157.1 km) showing a more accurate depiction of mobility behaviour including both urban and interurban trips. Regarding the observed modes, 80.4% were car trips, 5.1% bus, 1.2% rail, 1.2% taxi, 1.9% cycling and 10.1% walking trips. Finally, commuting and business trips account for 17.4% and 4.0% of NTS trips, respectively. Detailed descriptive statistics of DECISIONS and NTS trips per mode and purpose are presented in Table 1.

4. Modelling framework

The VTT estimates presented in the current study are derived from a behavioural model based on the Discrete Choice Modelling (DCM) framework (Ben-Akiva and Lerman, 1985; Train, 2009). A DCM framework based on Random Utility Maximisation assumes that each individual n has a preference for a specific alternative i among a set of J alternatives in a choice task t represented as a latent utility U_{int} consisting of a deterministic part V_{int} and a disturbance term ϵ_{int}. Different distributional assumptions about the disturbance term would yield a different specification form. The most commonly used specification is the Multinomial Logit model (MNL) assuming a type-I (Gumbel) Extreme Value distributed disturbance term (McFadden, 1973). The deterministic part V_{int} consists of alternative- and individual-specific attributes, x_{int} and z_n, respectively, as shown in Eq. (1). The choice probabilities of an MNL model are derived from Eq. (2).

$$U_{int} = V_{int} + \epsilon_{int} = f(\beta, x_{int}, z_n) + \epsilon_{int} \tag{1}$$

$$P_{int}(\beta) = \frac{e^{V_{int}}}{\sum_{j=1}^{J} e^{V_{jnt}}} \tag{2}$$

[3] Details can be found here: https://www.gov.uk/government/collections/national-travel-survey-statistics.

[4] The NTS dataset was acquired from https://beta.ukdataservice.ac.uk.

Table 1
Number of DECISIONS and NTS trips per mode and purpose.

Mode	Commuting	Business	Other (non-work)	Total
DECISIONS trips				
Car	1015 (8.1%)	693 (5.5%)	4253 (34.0%)	5961 (47.6%)
Bus	510 (4.1%)	201 (1.6%)	1117 (8.9%)	1828 (14.6%)
Rail	243 (1.9%)	55 (0.4%)	350 (2.8%)	648 (5.2%)
Taxi	23 (0.2%)	30 (0.2%)	352 (2.8%)	405 (3.2%)
Cycling	121 (1.0%)	19 (0.2%)	269 (2.1%)	409 (3.3%)
Walking	477 (3.8%)	214 (1.7%)	2582 (20.6%)	3273 (26.1%)
Total	2389 (19.1%)	1212 (9.7%)	8923 (71.2%)	**12,524 (100%)**
NTS trips				
Car	62,750 (13.8%)	16,199 (3.6%)	285,594 (63.0%)	364,543 (80.4%)
Bus	5054 (1.1%)	406 (0.09%)	17,580 (3.9%)	23,040 (5.1%)
Rail	2133 (0.5%)	488 (0.1%)	2916 (0.6%)	5537 (1.2%)
Taxi	725 (0.2%)	143 (0.03%)	4786 (1.0%)	5654 (1.2%)
Cycling	3475 (0.8%)	254 (0.1%)	4927 (1.1%)	8656 (1.9%)
Walking	4729 (1.0%)	487 (0.1%)	40,792 (9.0%)	46,008 (10.1%)
Total	78,866 (17.4%)	17,977 (4.0%)	356,595 (78.6%)	**453,438 (100%)**

where β is a vector of parameters to be estimated.

The basic MNL specification assumes that individuals will have the same sensitivity to the specified parameters. Deterministic taste variation in response to specific attributes can be captured as shifts from their base level for specific types of individuals or choice tasks. In the present study, deterministic heterogeneity was captured by specifying shifts from the base level of the alternative specific constants (ASCs) for specific sociodemographic attributes. Furthermore, shifts were also included for the base time and cost parameters of level-of-service (LOS) variables for business and commuting trips. An elasticity specification was used for interactions with continuous sociodemographic attributes, such as age and income, with a separate beta being specified for respondents who did not provide any income information.

Even in the case of accounting for deterministic heterogeneity, however, it is reasonable to assume that some degree of heterogeneity would still remain uncaptured among and/or within individuals leading to biased estimates. Mixed Logit models (McFadden and Train, 2000) can be used to account for that, offering a more flexible specification, where parameters are allowed to vary randomly across individuals. Mixed Logit models are considered as the most general form of a Logit, since they are able to approximate any other specification (McFadden and Train, 2000). The results, however, will largely depend on the distributional assumptions for each random parameter, a task bestowed on the analyst.

The choice probabilities in a mixed Multinomial Logit Model (MMNL) model are now given by an integral over the distribution of individuals' sensitivities (which follow a density function $\phi(\beta|\Omega)$), where this integral does not offer a closed form solution. Simulated log-likelihood estimation is an alternative way of calculating the integral of choice probabilities, based on drawing random numbers from a pre-specified distribution. From that process, the choice probabilities can be calculated as the average over the draws (Eq. (3)) and the simulated log-likelihood can be computed as shown in Eq. (4).

$$\widehat{P_{int}(\Omega)} = \frac{1}{R}\sum_{r=1}^{R} P_{int}(\beta^r) \tag{3}$$

$$SSL(\Phi) = \sum_{n=1}^{N} ln(\widehat{P_{int}(\Phi)}) \tag{4}$$

where β^r is a random draw from a distribution with $\phi(\beta|\Omega)$.

It is reasonable to assume that the impact of LOS parameters should be strictly negative, indicating that an additional minute spent travelling or an additional unit of cost spent for a trip will decrease the utility and therefore the choice probability for a certain mode alternative. The specified distribution for the random LOS parameters should be able to account for that, with the negative log-normal distribution being the most applied one for that purpose. In the current study, the long tails of the log-normal distribution resulted in numerical issues during estimation, prompting us to take a different approach. As a consequence, the negative log-uniform distribution was chosen instead, with its shorter tails ensuring no problems during estimation, similarly to the official UK study, which provided the first large scale application of that distribution (Hess et al., 2017). Under that distribution, a variable x is log-uniformly distributed, if $y = log(x)$ is uniformly distributed. The log-uniform distribution is defined by two additional parameters, a and b denoting its lower bound and spread, respectively. The mean and the variance of the log-uniform distribution are calculated as following (Hess et al., 2017):

$$E(\beta_0) = \frac{e^{a+b} - e^a}{b} \tag{5}$$

$$Var(\beta_0) = e^{2a}\left[\frac{e^{2b}-1}{2b} - \frac{(e^b-1)^2}{b^2}\right] \tag{6}$$

In total, nine parameters were specified as random, namely travel time for car, taxi, walking and cycling, in-vehicle (IVT) and out-of-vehicle (OVT) travel times for bus and rail and finally travel cost. Due to the multidimensionality of the integral, Modified Latin Hypercube Sampling (MLHS) draws were chosen over Halton draws to avoid the multicollinearity issues identified with multidimensional Halton sequences (Hess et al., 2006). For the simulated log-likelihood estimation, 1000 MLHS numbers r_Us were drawn from a uniform distribution for each randomly distributed β, which was specified as $\beta_{LU(a,b)} = e^{a+b*r_U}$. At that number of draws, a sufficient level of stability was observed among the estimates and model fit, hence it was decided not to increase the number of draws any further.

Finally, another issue worth addressing is the potential presence of heteroscedasticity in the choices occurring from the variance differences across the choice tasks. In the official SP-based VTT study, a multiplicative error term was used instead of the additive one in Eq. (1) following the proposed specification in Fosgerau and Bierlaire (2009), such as $U_{int} = V_{int}\epsilon_{int}$. That specification can be simplified by taking a logarithmic transformation as $log(U_{int}) = log(V_{int}) + log(\epsilon_{int})$ (Fosgerau and Bierlaire, 2009), which requires a strictly positive V_{int}. Instead of that specification, we tried to capture heteroscedasticity by assuming that more uncertainty, hence variance, will exist for choice tasks/trips of longer distances. Therefore, for those trips, the systematic part of the utility, V_{int}, will be smaller compared to trips of shorter distances. In order to capture that, additional scale parameters ϕ_l are specified for different distance bands l and multiplied with the utility function with one scale parameter ϕ_{l_0} being fixed to 1.0. If the base ϕ_{l_0} refers to trips in the shortest distance band then the remaining estimated ϕ_l should be smaller and ideally decreasing as distance increases.

