Taylor & Francis
Taylor & Francis Group

Transportmetrica A: Transport Science

ISSN: 2324-9935 (Print) 2324-9943 (Online) Journal homepage: www.tandfonline.com/journals/ttra21

Imputing missing trip purposes in smartphone-based travel surveys using Bayesian posterior analysis

Azam Ali, Stephane Hess, Thijs Dekker & Charisma F. Choudhury

To cite this article: Azam Ali, Stephane Hess, Thijs Dekker & Charisma F. Choudhury (26 Apr 2026): Imputing missing trip purposes in smartphone-based travel surveys using Bayesian posterior analysis, Transportmetrica A: Transport Science, DOI: 10.1080/23249935.2026.2660827

To link to this article: https://doi.org/10.1080/23249935.2026.2660827

Published online: 26 Apr 2026.

TRANSPORTMETRICA A: TRANSPORT SCIENCE
https://doi.org/10.1080/23249935.2026.2660827

Imputing missing trip purposes in smartphone-based travel surveys using Bayesian posterior analysis

Azam Ali, Stephane Hess, Thijs Dekker and Charisma F. Choudhury

Choice Modelling Centre & Institute for Transport Studies, University of Leeds, Leeds, UK

ABSTRACT

Smartphone-based travel surveys are gaining popularity for collecting trip diaries, where much of the travel information (i.e. travel modes, purposes) is *inferred* and the participants are asked to *verify* or *correct* the inferred information. While this allows researchers to collect high-resolution mobility data from larger samples with minimal user input, in many cases, the participants do not verify all of their trips, resulting in a considerable amount of untagged or missing data. To better impute missing trip purposes for individuals with partially tagged trips, we propose using posterior analysis within a mixed multinomial logit (MMNL) modelling framework. Posterior analysis leverages Bayes' rule to predict missing data by conditioning on the individual's previously observed trip purposes. Using a two-week-long dataset collected in the UK, we find this approach provides a better predictive performance on a testing dataset in terms of predicting market shares and the probability of choosing the correct alternative when compared to not using posteriors in MMNL or using a standard multinomial logit model. We further investigate the impact of using the imputed trip purposes in a mode choice model. Our findings demonstrate that imputing trip purposes using MMNL with posterior analysis achieves a better model fit and values of travel time compared to mode choice models that use (a) other methods of imputing trip purposes, and (b) ignore trip purpose as an explanatory variable. This highlights the advantages of our imputation strategy, as well as the benefits of incorporating imputed trip purpose data in developing travel behaviour models.

ARTICLE HISTORY
Received 18 September 2025
Accepted 10 April 2026

KEYWORDS
Missing or untagged data; posterior analysis; trip purpose; mixed multinomial logit; value of travel time

1. Introduction

With the ubiquitous use of smartphones, many travel surveys are now carried out using smartphone-based travel survey applications (Calastri, Crastes dit Sourd, and Hess 2020; Cottrill et al. 2013; Molloy et al. 2022; Tabasi et al. 2024; Wang et al. 2019; Winkler, Meister, and Axhausen 2024). Typically, in these survey applications, travel details such as modes and purposes are automatically *inferred* from the GPS traces, but to ensure that the inferred trip details are correct, participants are generally asked to *verify* or *correct* the inferred information. There are also some cases where the respondents need to specify the modes and

CONTACT Azam Ali ts19aa@leeds.ac.uk Choice Modelling Centre & Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, UK

purposes from a drop-down menu for all trips, i.e. with no automatic inferring (McCool et al. 2021; Tabasi et al. 2024). Even though smartphone applications require minimal user input compared to traditional paper-based travel surveys, in many cases, the participants do not tag, verify, or correct all of their trips. This leads to a large proportion of untagged trips or trips with missing details. For example, 50% of the users who downloaded the survey application for a time-use and travel survey in Switzerland did not fully tag their activities (Winkler, Meister, and Axhausen 2024), half of the trips details were left unverified in a survey in Toronto (Faghih Imani et al. 2020), and nearly a third of the trip purposes were left untagged in a travel survey in Trondheim, Norway (Tørset and Svaboe 2020). Further, in long-duration panel surveys (over 8 weeks), tagging trips is encouraged but not made compulsory in order to reduce response burden (Heinonen et al. 2024; Meister et al. 2025; Molloy et al. 2022). To encourage higher levels of user participation and data completeness, most surveys offer incentives and set minimum thresholds for tagging. For instance, Tabasi et al. (2024) required respondents to tag at least 75% of their trips or activities, while Cottrill et al. (2013) required participants to tag at least 5 out of 14 days for respondents to receive monetary incentives. All of these aspects suggest that smartphone-based travel surveys often yield datasets with substantial amounts of missing or untagged trip details, primarily the modes and purposes. In this paper, we explore two interrelated aspects, i.e. methods for improving the imputation of missing travel information by exploiting the panel nature of the dataset, and the implications of using such imputed information in travel behaviour models.

While travel modes can typically be inferred with high accuracy levels using speed and acceleration profiles from the GPS traces, inferring or imputing trip purposes is considerably more challenging, as it relies heavily on contextual information about the user and the trip origins and destinations (Servizi et al. 2021). A large body of research has focussed on predicting trip purposes, often using geospatial context data such as points of interest data (Servizi et al. 2021). However, one of the key aspects which has been overlooked in the previous studies is the fact that the untagged trip purposes (which manifest as missing information in the data) are of two categories, i.e. there are some users who tag or verify some of their trips (but not all), while the rest do not tag or verify any of their trips. The former category opens up the possibility of using the tagged or verified part of the data of each *individual* for better predicting their missing information. To better leverage individuals' observed choices for predicting missing trip purposes, we propose using posterior analysis within the mixed multinomial logit (MMNL) modelling framework. MMNL is one of the most flexible frameworks in modelling discrete choices across various disciplines. These models are typically used to capture random or unobserved heterogeneity in preferences across individuals by specifying a mixing distribution. This yields sample level insights into the distribution of unobserved effects across individuals. Posterior analysis builds on this by applying Bayes' rule to obtain the most likely location of an individual's taste sensitivities on the sample level distribution, conditional on their observed choices. This allows researchers to more precisely infer each individual's location, i.e. with tighter confidence intervals, compared to the sample level distribution alone (Train 2009).

Posterior analysis has previously been used to obtain insights into the taste sensitivities of individuals within the sample (Campbell, Hutchinson, and Scarpa 2006; Hensher, Greene, and Rose 2006; Richter and Pollitt 2018; Sillano and de Dios Ortúzar 2005) and for prediction purposes (Danaf et al. 2019; Dumont, Giergiczny, and Hess 2015; Krueger et al. 2021; D. Revelt and Train 1999; Song et al. 2018). Past work using posterior analysis has tended to rely on stated preference data with a limited number of choices per individual on which to condition the distribution. For instance, Danaf et al. (2019) model 8 stated choice per individual which are used to predict the mode choice, Dumont, Giergiczny, and Hess (2015) compare prediction accuracy of holiday destination location with a maximum of 36 stated choices, and Song et al. (2018) use revealed preference data of 4 tours in a day in the context of mode choice. The applicability of the framework on a longer panel of choices, such as multi-day GPS travel diary data, remains unexplored. Further, mixed multinomial logit models are typically limited to modelling behavioural decisions such as mode choices, while the applicability of the modelling framework to impute partially missing trip information, specifically trip purposes, remains less investigated.

In terms of modelling and predicting trip purposes, most studies employ machine learning (ML) algorithms such as gradient boosting trees and neural networks (Cui et al. 2018; Ectors et al. 2017; Gao, Molloy, and Axhausen 2021; Y. Liu, Miller, and Habib 2023; Xiao, Juan, and Zhang 2016; Yazdizadeh, Patterson, and Farooq 2019). There are also a few studies that use econometric or statistical models such as multinomial logit (MNL), nested logit, and MMNL (Ermagun et al. 2017; Y. Liu, Miller, and Habib 2023; Oliveira et al. 2014) to predict trip purposes; however, they do not use posterior analysis. Machine learning (ML) techniques are often preferred for modelling trip purposes because they provide a data-driven approach capable of capturing potentially non-linear and complex relationships between explanatory variables and observed trip purposes. In contrast, econometric models are typically favoured for their interpretability. There are some studies that compare the two modelling frameworks, i.e. econometric and machine learning models, specifically for predicting trip purposes. For instance, Ermagun et al. (2017) report that prediction accuracy is better for ML models compared to econometric models for modelling trip purposes. On the other hand, Y. Liu, Miller, and Habib (2023) note that econometric models better predict market shares compared to ML models. Nevertheless, there is substantial evidence in the transport literature that finds that there is no significant difference in the out-of-sample predictive performance between econometric models and ML models in terms of probabilistic goodness of fit metrics such as log-likelihood and predicting market shares (Ali, Kalatian, and Choudhury 2023; Hillel 2019; Zhao et al. 2020). Accordingly, we restrict our analysis to econometric models, namely the multinomial logit and mixed multinomial logit model, for predicting missing trip purposes and evaluate the benefits of incorporating posterior analysis in MMNL.