5. Modelling results

The outputs of the behavioural models estimated on the DECISIONS dataset, base MNL, scaled MNL and scaled MMNL, are presented in Table 2. The scaled MNL model showed significant model fit improvements over the base MNL of 15.5 LL units with 2 additional parameters, namely the scaling parameters ϕ. Those scaling parameters are significantly lower than 1.0 and are decreasing as the distance band increases. Therefore they are able to uncover significant heteroscedasticity among the choices based on distance, hence conforming to our initial hypothesis.

The scaled MNL model was used as the initial point of departure for the MMNL specification with the purpose of capturing unobserved heterogeneity. The MMNL model with 9 additional parameters provided significant improvements in model fit reducing the LL by 1166.5 units from the scaled MNL model. The adjusted ρ^2 of 0.7723 also signifies that the model is able to explain a significant portion of the variation in the dataset. The Alternative Specific Constants (ASCs) reveal that, all else held equal, individuals of higher income or those who are employed have a negative inherent preference for bus, while also bus is less preferred for trips over the weekend. Interestingly, individuals have an inherent positive attitude for rail compared to car signifying the perceived quality superiority of rail over bus, despite both being public transport modes. Individuals have a negative preference for taxi, although that is not the case for younger individuals below 30 years old. Individuals of lower education (with no undergraduate degree), which can be considered a proxy of income, have a significantly higher dispreference for taxi. Furthermore, there is a negative preference for cycling and an even higher dispreference from unemployed individuals, but less so from male individuals and students. Cycling is also more preferred by both the lowest and the highest personal income bands. Finally, all else held equal, individuals and specifically those of younger ages and students have a positive inherent preference for walking.

The importance of capturing non-linear sensitivities with a Box–Cox transformation (Box and Cox, 1964) of travel time and cost attributes for a more accurate VTT estimation has been stated before in the literature (Koppelman, 1981; Gaudry et al., 1989). A common approach is to have either linear or logarithmic specifications of time and cost attributes in the utility function assuming a linear increase of sensitivities or a decreasing one, respectively, as the attribute values increase. Such a specification can be limiting as it assumes that one of those two extreme cases exist in the sample. On the other hand, a Box–Cox transformation provides a more generalised specification as it allows the analyst to capture non-linearities across the whole spectrum of possible values. Using a Box–Cox transformation, an attribute x is specified as $\frac{x^\lambda - 1}{\lambda}$ with λ being an estimable parameter capturing the degree of non-linearity. If λ is not statistically different than 1.0, then the sensitivities for that particular attribute are indeed increasing linearly. If λ is not statistically different than 0, then the sensitivities take a logarithmic form leading to a steep increase at first for small values followed by a decreasing rate for high attribute values (decreasing marginal disutilities). In most cases, the estimated λ will be between 0 and 1.0, however empirical evidence shows that it can also take values above 1.0 capturing a slow increase in sensitivities for small values and followed by a steeper increase for higher values (increasing marginal disutilities) as shown in Gaudry et al. (1989).

In the current study, a Box–Cox transformation was specified for all LOS attributes in order to assess the estimated λ parameters and their behavioural meaning. For the final model, it was decided to keep a Box–Cox transformation for car, bus, taxi, cycling, walking, OVT bus and OVT rail, which resulted in $\lambda = 0.4581$ significantly different than 1.0 and 0 capturing significant decreasing marginal disutilities, but still not at the extreme of a logarithmic specification. Contrary to the above travel times, a linear specification was used for rail IVT. A behavioural meaning of those specifications and the respective uncovered sensitivities could be that individuals are more sensitive for higher in-vehicle travel times when travelling by rail compared to travelling by car, for example. That can be attributed to an increased discomfort caused by longer distance rail trips or to an increased time restriction for the rest of the daily schedule imposed by the longer distance trip, which prompts the individuals to choose faster and more expensive services. Commuters and business car travellers have also increased time sensitivities due to the importance of arriving on time for those activities compared to non-work trips. Increased car time sensitivities were found for trips during the am peak period (7.00–10.00), while the opposite was true for the pm period (16.00–19.00), relative to all other time periods of the day. That denotes the increased time restrictions of morning trips and the need to arrive on time to the various destinations (mostly work locations), compared to evening trips, most of which are either leisure, shopping or returning trips to home. The decreased time

Table 2
Fit statistics and modelling outputs.

Fit statistics	Base MNL	Scaled MNL	Scaled MMNL
Log-likelihood (0)	−14,974.45		
Log-likelihood (model)	−4535.41	−4519.92	−3353.41
Adjusted ρ^2	0.6941	0.6950	0.7723
AIC	9162.81	9135.83	6820.82
BIC	9504.84	9492.73	7244.64
Number of parameters	46	48	57
Number of individuals	540		
Number of observations	12,524		
Parameter	Estimate (Rob. t-rat. 0) [Rob. t-rat. 1]		
	Base MNL	Scaled MNL	Scaled MMNL
Alternative-specific constants			
Constant Bus	−0.0929 −0.26	0.3195 (0.80)	−0.0905 (−0.20)
Constant Bus shift for personal income>50k	−1.1902 −1.44	−1.4445 (−1.46)	−4.4972 (−5.77)
Constant Bus shift for weekend	−0.7165 (−4.05)	−0.8230 (−3.99)	−1.1573 (−4.19)
Constant Bus shift for unemployed individuals	0.7378 (1.08)	0.7066 (0.94)	0.7765 (1.43)
Constant Rail	2.4421 (2.21)	3.1648 (2.58)	2.3179 (2.23)
Constant Taxi	−1.8075 (−4.09)	−1.3129 (−2.37)	−1.8398 (−2.91)
Constant Taxi shift for male	−0.6434 (−2.10)	−0.7479 (−2.19)	−0.5739 (−1.32)
Constant Taxi shift for age 18–24	1.5014 (4.67)	1.8120 (4.66)	2.8605 (6.74)
Constant Taxi shift for age 25–29	0.9324 (2.64)	1.1385 (2.79)	2.0238 (3.86)
Constant Taxi shift for lower education levels	−1.3979 (−2.46)	−1.5940 (−2.44)	−2.4137 (−2.66)
Constant Taxi shift for personal income 40k–50k	−0.7975 (−2.18)	−0.9311 (−2.26)	−1.3510 (−2.35)
Constant Cycling	−4.0728 (−7.85)	−4.1352 (−7.21)	−4.3336 (−6.71)
Constant Cycling shift for male	1.1047 (2.69)	1.2663 (2.80)	2.0747 (4.89)
Constant Cycling shift for personal income 10k–20k	0.8020 (1.93)	0.9510 (2.08)	2.0946 (4.97)
Constant Cycling shift for personal income 75k–100k	3.5150 (3.49)	4.0077 (3.61)	5.7908 (4.96)
Constant Cycling shift for not reported income	−2.4544 (−1.69)	−2.9037 (−1.72)	−2.9202 (−2.89)
Constant Cycling shift for weekend	−0.6604 (−2.01)	−0.7803 (−2.06)	−1.5423 (−2.51)
Constant Cycling shift for student	1.1559 (2.23)	1.4223 (2.49)	2.8486 (5.57)
Constant Cycling shift for unemployed individuals	−1.1964 (−2.22)	−1.5084 (−2.59)	−5.5566 (−10.02)
Constant Walking	3.0294 (5.73)	3.1564 (5.34)	3.4426 (4.95)
Constant Walking shift for age 18–29	0.6964 (2.62)	0.6948 (2.55)	0.8225 (2.82)
Constant Walking shift for lower education levels	−0.7563 (−3.47)	−0.7843 (−3.43)	−1.1039 (−4.09)
Constant Walking shift for weekend	−0.5704 (−2.87)	−0.6310 (−2.90)	−0.8050 (−3.11)
Constant Walking shift for student	0.7277 (2.37)	0.8039 (2.53)	1.1111 (2.92)
LOS parameters			
Travel time Car (Box–Cox)	−0.2426 (−3.02)	−0.1680 (−2.33)	–
Travel time Car shift for commuting	−0.1478 (−2.92)	−0.1709 (−2.95)	−0.2347 (−3.37)*
Travel time Car shift for business	−0.0408 (−0.97)	−0.0308 (−0.65)	−0.1180 (−1.79)
Travel time Car shift for am peak	−0.1637 (−3.08)	−0.1817 (−3.07)	−0.2315 (−3.48)*
Travel time Car shift for pm peak	0.1282 (2.47)	0.1279 (2.43)	0.1720 (2.41)*
IVT Bus (Box–Cox)	−0.1281 (−3.19)	−0.1363 (−3.32)	–
IVT Bus shift for am–pm peak	−0.0998 (−3.33)	−0.1010 (−3.29)	−0.0902 (−2.51)*
IVT Rail (linear)	−0.0080 (−0.66)	−0.0105 (−0.71)	–
IVT Rail shift for am–pm peak	−0.0327 (−2.91)	−0.0366 (−2.65)	−0.0472 (−2.54)*
Travel time Taxi (Box–Cox)	−0.4525 (−3.22)	−0.5299 (−3.40)	–
Travel time Taxi shift for am peak	−0.1709 (−2.71)	−0.1918 (−2.85)	−0.2877 (−3.70)
Travel time Taxi shift for pm peak	−0.1423 (−3.24)	−0.1346 (−2.96)	−0.1305 (−2.31)
Travel time Cycling (Box–Cox)	−0.3343 (−3.15)	−0.3478 (−3.22)	–
Travel time Walking (Box–Cox)	−0.6774 (−3.26)	−0.6116 (−3.32)	–
Box–Cox lambda for Travel time for Car, Bus, Taxi, Cycling, Walking	0.5424 (5.25) [−4.43]	0.5942 (5.91) [−4.03]	0.5932 (5.57) [−3.82]*
OVT Bus (Box–Cox)	−1.1484 (−5.59)	−1.2007 (−5.57)	–
OVT Bus for income non respondents	−1.1920 (−5.14)	−1.2663 (−5.05)	−1.4745 (−6.62)

(*continued on next page*)

Table 2 (*continued*).

OVT Rail (Box–Cox)	−1.7365 (−3.50)	−1.8308 (−3.44)	–
OVT Rail for income non respondents	−1.4853 (−3.04)	−1.5539 (−2.93)	−1.4900 (−4.48)
Box–Cox lambda for OVT Bus, OVT Rail	0.1452 (1.59) [−9.35]	0.1884 (1.97) [−9.59]	0.3131 (4.67) [−10.25]
Income elasticity for OVT Bus, OVT Rail	0.0880 (2.26)	0.0962 (2.36)	0.0572 (1.11)
Travel cost (log)	−0.8362 (−10.73))	−0.9428 (−10.43)	–
Random LOS parameters			
a of Travel time Car (Box–Cox)	–	–	−0.1553 (−0.60)*
b of Travel time Car (Box–Cox)	–	–	−3.0508 (−5.60)*
a of IVT Bus (Box–Cox)	–	–	−0.2598 (−0.97)*
b of IVT Bus (Box–Cox)	–	–	−2.9434 (−5.64)*
a of IVT Rail (linear)	–	–	−0.9137 (−7.59)*
b of IVT Rail (linear)	–	–	−7.4859 (−4.01)*
a of Travel time Taxi (Box–Cox)	–	–	0.9373 (4.86)
b of Travel time Taxi (Box–Cox)	–	–	−1.7604 (−8.33)
a of Travel time Cycling (Box–Cox)	–	–	2.2121 (8.32)
b of Travel time Cycling (Box–Cox)	–	–	−3.6775 (−8.09)
a of Travel time Walking (Box–Cox)	–	–	−0.6679 (−1.81)
b of Travel time Walking (Box–Cox)	–	–	0.8886 (6.97)
a of OVT Bus (Box–Cox)	–	–	1.3095 (6.15)
b of OVT Bus (Box–Cox)	–	–	−1.8183 (−7.20)
a of OVT Rail (Box–Cox)	–	–	0.0579 (0.22)
b of OVT Rail (Box–Cox)	–	–	1.1181 (7.41)
a of Travel cost (log)	–	–	−1.8271 (−4.53)*
b of Travel cost (log)	–	–	3.0789 (5.83)*
Scale parameters			
Scale ϕ_1 for trip distances <3 km	–	1.0000 (–)	1.0000 (–)
Scale ϕ_2 for trip distances 3 km–20 km	–	0.8313 (16.50) [−3.35]	0.8031 (14.84) [−3.64]
Scale ϕ_3 for trip distances >20 km	–	0.6805 (8.54) [−4.01]	0.6560 (7.09) [−3.72]

sensitivity for car in the pm period could also denote the individuals' desire to take that extra time to diffuse themselves from the stress of work in the privacy of their car. That difference in time sensitivities between am–pm periods, however, was not evident for public transport and taxi trips. More specifically, individuals showed the same increased sensitivities in both periods for bus and rail trips possibly due to similar crowding levels in those modes from people going to and returning from work resulting in similar unpleasant conditions. For taxi trips, individuals had a lower sensitivity for pm compared to am period, but still higher relative to all other periods of the day.

A Box–Cox λ not statistically different than 0 was found for travel cost, which led us to keep a logarithmic specification of travel cost for the final models presented here signifying the presence of cost damping effects in the sample (Daly, 2010). That also signifies that cost sensitivity decreases at a faster rate than the time sensitivities, which will result in higher valuations for larger potential time savings, i.e. in longer distance trips (De Borger and Fosgerau, 2008; Small, 2012). Different combinations of personal and household income were specified as elasticities for the LOS parameters, with only the elasticity of personal income and OVT for bus and rail resulting in statistically significant estimates for the MNL models, at least. The sign of the parameter is positive, meaning that as income increases, the sensitivity to OVT also increases, which is behaviourally sensible. That OVT-income elasticity, however, became statistically insignificant in the MMNL model. Finally, all of the β parameters of the log-uniform distributions, i.e. the spread, were found to be statistically significant capturing significant inter-individual heterogeneity.

6. Values of travel time estimates

The estimated parameters of the scaled MMNL model (see Table 2) were applied to the NTS dataset, which provides mobility-related information on a sample of the UK population. Acknowledging the fact that the NTS dataset is still a sample of the UK population, we further adjusted the resulting VTTs using appropriate distance-based factors in order to calculate VTT values representative for the population of the UK (with the exclusion of London). Distance correction was performed using mode-specific information on travel distances for work trips obtained from the Census of 2011. The process is detailed below and also depicted in Fig. 4.

Sample level VTT calculation

VTTs were computed using sample enumeration over the car, bus and rail trips of the NTS sample. VTTs are calculated as the relative importance of one unit of change in time relative to one unit of change in cost. In mathematical terms, that is represented as the ratio of the partial derivatives of travel time over travel cost, as shown in Eq. (7).

$$VTT = \frac{\dfrac{\partial V_i}{\partial tt_i}}{\dfrac{\partial V_i}{\partial tc_i}} = \frac{\partial V_i}{\partial tt_i}\frac{\partial tc_i}{\partial V_i} \tag{7}$$

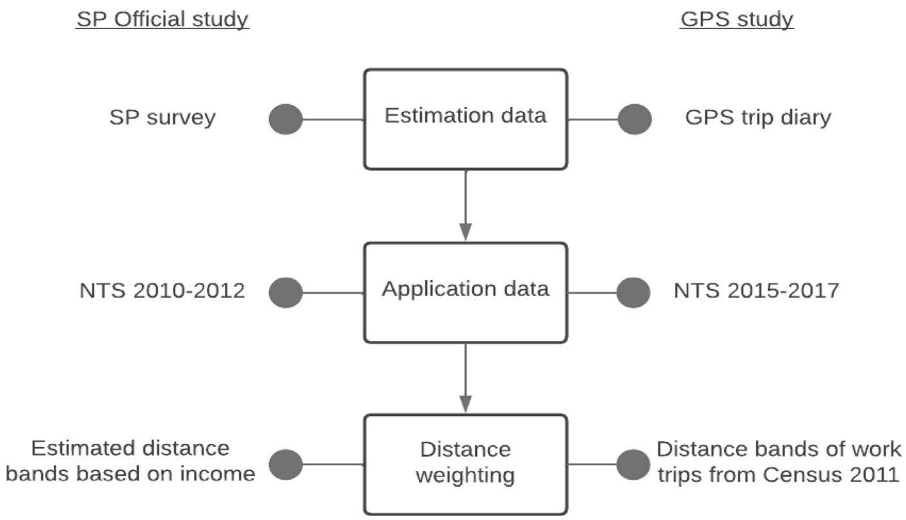

Fig. 4. Comparison approach between SP and GPS-based VTTs.

Sample enumeration is required because the attributes of travel time for car and bus and travel cost for all three modes enter the VTT calculation due to their specification in the respective utility functions. Taking car travel time as an example, its specification using a Box–Cox transformation leads to a partial derivative of $\beta_{tt}^{car} tt_{car}^{(\lambda_{time}-1)}$. Furthermore, travel cost, specified using a logarithmic transformation, leads to a partial derivative of $\beta_{tc} \frac{1}{tc_{car}}$. Therefore, the VTTs are computed for each choice task of the sample taking into account their respective trip attributes of travel time and cost for the chosen mode.