Further, although prior research assesses the feasibility of utilising untagged datasets to develop travel behaviour models based on accuracy levels of trip purpose and mode inference algorithms on testing or validation datasets (Harding et al. 2021; Marra et al. 2019; Yazdizadeh, Patterson, and Farooq 2019), it should be recognised that reporting prediction or inference accuracies is of limited practical value for transport practitioners. The ultimate goal of imputing trip purposes should be to assess whether such imputations enhance the estimation of downstream travel behaviour models. Trip purposes are valuable explanatory variables in mode and route choice models, and consequently in valuation and appraisal, where values of travel times are differentiated by purpose (Hess et al. 2017; Wardman, Chintakayala, and de Jong 2016). Failing to account for these preference differences across purposes can result in significant biases in demand forecasts and welfare analysis.

Further, trip purposes are essential components in understanding daily travel and activity patterns, and act as dependent variables in various models such as trip generation and time-use models (Arentze, Ettema, and Timmermans 2013; Nurul Habib and Miller 2009). Vij and Shankari (2015) is a notable exception, as they investigate the impact of utilising inferred or untagged modes in a mode choice model and highlight that there is a trade-off between inference accuracy and dataset size. Their findings suggest that even with a high mode inference accuracy of 95%, there is a 13% bias in predicting the value of travel time, indicating limited benefits of utilising inferred untagged datasets. However, there has been limited research evaluating the impact and benefits of using imputed trip purposes in travel behaviour models, especially when incorporating untagged data alongside tagged data from smartphone-based travel surveys. One of the most important uses of trip purpose information is in mode choice decisions and in the estimation of the values of travel time. Hence, in this study, we also check if imputing trip purposes using MMNL with posteriors leads to a better model fit and more reliable values of travel time (VTT) compared to mode choice models that do not use trip purpose as an explanatory variable or use other methods of imputing trip purposes. It may be noted that our context differs fundamentally from that of Vij and Shankari (2015), who focus on imputing a discrete dependent variable, i.e. modes in a mode choice model. In contrast, our study evaluates the implications of imputing a discrete independent variable, i.e. trip purpose, in a mode choice model. The main research questions addressed in this study are as follows:

(1) To what extent does posterior analysis in a mixed multinomial logit model improve the predictive performance of a trip purpose prediction model in a multi-day panel dataset compared to other methods of imputation?
(2) What is the implication of imputing trip purposes using mixed multinomial logit model with posterior analysis, in terms of model fit and values of travel time, when subsequently used in a downstream mode choice model?

The remainder of this paper is divided as follows; where Section 2 provides a brief literature review of posterior analysis and trip purpose prediction models, Section 3 explains the methodology used to carry out the study, Section 4 describes the data, Section 5 the results and discussion. The conclusions are summarised in Section 6.

2. Literature review

In this section, a brief overview is provided of the use of posterior analysis in the discrete choice modelling literature. This is followed by a review of how trip purposes have been modelled in recent studies.

2.1. Use of posterior analysis in transport literature

Posterior analysis has first been used in the transport or discrete choice modelling literature context by D. Revelt and Train (1999) inspired by the hierarchical Bayesian estimation method that relies on taking draws of individual level parameters from sample level parameters in the case of random coefficients, which is later explained in Section 3.

There are several studies that examine the impact of incorporating posterior analysis on the predictive performance of mixed multinomial logit (MMNL) models. D. Revelt and Train (1999) model customer's choice of electricity suppliers using a mixed multinomial logit, utilising a stated preference dataset that consists of 8 to 12 choice tasks, where the last choice task is set aside for testing. D. Revelt and Train (1999) find that the average probability of the chosen alternative (APCA) increases from 0.35 to 0.50 by making use of posteriors. Further, the APCA improves by 0.25 for nearly three-quarters of the individuals. For the rest of the individuals, the average probability decreases by 0.11. One of the possible reasons for some individuals having a decrease in the APCA is that in stated preference surveys, the choice tasks are designed to have new trade-offs in each scenario; therefore, the taste sensitivities for all individuals may have not been fully captured. This aspect would be interesting to examine in the context of this study, as we make use of multi-day revealed preference data. Furthermore, a greater increase in APCA might be observed as real-world journey purposes may be more habitual.

Song et al. (2018) use posterior analysis to generate a personalised list of potential mode choices, i.e. the mode choice menu in a smartphone-based travel survey application. Using mode choice data from 246 users, different models are estimated where the training dataset contains 2, 3 or 4 observed tours per individual. A hit rate is calculated if the observed mode choice is present in the predicted mode choice menu. Results indicate that the hit rate increases by making use of posteriors and also with an increase in the number of observations per individual in the training dataset. Song et al. (2018) further proposes the use of posteriors to generate personalised recommendations for mode choice menus for individuals in smartphone-based travel surveys. Building on this idea, the present study similarly exploits posterior analysis, but applies it to a different problem, i.e. imputing untagged trip purposes in smartphone-based travel surveys rather than generating personalised mode choice menus.

Another branch of literature examines the impact of updating posteriors dynamically and the benefits of capturing intra-personal heterogeneity. Danaf et al. (2019) develop an MMNL framework that updates individual-level parameters as new data arrives, capturing both inter-and intra-personal heterogeneity as opposed to the approach by Train (2009) where the posteriors are calculated using all the observed data statically. The proposed framework is tested to predict the mode choice on the Swissmetro dataset, where 8 out of the 9 stated preferences are in the training dataset and the last choice scenario is retained for testing. It is observed that the best predictive performance is observed when all of the observed data is used in the estimation of the posteriors rather than the dynamic updating procedure. Nonetheless, the results indicate that by making use of posteriors in MMNL, the average probability of the chosen alternative increases from 0.509 to 0.717 (Danaf et al. 2019). Krueger et al. (2021), on the other hand, check the predictive performance of a mixed multinomial logit model with inter-personal heterogeneity compared with both inter-and intra-personal heterogeneity. The study finds that there is no significant improvement by making use of both inter-and intra-personal heterogeneity compared to using interpersonal heterogeneity only on a synthetic dataset and a stated preference dataset, which contains 7 choice scenarios.

While Danaf et al. (2019), D. Revelt and Train (1999), Song et al. (2018), and Krueger et al. (2021) utilise data with 9, 12, 4, and 7 observations; only Dumont, Giergiczny, and Hess (2015) assesses the impact of utilising a longer panel of stated preference data.

Dumont, Giergiczny, and Hess (2015) compare the predictive performance of mixed multinomial logit with individual-level models. Using stated preference data of holiday destination choice, different models are estimated where out of 48 stated choices, 12, 24, and 36 choices per individual are used for the training dataset and the rest are set aside for testing. It is observed that when 12 observations are used in the training dataset, the predictive performance on the testing dataset of the MMNL model with posteriors is lower compared to the model without posteriors. However, with 24 and 36 observations per individual in the training dataset, the use of posteriors leads to an 11% and 22.8% increase in log-likelihood on the testing dataset, respectively, compared to not using posteriors. Despite these insights, it remains unclear how posterior analysis performs when applied to multi-day revealed preference (RP) data.

Further, several studies utilise posterior analysis to examine willingness-to-pay (WTP) measures at an individual level (Campbell, Hutchinson, and Scarpa 2006; Hensher, Greene, and Rose 2006; Hess 2007; Richter and Pollitt 2018; Sillano and de Dios Ortúzar 2005). All of these studies rely on stated preference data and find that the posterior estimates exhibit a narrower range compared to sample level estimates. Additionally, D. A. Revelt (1999) performs a cluster analysis on the posteriors to gain deeper insight into individual behaviour. Therefore, using posterior analysis within a mixed multinomial logit modelling framework for imputing partially missing trip purpose data in a multi-day GPS travel diary dataset remains unexplored in the literature. Overall, while posterior analysis has been widely applied in stated preference contexts for prediction and WTP analysis, its use in MMNL models to impute partially missing trip purpose data in multi-day GPS-based RP datasets remains unexplored. This study addresses this gap by leveraging posterior information to improve inference for imputing trip purposes in a real-world, multi-day setting.