Deriving representative VTT for different distance-bands

In the guidance provided by DfT, the VTT values were both adjusted based on covariates included in the NTS data and were further weighted based on trip distances of different synthetic household income bands acknowledging the impact of distance on the VTT (Zhang and Laird, 2014; Batley et al., 2019). In the current study, due to the absence of any information relating to the distance weighting per income band, only the first correction was performed. According to that, distance-based factors were derived based on the relative importance of the mode-specific distance bands from the Census 2011 (excluding London) over the NTS distance bands and applied on the sample level VTTs (Eq. (8)). For that purpose, the NTS car, bus and rail trips were allocated to the same eight distance bands as the ones in the Census, namely 0–2 km, 2–5 km, 5–10 km, 10–20 km, 20–30 km, 30–40 km, 40–60 km and over 60 km, and the trip distributions over those distance bands were calculated for both datasets. The distance-based correction factor w_t for trip t was computed as $w_t = \frac{perc^{C}_{d_{t,i,p,m}}}{perc^{NTS}_{d_{t,i,p,m}}}$, where $perc^{C}_{d_{t,i,p,m}}$ and $perc^{NTS}_{d_{t,i,p,m}}$ are the Census and the NTS distributions of distance band i for trip t of purpose p and mode m, respectively.[5] Due to the lack of any additional explicit information regarding distances for non-work trips, the same correction factors were applied in those cases, as well. The distance band distributions of the NTS and the Census 2011 per mode and purpose, as well as the respective correction factors applied are presented in Table 3.

$$\overline{VTT_{p,m}} = \frac{\sum_t (w_t VTT_{t,p,m})}{\sum_t w_t} \tag{8}$$

where $VTT_{t,p,m}$ and $\overline{VTT_{p,m}}$ are the VTT for choice task t, purpose p and mode m and the weighted average VTT, respectively.

Due to the complex specification of the utility function and the resulting VTT calculation, we rely on simulation for the calculation of standard errors for the different VTTs and not on the most commonly used approach of the Delta method (Daly et al., 2012). Confidence intervals and standard errors for the estimated VTTs were calculated using multivariate normal draws based on the estimated parameters and the covariance matrix of the behavioural model (Train, 2009). Specifically, 3000 draws for the estimated parameters were drawn and simulated VTTs were calculated for each of the 3000 samples. At that number of draws the distance-adjusted means of the simulated VTTs per mode and purpose had only small discrepancies from $\overline{VTT_{p,m}}$, so no further draws were deemed necessary for the analysis. The 95% confidence interval was then calculated for the resulting simulated VTT distribution per mode and purpose using the percentile interval method and the standard errors were calculated as the standard deviations of the simulated VTT distributions. Finally, the t-stat of the difference between the estimated VTT means were calculated based on

[5] Details can be found here: https://www.nomisweb.co.uk/census/2011/dc7701ewla.

Table 3
Distance band distributions and distance-based correction factors per mode.

Distance band	Census 2011 (%)	NTS (%)	Correction factor
Car			
0–2 km	13.52	7.65	1.77
2–5 km	21.83	20.57	1.06
5–10 km	22.67	21.48	1.06
10–20 km	21.97	23.31	0.94
20–30 km	9.04	11.19	0.81
30–40 km	3.93	5.03	0.78
40–60 km	3.18	5.89	0.54
Over 60 km	3.87	4.88	0.79
Bus			
0–2 km	12.39	3.91	3.17
2–5 km	40.22	30.13	1.34
5–10 km	28.18	37.17	0.76
10–20 km	11.91	23.32	0.51
20–30 km	2.67	4.08	0.65
30–40 km	1.00	0.91	1.11
40–60 km	1.05	0.34	3.08
Over 60 km	2.57	0.15	16.74
Rail			
0–2 km	3.55	–	–
2–5 km	5.74	2.97	1.93
5–10 km	10.84	9.09	1.19
10–20 km	17.62	22.64	0.78
20–30 km	15.75	18.63	0.85
30–40 km	12.10	17.28	0.70
40–60 km	16.77	10.08	1.66
Over 60 km	17.65	19.31	0.91

Eq. (9).

$$t-stat_{diff} = \frac{\overline{VTT_{GPS}} - \overline{VTT_{SP}}}{\sqrt{s.e._{GPS}^2 + s.e._{SP}^2}} \tag{9}$$

VTTs in the official study were segmented into distance bands of 0–20 miles, 20–100 miles and over 100 miles, however, those were later revised by the DfT to 0–50 miles, 50–100 miles, 100–200 miles and over 200 miles (Batley et al., 2019). The behavioural model in the current study was estimated on a dataset where the maximum trip distance was 61.1 miles. Despite that, we decided to apply the estimates on NTS trips of longer distances, as well, to present a more complete comparison across the different distance bands with the official VTT study. As a result, the same distance bands of <20 miles, 20–100 miles and over 100 miles have been used. In addition to not having long distance trips in the estimation data, the application data of NTS also includes only a small number of medium and long distance trips, namely 8.09% of trips are above 20 miles, 2.16% above 50 miles and 0.78% above 100 miles. The same issue of a small number of long distance NTS trips was tackled in the official SP study by conducting additional intercept sampling favouring trips of longer distance and specifically business trips (Batley et al., 2019), however that was not possible in our case. Therefore, we should acknowledge the fact that calculated VTTs for medium (20–100 miles) and longer distances (above 100 miles) might contain higher estimation errors. Furthermore, when comparing the derived VTTs of the current study with the official ones, the exclusion of London is an important factor to be taken under consideration, as well. As a consequence, in the current study there is no "Other PT" as a mode alternative, which was included in the official VTTs and mainly referred to London-specific mode alternatives, such as light rail and the underground. The exclusion of London-based trips from the NTS data, also has an impact on the total sample size for our GPS-based VTTs, which is much smaller than the NTS sample size used in the official SP study.

The official VTT estimates based on the latest nationwide UK SP survey (adjusted for 2016 prices) are presented in Table 4, both overall and distance segmented values, and are compared with the respective GPS-based VTT estimates of the current study. In the official VTT study, bus was not included as an alternative for business trips, hence we decided to follow the same approach here, as well, for consistency reasons. The overall VTT values, which are to be used for appraisal, show only negligible differences with the official ones, despite that only a limited number of longer distance trips was used in the calculation of the GPS-based VTTs and that a different distance weighting method was applied without any explicit information on non-work trips. The distance segmented values are mainly used for reporting purposes (Daly et al., 2014), however, interesting findings can be extracted by their examination. Firstly, there are very small differences between the GPS-based and SP-based VTTs for the shortest distance band below 20 miles, and there is a higher estimation accuracy on the VTTs of that distance band due to the larger sample size in the NTS data. Another reason for these small discrepancies could be the that the SP surveys were able to sufficiently capture individual

Table 4
Official VTT estimates per mode, purpose and distance band based on the latest SP survey (Batley et al., 2019) and the respective derived GPS-based VTT estimates (£/hour) (2016 prices).

Distance band	Commute trips	Other trips	Business trips			
	All modes	All modes	All modes	Car	Other PT	Rail
Official SP values						
All distances	11.69	5.34	19.01	17.46	8.69	28.80
<20 miles	8.63	3.78	8.67	8.56	8.69	10.54
20–100 miles	12.67	6.77	16.74	16.53	8.69	30.23
≥100 miles	12.67	9.67	29.85	26.84	8.69	30.23
GPS-based values						
All distances	12.90	5.40	17.13	16.60	–	33.43
<20 miles	11.24	4.68	11.16	11.20	–	9.78
20–100 miles	30.52	15.67	36.59	36.21	–	43.78
≥100 miles	75.57	24.29	82.25	70.12	–	187.96

Table 5
Confidence intervals and standard errors of the mean estimates for the overall official VTT estimates per mode and purpose (Batley et al., 2019; Hess et al., 2017) and the respective derived GPS-based VTT estimates (£/hour) (2016 prices).

Mode-Purpose	VTT (St.error) [95% C.I.]		t-stat diff
	SP-based values	GPS-based values	
Car			
Commuting	12.20 (2.05) [8.18–16.23]	13.52 (2.68) [9.17–19.78]	0.39
Business	17.46 (2.04) [13.47–21.45]	16.60 (4.60) [9.94–27.40]	−0.17
Other	5.12 (1.84) [1.53–8.72]	5.45 (1.28) [3.50–8.32]	0.15
Bus			
Commuting	3.29 (0.48) [2.34–4.23]	5.46 (1.14) [3.59–7.97]	1.75
Other	3.40 (0.43) [2.56–4.24]	3.99 (0.88) [2.60–5.94]	0.60
Rail			
Commuting	12.95 (0.91) [11.17–14.73]	12.41 (3.28) [7.26–20.00]	−0.16
Business	28.80 (2.49) [23.91–33.68]	33.43 (9.11) [19.82–55.09]	0.49
Other	9.05 (0.64) [7.80–10.30]	14.26 (3.94) [7.32–23.39]	1.31

mobility behaviour in such hypothetical scenarios of small trip distances. On the contrary, starker differences are observed across the remaining VTTs for medium to long distance trips, which to a large extend can be attributed to the small number of trips in those bands. Nonetheless, those significant differences could also be attributed to the inherently more unpredictable nature of longer trips that is more challenging to be sufficiently accounted for in the hypothetical setting of an SP survey. As it is evident from Table 4, the GPS-based values might be able to capture a more behaviourally accurate depiction of the VTTs, which increase significantly for longer distance trips capturing the individuals' increased time restrictions during their daily activity schedule, as it is also supported by the literature. Having said that, however, it is important to note that a more balanced sample would be required in terms of trip distances in order to draw more robust conclusions about that.

The standard errors of the estimated mode- and purpose-specific VTT means (across all distances) and their 95% confidence intervals are presented in Table 5 along with the t-statistic of the difference of the means between the GPS-based and the official SP-based VTTs. Overall, standard errors of the GPS-based VTTs are higher and that can be attributed to a higher degree of heterogeneity captured in our study compared to the SP survey or it could also simply be due to the smaller sample size. For all VTTs presented in Table 5, we cannot reject the null hypothesis of difference from the official valuations, at the 95% confidence level. Furthermore, the GPS-based VTTs follow the general trends of the SP values. Values for rail are higher than car and bus, with the latter has the lowest values overall. Furthermore, business VTTs are higher than commuting and other non-work trips. Commuting values are in general higher than non-work trips, however, the opposite was found for SP bus and GPS rail VTTs for non-work trips. Bus VTT for commuting trips show the largest difference compared to the rest, followed by other non-work rail VTTs, but those are still not statistically significant differences. Finally, in most cases, the GPS-based VTTs are higher than the SP-based ones conforming with previous evidence in the literature (Wardman et al., 2016), however the downward hypothetical bias for SP is less significant in that case.

The distributions of the simulated VTTs are also presented in the box plots in Figs. 5–7. In those plots, the impact of the smaller sample size for rail trips becomes evident as it leads to wider distributions highlighting the uncertainty around the estimation of those VTTs, contrary to the more compact distributions of commuting and non-work VTTs.

7. Discussion

The results of this study clearly demonstrate that the argument of RP data collection limitations of the past does not hold anymore in the current age of data revolution. Semi-passively collected emerging data sources have the ability to provide the analysts with

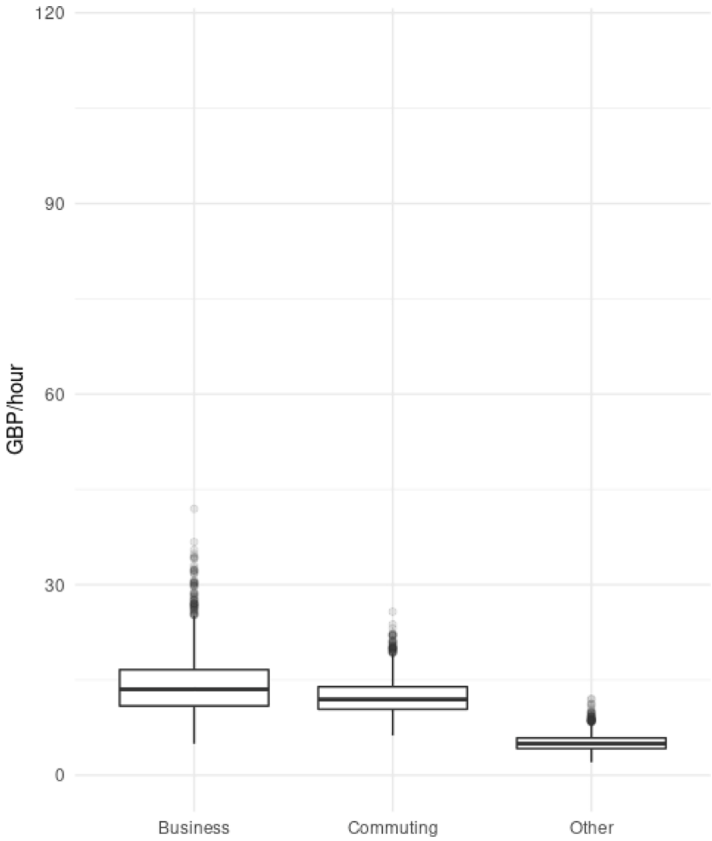

Fig. 5. Box plots of GPS-based Car VTTs per purpose.

large panels of observed mobility behaviour at a high spatio-temporal resolution and at a relatively low cost. Those types of datasets, in that case a GPS trip diary coupled with a background household survey, have the ability of providing robust behavioural models and VTT estimates statistically equal with official national values derived from traditional SP surveys.

It may be noted though that the overall sample size of the finally utilised estimation dataset was smaller compared to the SP survey, as a large share of trips had to be removed from the original DECISIONS data during the cleaning phase in order to exclude inconsistent, incomplete and untagged trips. Furthermore, the trips recorded in the DECISIONS dataset were mostly urban trips, while the official SP survey included longer distance trips, as well. The aforementioned limitations, however, can be easily overcome in nationally important studies by designing a more comprehensive data collection process. Those limitations can be partly justified since the DECISIONS dataset was not collected with the purpose of estimating nation-wide VTT values in mind. Despite those limitations, the study has two key findings which are of importance to transport planners and policy makers:

1. The study demonstrates that the overall VTT estimates were similar to the official SP-based values used for appraisal, even with a smaller sample size. As a result, there is a smaller hypothetical bias in the official SP study compared to the usual RP/SP documented in the literature across a range of studies (Wardman et al., 2016).
2. Segmenting the VTTs by distance bands, larger discrepancies start to become evident among longer distance bands with the SP-based VTTs being smaller in most cases. That hints to a downward bias for SP surveys potentially originating from the hypothetical nature of the longer trips, which made them difficult to comprehend. In contrast, the smaller differences of VTTs for shorter distance bands could mean that the design of the SP survey was sufficient enough to capture realistic mobility behaviour.

The results hence demonstrate that by harnessing recent technological advances in data collection, transport planners and policy-makers can make a successful shift to RP data sources, which have more behavioural validity compared to SP. Furthermore, the findings of the current study also demonstrate that smaller sample sizes derived from GPS smartphone data could be sufficient for the estimation of behaviourally accurate VTTs for the whole population. That finding could lead to a more frequent data collection process for the purpose of updating the national VTT estimates, compared to the so far slower update rate of traditional SP-derived VTTs (approximately every 10–20 years for most countries with some exceptions, e.g. in Sweden and Norway).

GPS-derived VTTs could also be used by policy-makers to complement the official SP-derived VTTs, since the more frequent GPS studies could help to detect any significant deviations from the previously SP-based estimated VTTs due to income increase or

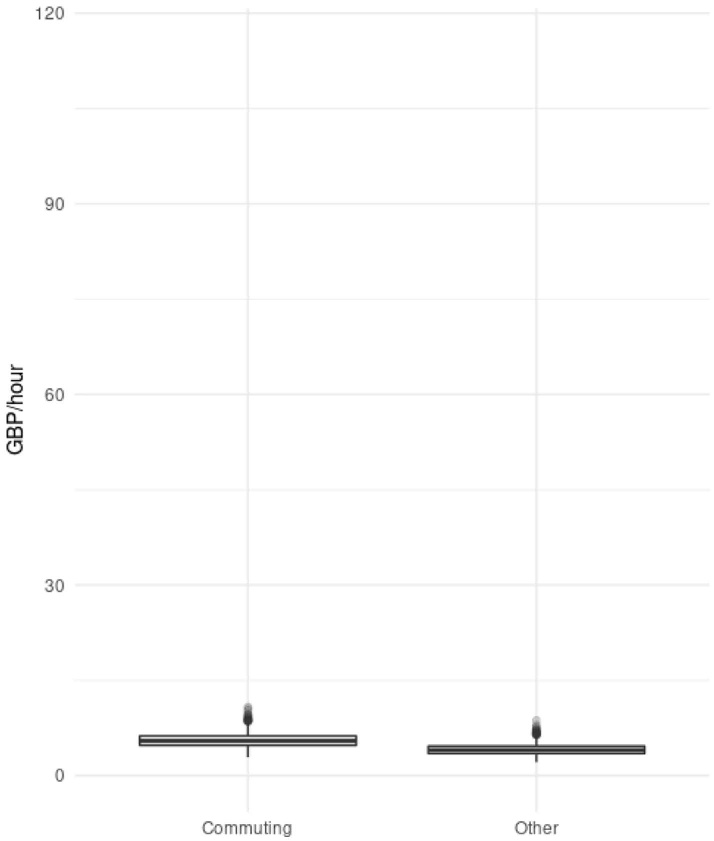

Fig. 6. Box plots of GPS-based Bus VTTs per purpose.

other unforeseen circumstances that could occur in the meantime, such as economic recession or the introduction of new disruptive modes/technologies into the transport market. Technologies like online shopping and its ever-increasing popularity especially in a post-Covid world, electric vehicles, shared ride modes (Uber, Lyft etc.) and policy initiatives like Mobility as a Service are constantly changing the transport sector, which has become more volatile than ever before. Transport is rapidly changing and the usual update rate of nation-wide official VTTs might be too slow nowadays to provide insights into the current trade-offs or even capture the sensitivities on new technologies in hypothetical scenarios and in a behaviourally realistic manner. As a result, new transport projects might not be properly evaluated if the appraisal is based on individual trade-offs that no longer represent the current behaviour of the target population.