2.2. Modelling trip purposes

Smartphone-based travel survey applications typically record timestamped GPS traces of users. Based on the spatial and temporal data, the trip details such as trip stops, modes and purposes are inferred using various algorithms. For a recent review of methods for extracting trip information from smartphone-based travel survey data, readers are referred to Servizi et al. (2021). Beyond smartphone-based GPS data, other big data sources, such as call detail records (CDRs) and public transport smartcard data, have also been increasingly used to infer trip purposes or activity patterns as detailed in a recent review (Fu et al. 2025). In this section, we mention algorithms and variables used to model trip purposes.

In terms of the modelling framework used, most of the researchers use machine learning algorithms such as decision tree (Deng and Ji 2010; Lu, Zhu, and Zhang 2012; Oliveira et al. 2014), random forest (Gao, Molloy, and Axhausen 2021; Gong, Kanamori, and Yamamoto 2018; Montini et al. 2014; Yazdizadeh, Patterson, and Farooq 2019), gradient boosting trees (Nair et al. 2019; Zahroh et al. 2025), and neural networks (Ali et al. 2026; Cui et al. 2018; Xiao, Juan, and Zhang 2016; Zahroh et al. 2025). There are some studies which use econometric or statistical techniques such as multinomial logit model, nested logit, and mixed multinomial logit models (Ermagun et al. 2017; Hossain and Habib 2021; Y. Liu, Miller, and Habib 2023). However, these studies do not propose the use of posterior analysis to predict missing data. There are also a few studies that use rule-based algorithms, such as if-else statements, to model trip purposes (Bohte and Maat 2009; Shen and Stopher 2013).

The variables normally used to model trip purposes are divided into three main categories, i.e. personal characteristics, trip-specific features, and spatial information at the trip destination. Personal characteristics normally include socio-demographic features such as age, gender, employment status, etc. Trip-related features include mode choice, trip starting time, duration of the trip, time spent at the destination, and the day when the trip was carried out. There are some studies which make use of features of the previous trip as well (Cui et al. 2018; Lu, Zhu, and Zhang 2012). Spatial information typically includes the land-use data and the locations of the most visited places, i.e. home and work. Some researchers use land-use databases specified by the development authorities. For example, Y. Liu, Miller, and Habib (2023) uses a dataset which contains the location and categories of businesses in Canada, whereas others use online mapping services such as Google Maps (Cui et al. 2018; Ermagun et al. 2017). Some studies also use social media platforms to supplement the land-use data. For instance, Yazdizadeh, Patterson, and Farooq (2019) and Cui et al. (2018) use Foursquare and Twitter to infer information regarding the land-use type. There are some studies which do not use spatial location data as it is quite costly to use Google Maps and does not always lead to substantial improvements in the accuracy of the trip purpose algorithm (Gao, Molloy, and Axhausen 2021). More recently, some studies have adopted advanced methods, such as graph attention networks, to more effectively exploit points-of-interest data (Liao et al. 2022).

The number and type of trip purposes modelled depend on the research objectives and typically lack uniformity (Ectors et al. 2017; Fu et al. 2025). There are some research studies which classify trip purposes into three main categories, such as home, work and other; whereas, some studies have a large number of trip purposes, for example, Bohte and Maat (2009) and Gao, Molloy, and Axhausen (2021) model 13 and 8 different purposes, respectively. Y. Liu, Miller, and Habib (2023) follow a two-stage approach, where they first predict three main purposes: home, work and other, and then the 'other' purpose is further divided into 13 categories. In addition to differences in purpose granularity, some studies define trip purposes in a context-specific manner, such as focussing on long-distance travel (Z. Liu et al. 2026). Typically, goodness of fit metrics vary depending on the number of trip purposes and data collected; however, the trip purpose prediction accuracy ranges from 40% up to 96% in the literature.

An often-overlooked aspect of the literature is the potential benefit of incorporating imputed trip purposes into downstream travel behaviour models. Existing studies typically focus on predictive performance on testing or validation datasets (Harding et al. 2021; Marra et al. 2019; Yazdizadeh, Patterson, and Farooq 2019). For example, Yazdizadeh, Patterson, and Farooq (2019) argue that GPS traces can be as complete as traditional surveys, reporting inference accuracies of 87% for travel modes and 71% for trip purposes on a testing dataset. However, such evaluations provide limited insight for practitioners, who are ultimately concerned with the implications of using imputed information in subsequent behavioural models. Therefore, in this study, we address this research gap by (a) applying posterior analysis within a mixed multinomial logit framework to impute partially missing trip purposes in a multi-day GPS-based dataset, and (b) quantifying the impact of such imputation on a downstream mode choice model.

3. Methodology

To carry out the study, models for two different types of dependent variables are estimated, i.e. trip purpose prediction models and mode choice models, which are explained in the following sub-sections.

3.1. Trip purpose models

This section describes a succinct background theory for multinomial and mixed multinomial logit model, and the methodology used to predict missing trip purposes in smartphone-based travel survey data.

3.1.1. Background theory

In a random utility context (see e.g. Train 2009), a decision maker n associates a utility $U_{n,i,t}$ for alternative i in choice situation t, which is the sum of a deterministic or observable utility $V_{n,i,t}$ and a random or error component $\varepsilon_{n,i,t}$ (cf. Equation (1)). The decision-maker then chooses the alternative with the highest utility.

$$U_{n,i,t} = V_{n,i,t} + \varepsilon_{n,i,t} \tag{1}$$

The observed component of the utility is a function of taste parameters β and explanatory variables x, i.e. $f(\beta, x)$. The assumptions about the error term have a major influence on the resulting model structure. In multinomial logit (MNL) model, the error term is specified using an independent and identically distributed type 1 extreme value distribution to obtain a closed-form solution for the probability of decision maker n choosing alternative i from a set of J alternatives (as shown in Equation (2)).

$$P_{n,i,t} = \frac{e^{\beta x_{n,i,t}}}{\sum_{j=1}^{J} e^{\beta x_{n,j,t}}} \tag{2}$$

The likelihood of a decision maker n observing a sequence of choices across choice situations T can be written as follows, where $y_{n,i,t}$ is equal to 1 if decision maker n chooses alternative i in task t, and 0 otherwise.

$$L_n(\beta) = \prod_{t=1}^{T} \prod_{i=1}^{J} \left(\frac{e^{\beta x_{n,i,t}}}{\sum_{j=1}^{J} e^{\beta x_{n,j,t}}} \right)^{y_{n,i,t}} \tag{3}$$

The model parameters, i.e. β, are estimated by maximising the full-sample log-likelihood function as described in Equation (4),

$$LL(\beta) = \sum_{n=1}^{N} \ln(L_n(\beta)) \tag{4}$$

In mixed multinomial logit models, the taste parameters β follow a random distribution across decision-makers, such that $\beta_n \sim g(\beta_n \mid \Omega)$, where Ω are the parameters of the multivariate distribution of β. The conditional or posterior distribution of an individual or decision maker n based on observed choices y_n can be calculated using Bayes' rule and is denoted by $h(\beta_n | y_n, x_n, \Omega)$. As this *posterior* distribution makes use of more information (namely past choices), it is possible to better infer the likely location of an individual's taste sensitivities on the sample level distribution with a tighter confidence interval (Train 2009).

In this study, we use the hierarchical Bayesian method to estimate the mixed multinomial logit along with the posteriors. The hierarchical Bayesian method is commonly associated with marketing and consumer research literature (Rossi and Allenby 2003). Bayesian methods are more suited for posterior analysis, as the posterior distributions are standard outputs. Further, Bayesian methods are faster to estimate compared to the classical or frequentist estimation procedures when there are a large number of random parameters (Train 2009).

Bayesian estimation also offers the analyst flexibility in terms of the distributional assumptions for β such as the log-normal or a triangular distribution. In the context of using the MMNL model for predicting trip purposes, we rely on a multivariate normal distribution for β in the absence of clear a priori sign assumptions, i.e. $\beta_n \sim N(b, W)$, where b are means and W the covariance matrix. Independent prior densities are assumed for b and W, and we follow Train (2009) by adopting diffuse priors which are respectively normal and inverse Wishart (IW) in shape. The choice for diffuse priors limits the impact of the prior on the posterior density. To learn about the model parameters, Gibbs sampling is used to conditionally sample from three layers. First, draws are taken from the conditional posterior density for the individual level parameter β_n, which represents the taste sensitivities of individual n. The conditionality is based on the individual's observed data y_n and current draws for b and W, as indicated in Equation (5). The second and third layers, i.e. Equations (6) and (7), sequentially update b and W based on the updated draws of β_n. This estimation framework is called hierarchical Bayesian as there is a hierarchy among the individual level parameters or posteriors with the sample level parameters.

$$K(\beta_n \mid b, W, y_n) \propto \prod_t \frac{e^{\beta_n x_{n,y_{nt},t}}}{\sum_j e^{\beta_n x_{n,y_{njt}}}} \phi(\beta_n \mid \; b, W) \tag{5}$$

$$K(b \mid \beta_n, W) \text{ is } N(\overline{B}, W/N) \text{ where } \overline{B} = \sum_n \beta_n / N \tag{6}$$

$$K(W \mid b, \beta_n) \text{ is } IW\left(K + N, \frac{KI + N\overline{S}}{K + N}\right) \quad \text{with } \overline{S} = \sum_n (\beta_n - b)((\beta_n - b)'/N \tag{7}$$

3.1.2. Predicting missing trip purposes

In our study, we predict missing trip purposes using the above mentioned methods, i.e. multinomial logit model, and mixed multinomial logit model with and without posteriors. Since a portion of the trip purposes are observed or tagged in smartphone-based travel survey datasets, we can estimate models using the observed trip purposes as the dependent variable with associated explanatory variables, such as socio-demographic features, land-use features, and trip-specific details such as the day (weekday vs weekend) and time when the trip is carried out. We can then utilise the estimated models to impute or predict the missing trip purposes in the dataset.

It is important to note that while MNL and MMNL models are typically applied to predict *behavioural* choices, such as mode or route choices, in our context, they are used as *statistical* models to capture the associations between trip purposes (discrete dependent variables) and explanatory variables. Further, all inference-based trip purpose models implicitly rely on a missing-at-random (MAR) assumption, which states that missing trip purposes are explainable by observed explanatory variables. Under this assumption, higher missing rates for certain purposes (e.g. leisure trips) do not induce systematic bias in estimates provided relevant explanatory variables are observed. If missingness depends on unobserved factors, i.e. under a missing not at random (MNAR) assumption, then no model-based imputation approach can fully recover missing trip purposes without explicitly modelling the missingness mechanism. For example, if all long-distance leisure trips are not systematically recorded and explanatory variables not observed, then the model will be unable to recover such trip purpose accurately, leading to biased imputations.

Let us consider that the trip purposes are missing in situations t^* with corresponding explanatory variables x_{n,t^*}. For a developed multinomial logit model, predicting the probability of each possible trip purpose at t^* is straightforward. This can be done by using Equation (2), where we apply the estimated parameters β to the explanatory variables x_{n,t^*} associated with that particular situation.

To calculate the probability of predicting trip purposes in the mixed multinomial logit model, the baseline approach would be to use the sample level model only, with the predicted probability of individual n making a trip for purpose i given by:

$$P(i \mid x_{n,t^*}, \Omega) = \int L_{n,t^*}(i \mid \beta_n) g(\beta_n \mid \Omega)\, \mathrm{d}\beta, \tag{8}$$

where $L_{n,t^*}(i \mid \beta_n)$ is the likelihood of individual n having trip purpose i in situation t^*, with explanatory variables x_{t^*} and with Ω giving the estimated sample or population level parameters. In simple terms, the predicted probabilities are computed using draws from the sample-level parameters distribution $g(\beta_n \mid \Omega)$, and the results are then averaged.

However, for individuals where we observe some past trip purpose information, we can instead make use of the posterior distribution, given by:

$$P(i \mid x_{n,t^*}, y_n, x_n, \Omega) = \int L_{n,t^*}(i \mid \beta_n) h(\beta_n \mid y_n, x_n, \Omega)\, \mathrm{d}\beta, \tag{9}$$

where this now also relies on the observed prior trip purposes (y_n) and the posterior distributions $h(\beta_n \mid y_n, x_n, \Omega)$. This implies that rather than relying solely on sample-level parameter estimates, we use draws from individual-specific posterior distributions $h(\beta_n \mid y_n, x_n, \Omega)$ to predict missing trip purposes. In our methodology, the MMNL model is estimated using a hierarchical Bayesian approach, under which posterior distributions are obtained directly as standard outputs. Equivalent posterior distributions can also be recovered under a classical maximum simulated likelihood framework and can be readily calculated using choice modelling libraries, including BIOGEME and Apollo (Bierlaire 2003; Hess and Palma 2019).

The main intuition behind using posterior analysis within the MMNL modelling framework to predict missing or untagged trip purposes is its ability to leverage the panel nature of the dataset to capture persistent individual-specific patterns or random effects. Standard models such as the MNL (or even machine learning classification algorithms, such as a neural network or random forest) primarily capture the average impact of an explanatory variable on the dependent variable, which are reflected by fixed β parameters for the complete sample. In contrast, MMNL models, estimate a distribution of individual-specific parameters, i.e. $\beta_n \sim g(\beta_n \mid \Omega)$ allowing effects to vary across individuals. For example, let's consider two individuals: where individual A typically travels 10 km from home to work before 9 am, and another individual B travels after 9 am from a distance of 3 km for work. Posterior analysis within an MMNL modelling framework, in which all parameters are random, captures and retains such associations or effects at the individual level, whereas an MNL model would only reflect average sample-level effects. If one of individual A's weekday 10 km morning trips is untagged, then posterior distribution being informed by the individual's previously observed trips, assigns a higher probability that the missing trip is work. As more observations are available per individual, posterior distributions $h(\beta_n \mid y_n, x_n, \Omega)$ can be estimated with greater precision, improving predictions of missing or untagged trip purposes, particularly for habitual trips (e.g. commuting to work, routine shopping, or leisure trips). In contrast, for individuals with fewer tagged trips, the posterior estimates are weakly informed and would converge toward the sample level average β_n parameters. It must be noted that if an individual's observed trip purposes are weakly related to the explanatory variables, the gains from posterior analysis will be limited and as a result, posterior analysis might not provide greater improvements over sample level estimates.

In real-world scenarios, we can only predict the probabilities of the missing trip purposes and cannot validate the results due to the absence of ground truth. However, to compare the three methods, we split the tagged dataset into training and testing subsets, allowing us to evaluate performance on an unseen testing data. We compare the predictive performance in terms of the average probability of the chosen (or correct) alternative and error in predicting the aggregated market counts, both of which are probabilistic goodness of fit metrics. To calculate the error in market counts, the mean absolute percentage error is used which is defined in Equation (10) for *k* purposes. Further, we use additional metrics including weighted MAPE and root mean percentage square error (RMPSE). As a secondary evaluation metric, we also report the model's accuracy or hit rates, which is a deterministic goodness of fit metric commonly used in the machine learning literature to assess hard classifications. Accuracy is defined as the percentage of correct predictions, where the trip purpose which has the highest probability is considered as the *predicted* purpose.

$$\frac{1}{K}\sum_{k=1}^{K}\left|\frac{\textit{predicted counts} - \textit{observed counts}}{\textit{observed counts}}\right| * 100 \tag{10}$$

3.2. Mode choice model

In the mode choice model, it is assumed that trip purpose (a discrete independent variable) is missing or uncertain. There have been some previous research studies which formulate methods to estimate discrete choice models where the missing or uncertain data is a continuous independent variable such as income (Sanko et al. 2014) or travel time (Biswas et al. 2024), or a dependent variable (Andersson et al. 2022; Vij and Shankari 2015). There are also some studies which have focussed on inferring demographics using a latent variable approach (Bwambale, Choudhury, and Hess 2019). However, there are no specific case studies in the transport literature where only a single discrete independent variable is missing.