8. Conclusions

The current paper presents a study of deriving VTT estimates in a manner comparable to the official values currently used in appraisal. Though the level of detail included in the initial GPS trip diaries provided challenges during the data cleaning phase, there were significant advantages in terms of accuracy. For example, the time-stamped geo-locations provided the ability to better capture individual mobility behaviour by making it possible to get precise travel times for the chosen modes and to extract travel times between specific latitude/longitude pairs (opposed to between TAZ centroids as done in traditional RP survey data). It also enabled the estimation of more behaviourally rich models by offering a more comprehensive representation of tours, where even very short stops and/or trips have been included.

Despite those advantages, it is worth noting that limitations do exist in the utilised dataset, as well. Information on trip attributes were obtained several years after the initial data collection period from an API that does not provide historical network information, hence the actual traffic conditions for each trip cannot be retrieved. Furthermore, additional information on weather conditions and other intrinsic information that would influence both the formation of the consideration set and the choice itself are not accounted for in the present study. Future studies using GPS data for VTT estimation should aim to incorporate a probabilistic choice set formation framework to account for the inherent latent choice sets in RP datasets. An immediate extension of the current study is to analyse the impact of such a framework on the estimated VTTs and assess their discrepancies from the official SP-based VTTs. Further trip-specific attributes can be incorporated in the analysis to enrich the estimation dataset regarding historical weather

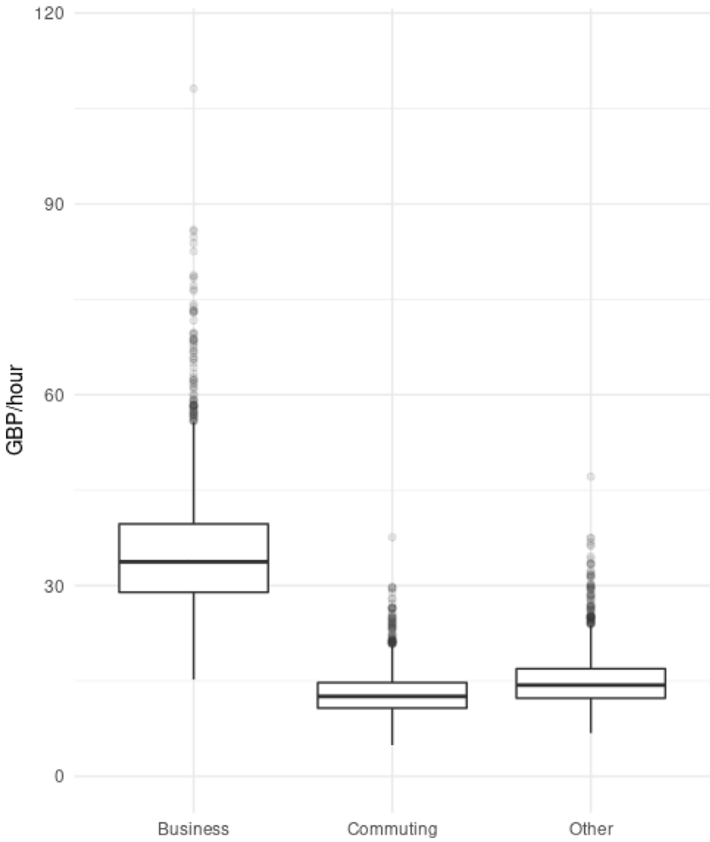

Fig. 7. Box plots of GPS-based Rail VTTs per purpose.

conditions, hilliness/slope and the type of land uses both around home locations and also along the route to the destination of each choice task.

The findings in the current study can provide practitioners and policy makers with additional confidence when it comes to using new emerging data sources for future nationwide VTT studies. The small differences across the overall VTT estimates, regardless of distance bands, showcase that RP data captured through new emerging data collection methods – GPS in this case – can provide behaviourally reasonable VTT estimates that are in line with the official SP-based values currently used in appraisal. Of course, a reader may ask why RP data should be used if the results are no different from SP data. The simple answer to this question is that RP data provides the truth, and the fact that the findings in this case are in line with the SP results thus arguably also serves as a validation of SP rather than RP and a result of the rigorous work of the researchers involved in the official study. Furthermore, our results are achieved using a much smaller sample size during estimation, compared to the SP study, which can lead to a reduced cost or more frequently performed surveys in general. Performing a large scale GPS-based RP study at the country level will result in a significantly more accurate representation of individual mobility behaviour, capturing choices over a large number of real-life scenarios, independent of the researcher's assumptions, while also resulting in less fatigue for the respondents.

This study comes at a time where ubiquitous sensing data sources are steadily gaining ground in transport research and provides empirical evidence for their further adoption into the field of practice. Nonetheless, more studies are required offering similar practical applications, even in small sample sizes, before we can see a departure from the current state of practice that has been dominant over the last several decades.

CRediT authorship contribution statement

Panagiotis Tsoleridis: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **Charisma F. Choudhury:** Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing. **Stephane Hess:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing.

Acknowledgements

The current research was funded by the Advanced Quantitative Methods (AQM) scholarship of the Economic and Social Research Council (ESRC), UK. Charisma Choudhury acknowledges the financial support of her UKRI Future Leader Fellowship MR/T020423/1-NEXUS. Stephane Hess acknowledges the financial support by the European Research Council through the consolidator Grant 615596-DECISIONS.

References

Accent and Hague Consulting Group, 1996. The Value of Travel Time on UK. Department of Environment, Transport and Regions, London.

Arifin, Z.N., Axhausen, K.W., 2012. Investigating commute mode and route choice variabilities in jakarta using multi-day GPS data. Int. J. Technol. 3 (1), http://dx.doi.org/10.14716/ijtech.v3i1.1087.

Axhausen, K.W., Koenig, A., Abay, G., Bates, J., Bierlaire, M., 2006. Swiss Value of Travel Time Savings. ETH Zürich.

Batley, R., Bates, J., Bliemer, M., Börjesson, M., Bourdon, J., Cabral, M.O., Chintakayala, P.K., Choudhury, C.F., Daly, A., Dekker, T., Drivyla, E., Fowkes, T., Hess, S., Heywood, C., Johnson, D., Laird, J., Mackie, P., Parkin, J., Sanders, S., Sheldon, R., Wardman, M., Worsley, T., 2019. New appraisal values of travel time saving and reliability in Great Britain. Transportation 46, 583–621. http://dx.doi.org/10.1007/s11116-017-9798-7.

Becker, G.S., 1965. A theory of the allocation of time. Econ. J. 75 (299), 493–517.

Beesley, M.E., 1965. The value of time spent in travelling: Some new evidence. Economica 32, 174–185. http://dx.doi.org/10.2307/2552547.

Ben-Akiva, M., Lerman, S., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge, Massachusetts.

Bliemer, M.C.J., Rose, J.M., 2009. Designing stated choice experiments: The state of the art. In: The Expanding Sphere of Travel Behaviour Research, Selected Papers from the 11th International Conference on Travel Behaviour Research. pp. 495–498.

Bohte, W., Maat, K., 2009. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. Transp. Res. C 17 (3), 285–297. http://dx.doi.org/10.1016/j.trc.2008.11.004.

Börjesson, M., Eliasson, J., 2014. Experiences from the Swedish value of time study. Transp. Res. A 59, 144–158. http://dx.doi.org/10.1016/j.tra.2013.10.022.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations. J. R. Stat. Soc. 26 (2), 211–252.

Bradley, M., Daly, A., 1994. Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. Transportation 21 (2), 167–184. http://dx.doi.org/10.1007/BF01098791.