In our study, we follow a two-stage approach where we first impute the probability of observing a trip purpose and then utilise the probability as weights in a weighted estimation of the mode choice model. The two-stage approach can be better visualised and

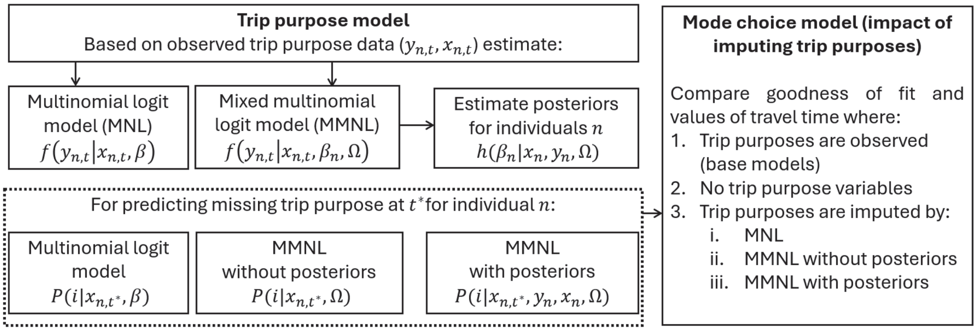

Figure 1. Workflow for the study.

assessed using the developed flowchart in Figure 1. The log-likelihood equation for estimating the mode choice model parameters (β) is given in Equation (11), where $P_{n,m,t|tp}$ represents the standard multinomial logit probability of observing mode choice alternative m (out of M modes) for individual n at choice situation t conditional on the trip purpose tp (out of TP trip purposes). The probability of observing the trip purpose tp based on the different imputation strategies (MNL, MMNL with and without posteriors as estimated in previous section) is assigned as a weight $\pi_{tp,n,t}$, and $y_{n,m,t}$ is 1, if individual n chooses alternative m in choice situation t, 0 otherwise. Simply put, if the observed mode choice is car, and there are 2 trip purposes, i.e. commute and business trips, with imputed trip purpose probabilities of 0.2 and 0.8. There would be a 0.2 weight for observing the mode choice as car with the trip purpose in the explanatory variable set as commute, and a 0.8 weight for observing the mode choice car with the trip purpose set as business. For any trips with an observed trip purpose, the weight would be 1; hence, there would only be a single contribution to the likelihood. This weighted likelihood approach resembles a traditional latent class choice model, where the latent classes correspond to trip purposes and the averaging is carried out at the observation level. The primary reason for adopting the weighted likelihood approach is that it enables the estimation of purpose-specific behavioural parameters, such as travel time sensitivities for commuting or business trips, allowing us to derive values of travel time (VTT) for different trip purposes.

$$LL(\beta) = \sum_{n=1}^{N} \ln\left(\prod_{t=1}^{T}\prod_{m=1}^{M}\left(\sum_{tp=1}^{TP} \pi_{tp,n,t} \cdot P_{n,m,t|tp}\right)^{y_{n,m,t}}\right) \tag{11}$$

4. Data

Travel diaries from a survey conducted primarily in the Yorkshire and Humber region in the UK are used to carry out the study. The survey was carried out using the smartphone-based app RMove for a 2-week period between November 2016 and March 2017. For more details regarding the collection of the data, readers are referred to Calastri, Crastes dit Sourd, and Hess (2020).

The travel survey application required users to select the trip purposes and modes from a drop-down menu. A total of 41,210 trips by 580 individuals were recorded. Out of these,

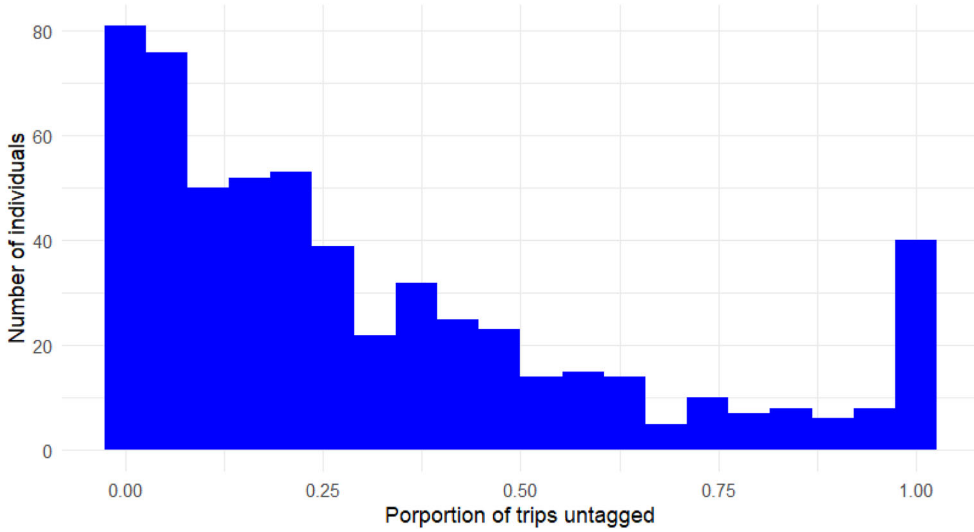

Figure 2. Histogram of missing trip purpose in dataset.

nearly a quarter of trip purposes (11,130 trip purposes or 27.8% of the total trips) were left untagged, leaving the remaining 72.8% as tagged. There is wide variability across individuals tagging behaviour as highlighted in Figure 2, which shows the proportion of untagged trip purposes per individual. It is observed that about 15% of the individuals completely tagged their trip (nearly 80 out of 580). However, a majority of individuals (50% of the total individuals) did not tag up to 30% of their trips. The histogram also shows that numerous individuals (33 out of 580 individuals or about 5% of the individuals in the sample) left all of their trips untagged. This indicates that there is a large proportion of trips that can be predicted by making use of posterior analysis in mixed multinomial logit models.

After removing the untagged trips, trips outside the Yorkshire and Humber region, and individuals with incomplete sociodemographic information or unidentified home locations, there are a total of 24,116 trips carried out by 484 individuals. The survey used a total of 31 trip purposes, which, for ease of analysis, were aggregated and classified into 5 main trip purposes, namely commute, business, home, shopping, and leisure trips. We opted for these categories to maintain a manageable number of trip purposes as increasing the number of categories substantially raises the computational burden of estimating the mixed multinomial logit model. In addition, business trips were separated from commuting trips because business travellers tend to have higher travel time values and are treated separately in mode choice models (Wardman, Chintakayala, and de Jong 2016). Nevertheless, future researchers may adopt alternative or more detailed trip purpose classifications, depending on the requirements of their specific study. Table 1 shows the count and percentages of the trip purposes. It can be observed that 31.42% of trips are home trips, 28.09% shopping, 19.87% leisure, 13.64% commuting and only 6.97% are business trips. Shopping/chores/service/education/escort trips were separated from leisure trips, as leisure activities and trips are quite different in nature. It was possible to add a separate category for educational trips; however, the total number of educational trips were low (418

Table 1. Aggregated trip purposes in dataset.

Trip purposes	Count	Percentage (%)
Commute	3,290	13.64
Business	1,682	6.97
Return to home	7,578	31.42
Shopping/services/education/caring/other	6,773	28.09
Leisure (Entertainment, Visit Friends, Hobbies)	4,793	19.87

total trips), so we decided to add them in the within the broader shopping category, which also encompasses other non-leisure, non-work activities.

To better model trip purposes, the dataset was augmented with the nearest points of interest data and home location. To find the home location, Density-Based Spatial Clustering of Applications with Noise (DBSCAN), which is a spatial clustering algorithm, was used. One of the benefits of using DBSCAN to find unique residential locations is that the clustering algorithm identifies outliers (or noise variables) which is the reason why this algorithm has been used in previous studies to model residential locations in the transport literature (Amaya, Cruzat, and Munizaga 2018; Lizana, Choudhury, and Watling 2023; Winkler, Meister, and Axhausen 2024). The clustering algorithm was used to find the unique home location for individuals who had specified their home trips. If there was only 1 cluster, then that centroid of the cluster was considered as the home location. However, if more than 1 cluster was found, the cluster with the highest number of origin or destination points was considered as the home location. For respondents who did not specify home trips at all, the origin location of their first trip in a day was checked. If only one cluster was formed, the centroid of the cluster was considered the home location, otherwise, observations from that individual were discarded (Ali et al. 2026). The minimum number of points to specify a cluster was 2 and the distance threshold between two clusters was set to 40 meters as these parameters produced the most consistent and plausible home locations for individuals across the sample.

Further, points of interest data published by Ordnance Survey (2023) were used to know the land-use type at the trip's destination. Ordnance Survey is the most comprehensive point of interest data in the UK, which contains geotagged details of businesses, services, and educational institutions that are updated quarterly. Since the survey was carried out between November 2016 and March 2017, the points of interest data published in April 2017 were used. A total of 335,832 points of interest were identified for the Yorkshire and Humber area. There are a total of 52 main classes and over 600 subclasses of land-use type in the Ordnance Survey; however, based on our judgement, these classes were aggregated into 10 land-use categories, including restaurants, services, shops and offices as mentioned in Table 2. The services points of interest category include post offices, banks, police stations, etc; entertainment areas include recreational and outdoor areas such as zoo and botanical gardens. Transportation hubs include airports, bus terminals, train terminals, and green areas including parks, and bodies of water. The nearest 20 points of interest, up to a distance of 250 meters, were used as this configuration yielded the best predictive performance.

Table 2 shows the descriptive statistics for the observed trip purpose data. It can be observed that nearly all of the commuting and business trips take place on weekdays as expected. Likewise, a higher proportion of shopping and leisure trips are carried out on the

Table 2. Descriptive statistics of trip purpose dataset.

Descriptive statistics	Business	Commute	Home	Leisure	Shopping
Proportion of trips in each explanatory variable					
Type of day					
Weekday	0.93	0.96	0.75	0.65	0.73
Weekend	0.07	0.04	0.25	0.35	0.27
Gender					
Male	0.52	0.44	0.40	0.39	0.37
Female	0.48	0.56	0.60	0.61	0.63
Occupation					
Full time employed	0.79	0.81	0.62	0.62	0.56
Not fully employed/other (e.g. retired)	0.21	0.19	0.38	0.38	0.44
Starting time of trip					
Start time 0:00 to 9:00	0.19	0.52	0.05	0.07	0.13
Start time 9:00 to 16:00	0.67	0.40	0.36	0.49	0.58
Start time 16:00 to 19:00	0.11	0.05	0.37	0.27	0.22
Start time 19:00 to 24:00	0.03	0.03	0.22	0.17	0.08
Nearest 20 point of interest land-use categories					
Restaurants	2.26	2.27	1.40	2.74	2.52
Services	8.56	9.97	7.61	7.71	8.40
Transportation hubs	2.66	2.32	2.49	2.33	2.52
Shops and offices	0.97	0.90	0.66	1.38	1.98
Entertainment	1.11	1.29	0.73	1.21	0.91
Industrial areas	0.96	1.08	0.74	0.70	0.72
Green areas	0.07	0.07	0.06	0.09	0.07
Educational institutes	0.50	0.52	0.47	0.34	0.43
Temporary accommodation	0.12	0.14	0.11	0.14	0.09
Agricultural	0.00	0.01	0.02	0.02	0.02
Average value of explanatory variable					
Distance to home (Km)	11.10	13.80	6.15	12.02	8.29

weekend. There is a near equal split of females in commuting and business trips, however, 60% of the social and leisure trips are carried out by females. Full time employed people are observed to carry out a higher proportion of commuting and business trips. Starting time also has a significant impact on trip purpose. For instance, most of the commuting and business trips take place before 16:00 hours, whereas home, leisure and shopping trips are carried out throughout the day. At night, i.e. after 19:00, most trips are home trips. In terms of land use, it can be observed that home trips are close to areas with a reduced number of restaurants, services, shops and offices areas which is intuitive. Commute and business trips are closer to locations that have more services and industrial areas. Shopping trips have a higher proportion of restaurants, shops and office areas at their destination. Similarly, leisure trips have a higher number of shops, offices, and entertainment areas at their destination compared to other trip purposes. Distance to home is a clear explanator for home trips compared to business and commute trips, which are on average to locations 11.10 km and 13.80 km from home locations, respectively. The average distance to home for home trips is relatively high at 6.15 km, primarily due to outliers, specifically, individuals with two homes who label distant locations as 'home.' While such trips could have been excluded from the analysis, we chose to retain them, as similar cases are likely to occur in other datasets as well. Nevertheless, to reduce their influence, we applied a logarithmic transformation to the distance-to-home variable while modelling trip purposes.

The mode choice dataset is a subset of the complete data, comprising 12,524 observations from 540 individuals. The mode choice analysis is restricted to six alternatives, i.e. car, bus, rail, taxi, cycling, and walking. Other modes (e.g. motorcycle, air travel, and multimodal trips) are excluded. Trips shorter than 100 metres are also removed, and only observations with at least one feasible alternative are retained (e.g. car or walking trips without any other available alternative are excluded), further reducing the sample size. In comparison, the trip purpose dataset contains 24,116 observations from 484 individuals; the smaller number of individuals largely reflects the exclusion of 66 participants for whom home locations could not be reliably inferred. To calculate the travel time for the choice set in the data, Google API were used. The dataset contains socio-demographic features of individuals typically used in mode choice decisions such as income, gender, age and occupation. For more details regarding the mode choice data and the derivation of values of travel time (VTT) based on the dataset, the readers are referred to Tsoleridis, Choudhury, and Hess (2022). The mode choice dataset has 12,524 trips in total, but only a reduced proportion of the dataset (4896 trips) was considered in this study to ensure that there is no overlap in the estimation or training dataset of trip purpose and mode choice data to avoid any endogeneity or measurement errors between the two models.

To carry out the study, the trip purpose dataset was divided into a training and testing dataset. 30% of each individual's trips with observed purposes were randomly set aside in the testing dataset, which led to 16,892 and 7,224 trips in the training and testing dataset for the trip purpose model, respectively. Of the 4,896 trips used in the mode choice model, 3296 or 67% of the trips are within the testing dataset of the trip purpose model. Hence, we estimate mode choice models where 3,296 out of 4,896 trip purposes are imputed from the trip purpose models and for the remaining 1600 trips, the observed trip purposes are used. Thus, simulating a case where a large proportion of trip purposes are uncertain.

5. Results and discussions

This section is divided into two parts. The first section deals with the trip purpose models and their performance in predicting missing purposes, while the second part focuses on the impact of imputing trip purposes in a mode choice model.

5.1. Trip purpose model

We use a multinomial logit model (MNL) as an initial step towards the development of the mixed multinomial logit (MMNL) model. The variables tested in the MNL model included socio-demographic features, trip features, and spatial attributes. The trip purpose 'home' was used as the base. All of the variables mentioned in Table 2 were tested in the utility specification. Some variables which led to statistically insignificant and non-intuitive results were not included in the final model. These include the land-use categories: transportation, agricultural and temporary accommodation, and gender. The starting time of trips was specified in 5 categories, namely from 0:00 to 9:00 hours, 9:00 to 16:00 hours, 16:00 to 19:00, and 19:00 to 24:00, where the latter was fixed as the base. It is possible to split the time into more categories; however, this increases the total number of variables and the time needed to estimate the MMNL models.

To estimate the mixed multinomial logit model, the parameters associated with all variables were considered to be normally distributed with a full covariance matrix. We used 30,000 burn-in and 30,000 post burn-in draws after inspecting the stability of the Monte Carlo Markov Chains (MCMC) in earlier runs. To improve the convergence of the draws, the variables were scaled as well. Further, to carry out thinning of the MCMC chains, every 10th draw was retained. All models were estimated using Apollo, which is a choice modelling library in R (Hess and Palma 2019). The multinomial logit model has a log-likelihood of −19,154.43 with 47 estimated parameters, whereas a calculation of the equivalent model fit of the mixed multinomial model using the sample level distributions gives a log-likelihood of −16,374.29, with the equivalent of 1,175 estimated parameters (47 means and 1,128 covariance matrix terms).

The mean (μ) and standard deviation (σ) of the parameters for the mixed multinomial logit model can be observed in Table 3.[1] The mean of alternative specific coefficients (i.e. μ parameters) are negative for the trip purposes commuting, business, and shopping, which is expected as the base trip purpose, i.e. home trips, has the highest number of observations in the dataset. The μ asc coefficient for leisure is positive but has high posterior standard deviation. The standard deviations of the ascs, i.e. the σ parameters, have high posterior means indicating a high extent of unobserved preference heterogeneity for different trip purposes. The weekday commute, business, and shopping μ coefficients are positive which indicates that on weekdays there is an increased propensity of these purposes being carried out compared to weekends, which is intuitive. The trend is opposite for leisure trips as expected, which indicates a lower propensity of leisure trips on weekdays. It is also observed that fully employed individuals are more likely to carry out trips with commuting and business purposes. All of the coefficients for distance variables for commuting, business, and shopping trips are positive, which is intuitive, as an increase in distance from home implies an increase in the probability of the purposes being other than a home trip.