Brownstone, D., Small, K.A., 2005. Valuing time and reliability: Assessing the evidence from road pricing demonstrations. Transp. Res. A 39, 279–293. http://dx.doi.org/10.1016/j.tra.2004.11.001.

Bwambale, A., Choudhury, C.F., Hess, S., 2019. Modelling long-distance route choice using mobile phone call detail record data: A case study of Senegal. Transportmetr. A: Transp. Sci. 15 (2), 1543–1568. http://dx.doi.org/10.1080/23249935.2019.1611970.

Calastri, C., Hess, S., Choudhury, C.F., Daly, A., Gabrielli, L., 2018. Mode choice with latent availability and consideration: Theory and a case study. Transp. Res. B.

Calastri, C., Crastes dit Sourd, R., Hess, S., 2020. We want it all: Experiences from a survey seeking to capture social network structures, lifetime events and short-term travel and activity planning. Transportation 47, 175–201. http://dx.doi.org/10.1007/s11116-018-9858-7.

Cascetta, E., Papola, A., 2001. Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. Transp. Res. C 9 (4), 249–263.

Champ, P.A., Welsh, M.P., 2006. Survey methodologies for stated-choice studies. In: Kanninen, B.J. (Ed.), Valuing Environmental Amenities using Stated Choice Studies. In: The Economics of Non-Market Goods and Resources, Springer, Dodrecht.

Cherchi, E., Ortuzar, J.D., 2006. Use of mixed revealed-preference and stated-preference models with nonlinear effects in forecasting. Transp. Res. Rec.: J. Transp. Res. Board 1977 (1), 27–34. http://dx.doi.org/10.1177/0361198106197700104.

Daly, A., 2010. Cost Damping in Travel Demand Models - Report of a Study for the Department for Transport. RAND Europe.

Daly, A., Hess, S., de Jong, G., 2012. Calculating errors for measures derived from choice modelling estimates. Transp. Res. Part B: Methodological 46, 333–341.

Daly, A., Tsang, F., Rohr, C., 2014. The value of small time savings for non-business travel. J. Transp. Econ. Policy 48 (2), 205–218.

Daly, A., Zachary, S., 1975. Commuters' Values of Time. Local Government Operational Research Unit, Reading.

De Borger, B., Fosgerau, M., 2008. The trade-off between money and travel time: A test of the theory of reference-dependent preferences. J. Urban Econ. 64 (1), 101–115. http://dx.doi.org/10.1016/j.jue.2007.09.001.

Department for Transport, 2014. Values of Time and Vehicle Operating Costs TAG Unit 3.5.6. Department for Transport.

Department for Transport, 2015. Provision of Market Research for Value of Travel Time Savings and Reliability - Phase 2 Report.

DeSerpa, A.C., 1971. A theory of the economics of time. Econ. J. 81 (324), 828–846. http://dx.doi.org/10.2307/2230320.

Domencich, T.A., McFadden, D., 1975. Urban Travel Demand: A Behavioral Analysis. North-Holland Publishing Company Limited, Oxford, England.

Ebrahimpour, Z., Wan, W., Garcia, J.L.V., Cervantes, O., Hou, L., 2020. Analyzing social-geographic human mobility patterns using large-scale social media data. Int. J. Geo-Inf. 9 (125), http://dx.doi.org/10.3390/ijgi9020125.

Ehreke, I., Hess, S., Weis, C., Axhausen, K.W., 2015. Reliability in the German value of time study. Transp. Res. Rec. 2495, 14–22.

Essadeq, I., Janik, T., 2021. Use of Mobile Telecommunication Data in Transport Modelling: A French Case Study. OECD Publishing, International Transport Forum Discussion Papers, Paris.

Forrest, T.L., Pearson, D.F., 2005. Comparison of trip determination methods in household travel surveys enhanced by a global positioning system. Transp. Res. Rec. J. Transp. Res. Board 1917 (1), 63–71. http://dx.doi.org/10.1177/0361198105191700108.

Fosgerau, M., 2006. Investigating the distribution of the value of travel time savings. Transp. Res. B 40 (8), 688–707. http://dx.doi.org/10.1016/j.trb.2005.09.007.

Fosgerau, M., Bierlaire, M., 2009. Discrete choice models with multiplicative error terms. Transp. Res. B 43 (5), 494–505. http://dx.doi.org/10.1016/j.trb.2008.10.004.

Fosgerau, M., Jensen, T.L., 2003. Economic appraisal methodology - controversial issues and danish choices. In: European Transport Conference, 4-6 October. Strasbourg, France.

Gallotti, R., Bazzani, A., Rambaldi, S., 2015. Understanding the variability of daily travel-time expenditures using GPS trajectory data. EPJ Data Sci. 4 (18), http://dx.doi.org/10.1140/epjds/s13688-015-0055-z.

Gaudry, M.J.I., Dagenais, M.G., 1979. The dogit model. Transp. Res. B 13B (2), 105–111.

Gaudry, M.J.I., Jara-Diaz, S.R., Ortuzar, J.D., 1989. Value of time sensitivity to model specification. Transp. Res. B 23 (2), 151–158. http://dx.doi.org/10.1016/0191-2615(89)90038-6.

Grant-Muller, S.M., Abdelrazek, M., Budnitz, H., Cottrill, C.D., Crawford, F., Choudhury, C.F., Cunningham, T., Harrison, G., Hodgson, F.C., Hong, J., Martin, A., O'Brien, O., Papaix, C., Tsoleridis, P., 2021. In: Vickerman, R. (Ed.), Technology Enabled Data for Sustainable Transport Policy. In: International Encyclopedia of Transportation, Elsevier.

Haghani, M., Bliemer, M.C.J., Rose, J.M., Oppewal, H., Lancsar, E., 2021. Hypothetical bias in stated choice experiments: Part I. Integrative synthesis of empirical evidence and conceptualisation of external validity.