Time also impacts the trip purposes significantly, as indicated by the coefficients. The purposes are more likely to be commuting, business, shopping, and leisure trips up to 19:00 hours compared to home trips. In terms of land-use variables, it can be observed that when the trip's destination has a high number of restaurants, there is an increased likelihood of observing leisure, shopping, and business trips. Likewise, with an increase in the number of services at a trip's destination, there are more commuting trips. With an increase in shopping and office points of interest, there is an increase in the probability of trips being for leisure, shopping, and commute purpose, which is intuitive. If the trip destination is closer to an industrial area, there is an increased likelihood of them being commute or business trips. It can also be observed that with an increase in the number of entertainment points of interest, there is an increase in the likelihood of trips being commute or leisure trips. One of the potential reasons why there is an increase in commuting trips is that many people work in entertainment areas leading to a positive coefficient for commuting trips. With an increase in the number of green areas in the trip destination, there is an increase in the probability of such trips have a shopping or leisure purpose. Since education trips are included in the shopping trip purpose, it can be observed that with an increase in the number of educational institute at the trip destination there is an increased likelihood of trips being for shopping purpose.

Tables 4 and 5 show the average probability of the chosen alternative, accuracy, and aggregated market counts calculated by applying the estimated models on the training and the testing dataset. It can be observed that in both the training and testing dataset, the MMNL model with posterior analysis has the best performance in terms of the average probability of the chosen alternative and accuracy. In the testing dataset, the MMNL model with posteriors also obtains the lowest mean absolute percentage error (MAPE) weighted

Table 3. Estimates for mixed multinomial logit trip purpose model (home treated as base).

Parameter		Commute		Business		Leisure		Shopping	
		Post means	Post std	Post means	Post std	Post means	Post std	Post means	Post std
asc	μ	−12.67	0.56	−9.73	0.42	0.13	0.085	−0.99	0.11
	σ	5.88	0.42	4.06	0.35	1.23	0.13	1.80	0.14
$\beta_{\mathtt{weekday}}$	μ	4.60	0.24	3.26	0.16	−0.41	0.05	0.06	0.01
	σ	2.72	0.20	1.86	0.14	0.52	0.07	0.07	0.01
$\beta_{\mathtt{fully\ employed}}$	μ	1.74	0.14	1.21	0.28	–	–	–	–
	σ	0.86	0.09	0.98	0.09	–	–	–	–
$\beta_{\mathtt{distance}}$	μ	1.59	0.09	1.56	0.07	0.91	0.03	0.71	0.03
	σ	1.05	0.07	0.96	0.07	0.51	0.03	0.46	0.03
$\beta_{\mathtt{time\ 0\ to\ 9\,hrs}}$	μ	6.63	0.21	4.29	0.24	0.71	0.11	2.22	0.11
	σ	3.03	0.24	2.04	0.21	1.19	0.11	1.25	0.15
$\beta_{\mathtt{time\ 9\ to\ 16\,hrs}}$	μ	3.44	0.26	3.26	0.28	0.15	0.05	1.31	0.11
	σ	2.06	0.16	1.50	0.14	0.51	0.08	0.90	0.09
$\beta_{\mathtt{time\ 16\ to\ 19\,hrs}}$	μ	0.24	0.01	0.48	0.14	−0.22	0.07	0.48	0.1
	σ	0.13	0.02	0.73	0.11	0.75	0.07	0.79	0.10
$\beta_{\mathtt{restaurants}}$	μ	0.09	0.01	0.13	0.02	0.17	0.01	0.13	0.01
	σ	0.12	0.02	0.13	0.01	0.10	0.01	0.05	0.01
$\beta_{\mathtt{services}}$	μ	0.08	0.01	0.04	0.01	−0.02	0.01	0.04	0.01
	σ	0.11	0.01	0.06	0.02	0.07	0.01	0.03	0.01
$\beta_{\mathtt{shops\ offices}}$	μ	0.01	0.01	−0.06	0.01	0.14	0.01	0.29	0.02
	σ	0.07	0.01	0.08	0.01	0.08	0.02	0.24	0.02
$\beta_{\mathtt{entertainment}}$	μ	0.25	0.06	0.12	0.05	0.26	0.03	0.03	0.01
	σ	0.63	0.08	0.39	0.05	0.32	0.03	0.07	0.01
$\beta_{\mathtt{industrial}}$	μ	0.03	0.05	0.04	0.02	–	–	–	–
	σ	0.53	0.05	0.09	0.01	–	–	–	–
$\beta_{\mathtt{green}}$	μ	–	–	–	–	0.29	0.07	0.40	0.09
	σ	–	–	–	–	0.62	0.12	1.03	0.14
$\beta_{\mathtt{education}}$	μ	–	–	–	–	–	–	0.07	0.04
	σ	–	–	–	–	–	–	0.49	0.05

Table 4. Predictive performance of estimated trip purpose models.

Metric	Multinomial logit	Mixed multinomial logit	
		Without posteriors	With posteriors
Training dataset			
Average probability of chosen alternative (APCA)	0.415	0.408	0.589
Accuracy (%)	54.89	44.88	70.42
Testing dataset			
Average probability of chosen alternative (APCA)	0.413	0.408	0.537
Accuracy (%)	54.59	44.11	62.50

Table 5. Observed and predicted aggregated market counts of trip purpose models.

Description	Observed market counts	Predicted market counts		
			Mixed multinomial logit	
		Multinomial logit	Without posteriors	With posteriors
Training dataset				
Commute	2, 332	2, 332	2, 185.54	2, 331.67
Business	1, 168	1, 168	1, 349.35	1, 169.08
Home	5, 319	5, 319	4, 950.49	5, 323.66
Shopping	4, 708	4, 708	4, 872.93	4, 714.11
Leisure	3, 365	3, 365	3, 533.70	3, 353.48
MAPE (%)	–	0.00	7.45	0.13
Weighted MAPE (%)	–	0.00	6.10	0.14
RMPSE(%)	–	0.00	6.42	0.14
Testing dataset				
Commute	958	973.64	917.86	958.19
Business	514	497.26	574	492.59
Home	2, 259	2, 254.08	2, 094.35	2, 260.83
Shopping	2, 065	2, 048.2	2, 116.6	2, 066.15
Leisure	1, 428	1, 450.81	1, 521.18	1, 446.27
MAPE (%)	–	1.50	6.44	1.12
Weighted MAPE (%)	–	1.06	5.67	0.59
RMPSE(%)	–	1.78	6.88	1.26

MAPE, and root mean percentage square error (RMPSE) to predict market counts, showing the benefits of the proposed approach. The only exception to this trend is in the MAPE on the training dataset, where the use of an MNL model with a full set of constants leads to the perfect recovery of overall market counts, which is expected due to the nature of the MNL model and does not necessarily translate to a better fit on a testing dataset.

There is a substantial increase in the predictive performance of MMNL when posteriors are used, with the average probability of the chosen alternative (APCA) showing a noticable increase in both the training and testing data. MNL has a slightly better average probability of the chosen alternative compared with MMNL without posteriors, and also a lower MAPE in both the training and testing data. The lack of advantage of the MMNL model with sample level parameters is not unexpected in prediction, given that the predictions need to be averaged across draws from the distribution. This also indicates that MMNL models, which often have better in-sample goodness of fit compared to MNL models, may not offer better predictive performance in testing datasets compared to MNL if the same individuals are not present within the dataset.

Figure 3 further shows the average probability of alternatives given the observed choice in the testing dataset. It can be observed that the purpose-specific average probabilities are better for the MMNL model with posteriors with a range from 0.34 to 0.70, which is superior compared to MNL and MMNL without posteriors, which have a range from 0.15 to 0.59. Both the MNL and MMNL without posteriors model have a low average probability of correctly inferring business trips, i.e. 0.15, which is less than a random guess. This indicates both of these models have not fully captured the data-generating process of business trips. There is a substantial increase in the average probability of correctly inferring business trips for the MMNL model with posteriors as the average probability is 0.34. This is quite intuitive as once we predict future choices conditioned on previously observed choices, we can better capture the fact that only some individuals (most likely to be employed individuals) make repeated business trips. These results clearly show that the use of posterior analysis in MMNL leads to the least predictive error at the sample and individual level compared to MNL and MMNL without posteriors. Further, we find that the MNL model has a superior predictive performance compared to the MMNL model without posteriors.