Halse, A.H., Flügel, S., Kouwenhoven, M., de Jong, G., Sundfør, H.B., Hulleberg, N., Jordbakke, G.N., Lindhjem, H., 2022. A minute of your time: The impact of survey recruitment method and interview location on the value of travel time. Transportation http://dx.doi.org/10.1007/s11116-022-10287-8.
Hensher, D.A., Greene, W.H., 2010. Non-attendance and dual processing of common-metric attributes in choice analysis: A latent class specification. Empir. Econ. 39 (2), 413–426. http://dx.doi.org/10.1007/s00181-009-0310-x.
Hess, S., Daly, A., Börjesson, M., 2020. A critical appraisal of the use of simple time-money trade- offs for appraisal value of travel time measures. Transportation 47, 1541–1570. http://dx.doi.org/10.1007/s11116-020-10097-w.
Hess, S., Daly, A., Dekker, T., Cabral, M.O., Batley, R., 2017. A framework for capturing heterogeneity, heteroscedasticity, non-linearity, reference dependence and design artefacts in value of time research. Transp. Res. B 96, 126–149. http://dx.doi.org/10.1016/j.trb.2016.11.002.
Hess, S., Hensher, D., Daly, A.J., 2012. Not bored yet - revisiting respondent fatigue in stated choice experiment. Transp. Res. A 46 (3), 626–644. http://dx.doi.org/10.1016/j.tra.2011.11.008.
Hess, S., Quddus, M., Rieser-Schüssler, N., Daly, A., 2015. Developing advanced route choice models for heavy goods vehicles using GPS data. Transp. Res. E 77, 29–44. http://dx.doi.org/10.1016/j.tre.2015.01.010.
Hess, S., Train, K.E., Polak, J.W., 2006. On the use of a modified latin hypercube sampling (MLHS) method in the estimation of a mixed logit model for vehicle choice. Transp. Res. B 40 (2), 147–163. http://dx.doi.org/10.1016/j.trb.2004.10.005.
Huang, Y., Gao, L., Ni, A., Liu, X., 2021. Analysis of travel mode choice and trip chain pattern relationships based on multi-day GPS data: A case study in Shanghai, China. J. Transp. Geogr. 93, 103070. http://dx.doi.org/10.1016/j.jtrangeo.2021.103070.
Huang, A., Levinson, D., 2015. Axis of travel: Modeling non-work destination choice with GPS data. Transp. Res. C 58 (B), 208–223. http://dx.doi.org/10.1016/j.trc.2015.03.022.
Huang, A., Levinson, D., 2017. A model of two-destination choice in trip chains with GPS data. J. Choice Model. 24, 51–62. http://dx.doi.org/10.1016/j.jocm.2016.04.002.
Kang, M.J., Camerer, C.F., 2013. fMRI evidence of a hot-cold empathy gap in hypothetical and real aversive choices. Front. Neurosci. 7 (104), http://dx.doi.org/10.3389/fnins.2013.00104.
Kelly, P., Krenn, P., Titze, S., Stopher, P., Foster, C., 2013. Quantifying the difference between self-reported and global positioning systems-measured journey durations: A systematic review. Transp. Rev. 33 (4), 443–459. http://dx.doi.org/10.1080/01441647.2013.815288.
Koppelman, F.S., 1981. Non-linear utility functions in models of travel choice behaviour. Transportation 10 (2), 127–146. http://dx.doi.org/10.1007/bf00165262.
Kouwenhoven, M., Jong, G.C. de, Koster, P., Berg, V.A.C. van der, Verhoef, E.T., Bates, J., Warffemius, P.M.J., 2014. New values of time and reliability in passenger transport in The Netherlands. Res. Transp. Econ. 47 (1), 37–49. http://dx.doi.org/10.1016/j.retrec.2014.09.017.
Krcal, O., Peer, S., Stanek, R., Karlinova, B., 2019. Real consequences matter: Why hypothetical biases in the valuation of time persist even in controlled lab experiments. Econ. Transp. 20 (100138), http://dx.doi.org/10.1016/j.ecotra.2019.100138.
Li, L., Adamowicz, W., Swait, J., 2015. The effect of choice set misspecification on welfare measures in random utility models. Resour. Energy Econ. 42, 71–92. http://dx.doi.org/10.1016/j.reseneeco.2015.07.001.
Li, H., Guensler, R., Ogle, J., 2005. Analysis of morning commute route choice patterns using global positioning system–based vehicle activity data. Transp. Res. Rec.: J. Transp. Res. Board 1926, 162–170. http://dx.doi.org/10.1177/0361198105192600119.
Li, Z., Hensher, D.A., Ho, C., 2020. An empirical investigation of values of travel time savings from stated preference data and revealed preference data. Transp. Lett. Int. J. Transp. Res. 12 (3), 166–171. http://dx.doi.org/10.1080/19427867.2018.1546806.
Loewenstein, G., 2005. Hot-cold empathy gaps and medical decision making. Health Psychol. 24 (4S), 49–56. http://dx.doi.org/10.1037/0278-6133.24.4.S49.
Lorenzo Varela, Juan Manuel, Börjesson, Maria, Daly, Andrew, 2018. Public transport: One mode or several?. Transp. Res. Part A Policy and Practice (ISSN: 0965-8564) 113, 137–156. http://dx.doi.org/10.1016/j.tra.2018.03.018.
Louviere, J.J., Hensher, D.A., 2001. Combining sources of preference data. In: Hensher, D.A. (Ed.), Travel Behaviour Research: The Leading Edge. Pergamon Press, Oxford, pp. 125–144.
Lu, H., Fowkes, A.S., Wardman, M., 2008. Amending the incentive for strategic bias in stated preference studies: Case study in users' valuation of rolling stock. Transp. Res. Rec.: J. Transp. Res. Board 2049 (2049), 128–135. http://dx.doi.org/10.3141/2049-15.
Mackie, P.J., Jara-Diaz, S., Fowkes, T., 2001. The value of travel time savings in evaluation. Transp. Res. E 37 (2–3), 91–106. http://dx.doi.org/10.1016/S1366-5545(00)00013-2.
Mackie, P.J., Wardman, M., Fowkes, A.S., Whelan, G.A., Nellthorp, J., Bates, J., 2003. Values of Travel Time Savings in the UK. Report to Department for Transport. Institute for Transport Studies, University of Leeds.
Manski, C.F., 1977. The structure of random utility models. Theory and Decision 8 (3), 229–254.
Marchal, P., Madre, J.L., Yuan, S., 2011. Postprocessing procedures for person-based GPS data collected in the french national travel survey 2007–2008. Transp. Res. Rec.: J. Transp. Res. Board 2246, 47–54. http://dx.doi.org/10.3141/2246-07.
Marschak, J., 1960. Binary choice constraints on random utility indications. In: Arrow, K. (Ed.), Stanford Symposium on Mathematic Models in the Social Sciences. Stanford University Press, Stanford, CA, pp. 312–329.
Martinez, F., Aguila, F., Hurtubia, R., 2009. The constrained multinomial logit: A semi-compensatory choice model. Transp. Res. B 43, 365–377.
McFadden, D., 1973. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in Econometrics. Academic Press, New York, pp. 105–142.
McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. J. Appl. Econometrics 15, 447–470.
Meginnis, K., Burton, M., Chan, R., Rigby, D., 2018. Strategic bias in discrete choice experiments. J. Environ. Econ. Manag. 109 (1), http://dx.doi.org/10.1016/j.jeem.2018.08.010.
Montini, L., Antoniou, C., Axhausen, K.W., 2017. Route and mode choice models using GPS data. In: Transportation Research Board 96th Annual Meeting. 8-12 January 2017. Washington D.C., United States.
MVA, ITS Leeds and TSU Oxford, 1987. The value of travel time savings, Newbury Berkshire. Policy J..
Nielsen, O.A., 2004. Behavioral responses to road pricing schemes: Description of the Danish AKTA experiment. Intell. Transp. Syst. 8 (4), 233–251. http://dx.doi.org/10.1080/15472450490495579.
Schuessler, N., Axhausen, K.W., 2009. Processing raw data from global positioning systems without additional information. Transp. Res. Rec.: J. Transp. Res. Board 2105 (2105), 28–36. http://dx.doi.org/10.3141/2105-04.
Shires, J.D., de Jong, G.C., 2009. An international meta-analysis of values of travel time savings. Eval. Program Plan. 32 (4), 315–325. http://dx.doi.org/10.1016/j.evalprogplan.2009.06.010.
Small, K.A., 2012. Valuation of travel time. Econ. Transp. 1 (1–2), 2–14. http://dx.doi.org/10.1016/j.ecotra.2012.09.002.
Stopher, P., Jiang, Q., FitzGerald, C., 2005. Processing GPS Data from Travel Surveys. Toronto.
Swait, J., Ben-Akiva, M., 1986. Constraints on individual travel behavior in a Brazilian city. Transp. Res. Rec. 1085.
Swait, J., Ben-Akiva, M., 1987. Incorporating random constraints in discrete models of choice set generation. Transp. Res. B 21, 91–102.
Thiene, M., Swait, J., Scarpa, R., 2017. Choice set formation for outdoor destinations : The role of motivations and preference discrimination in site selection for the management of public expenditures on protected areas. J. Environ. Econ. Manag. 81, 152–173. http://dx.doi.org/10.1016/j.jeem.2016.08.002.
Train, K., 2009. Discrete Choice Methods with Simulation. Cambridge University Press, Cambridge, Massachusetts.
Tsoleridis, P., Choudhury, C.F., Hess, S., 2021. Utilising activity space concepts to sampling of alternatives for mode and destination choice modelling of discretionary activities. J. Choice Model. 42 (1), 100336. http://dx.doi.org/10.1016/j.jocm.2021.100336.

Tversky, A., Kahneman, D., 1991. Loss aversion in riskless choice: A reference- dependent model. Q. J. Econ. 106 (4), 1039–1061.

Vij, A., Shankari, K., 2015. When is big data big enough? Implications of using GPS-based surveys for travel demand analysis. Transp. Res. C 56 (3), 446–462. http://dx.doi.org/10.1016/j.trc.2015.04.025.

Vrtic, M., Schuessler, N., Erath, A., Axhausen, K.W., 2010. The impacts of road pricing on route and mode choice behaviour. J. Choice Model. 3 (1), 109–126. http://dx.doi.org/10.1016/S1755-5345(13)70031-9.

Wardman, M., Chintakayala, V.P.K., de Jong, G.C., 2016. Values of travel time in Europe: Review and meta-analysis. Transp. Res. A 94, 93–111. http://dx.doi.org/10.1016/j.tra.2016.08.019.

Williams, H.C.W.L., 1977. On the formation of travel demand models and economic evaluation measures of user benefit. Environ. Plan. A: Econ. Space 9 (3), 285–344.

Wolf, J., 2004. Defining GPS and its capabilities. In: Hensher, D.A., Stopher, P. (Eds.), Handbook of Transport Geography and Spatial Systems. In: Handbooks in Transport, Elsevier, Amsterdam.

Yan, X.Y., Zhou, T., 2019. Destination choice game: A spatial interaction theory on human mobility. Sci. Rep. 9 (1), http://dx.doi.org/10.1038/s41598-019-46026-w.

Zhang, G., Laird, J.J., 2014. Local values of travel time savings - should we forget about them? In: The 19th International Conference of Hong Kong Society for Transportation Studies. Hong Kong, 13-15 December, 2014.