As additional benchmarks, we also estimated machine learning (ML) models (neural networks and random forests). The neural network and random forest models achieve APCA values of 0.410 and 0.400 on the testing dataset, respectively, indicating that differences in predictive performance for the machine-learning models relative to the MNL and the MMNL without posterior analysis are relatively small, while the MMNL with posterior analysis remains the best-performing model. To better generalise these findings, we also compare the results of models estimated on different training and testing data splits, i.e. 60–40% and 50–50% split as opposed to the 70–30% split described here. The complete results are shown in Appendix A.1; however, the findings are similar, i.e. the best predictive performance is for the MMNL with posterior analysis, followed by the MNL model, while MMNL without the use of posterior distributions has the least predictive performance.

Our results are similar to Krueger et al. (2021) where MMNL with posteriors performs better than both MNL and MMNL without posteriors. In this study, we observe that the APCA (average probability of chosen alternative) increases from 0.408 to 0.537 by making use of posteriors in the MMNL model, which is similar in magnitude to what was reported by D. Revelt and Train (1999) and Krueger et al. (2021) who show a change from 0.35 to 0.50, and from 0.44 to 0.58, respectively, by making using of posteriors in the MMNL model. Danaf et al. (2019) observes a larger increase in APCA from 0.509 to 0.717.

The impact of using posteriors in the mixed multinomial logit model was further analysed at a trip and individual level. It was observed that at a trip level, 73.3% (5,293 out of 7,224) trips have a 0.215 increase in the average probability of the actual observed purpose when posteriors are used compared to without using posteriors. Whereas for the remaining 26.7% of trips, there is an average decrease of 0.107 in the probability of chosen alternative compared to without using posteriors in MMNL. At an individual level, 91.1% (441 out of 484) of individuals show an increase in the average probability of chosen alternative by making use of posteriors in the MMNL model. It was further observed that individuals who have an increase in the average probability of the chosen alternative in the testing dataset by making use of posteriors have an average of 36.23 trips or observations in the training dataset. On the other hand, individuals whose predictive performance decreased due to posteriors have an average of 21.23 trips. These results indicate that for nearly all of the individuals, there is a gain in the predictive performance by making use of their previously observed choices. The fact that there is a drop in performance for those individuals with fewer trips is in line with expectation as there is less data from which to *learn* the posterior distribution.

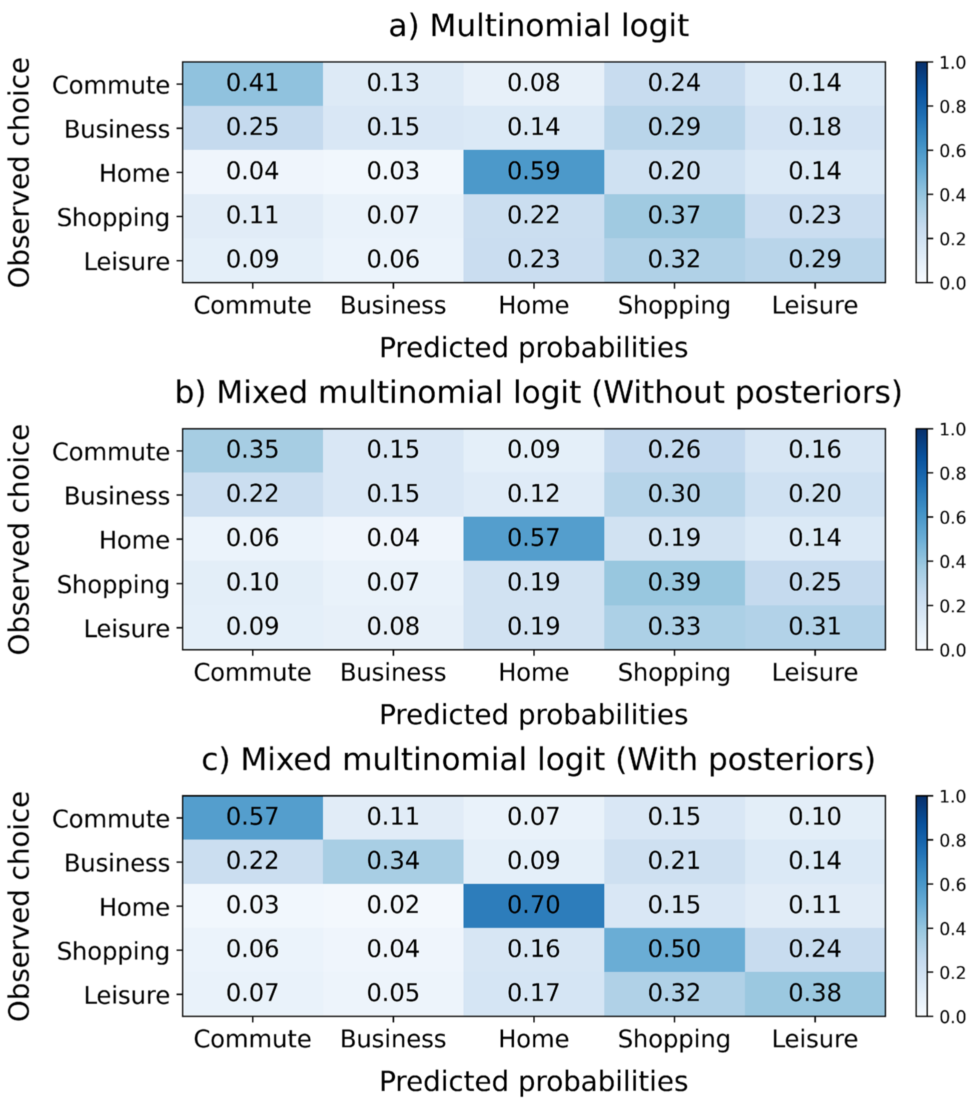

Figure 3. Predicted probabilities with respect to observed choices in testing dataset for (a) multinomial logit, (b) mixed multinomial logit without posteriors and (c) mixed multinomial logit with posteriors.

Overall the predictive performance for the trip purpose model is satisfactory compared to other studies which model trip purpose, as the model with the highest accuracy is 62.50%, i.e. the mixed multinomial logit model with posteriors, which is in the range of previous trip purpose studies, i.e. with a range of 40 to 96% (Servizi et al. 2021). The value is a bit lower than Montini et al. (2014), Yazdizadeh, Patterson, and Farooq (2019), Y. Liu, Miller, and Habib (2023), and Ermagun et al. (2017) which have accuracy of 79.8%, 71%, 79%, and 64.17% respectively. While it is difficult to compare the predictive accuracies across different studies since most studies have different number of purposes and sample size, one of the reasons why this study has lower accuracy is because the dataset is imbalanced due to the lower number of business trips. Another reason is the spatial context as most previous studies (Ermagun et al. 2017; Yazdizadeh, Patterson, and Farooq 2019) were conducted in regions such as Canada and the USA, which are characterised by relatively homogeneous land-use patterns. In contrast, land-use in the UK is denser and more heterogeneous, reducing the extent to which land-use characteristics can reliably predict trip purposes. Nevertheless, our approach is effective, as the use of a posterior analysis in an MMNL achieves substantially better predictive performance compared to standard modelling strategies, including multinomial logit and machine learning models.

5.2. Impact of imputing trip purposes in mode choice model

In this section, the impact of imputing trip purposes in the mode choice model is presented. Six different models are estimated, where the base model contains all of the dataset with observed trip purposes. In the second model, the complete dataset is used without any trip purpose variables. In the remaining four models, nearly two-third of the trip purposes, i.e. 3,296 out of the 4,896 trips, are imputed using the different strategies. The first of these uses a uniform probability of observing each trip purpose, i.e. business, commute, and 'other trips' based on the observed market shares which are 6.1%, 18.5% and 70.4%, respectively (labelled as uniform weights). The other three models use the predicted purposes using MNL, MMNL without and with posteriors. Imputation is only used for 3296 trips, while, for the other 1600 trips, the observed trip purposes are considered – thus simulating a case with a high share of missing purpose information.

To specify the mode choice model, a multinomial logit model is considered with similar model specification as mentioned in Tsoleridis, Choudhury, and Hess (2022). The travel cost is specified using a logarithm scale, while the travel times for all modes, except for rail, are modelled using a box-cox transformation. The travel time for rail was specified on a linear scale. The trip purpose variables, i.e. commute and business trips, enter the utility equation as shifts (i.e. Δ variables) in the travel time and alternative specific constants (asc) as boldfaced in Table 6. There are a total of five variables which interact with the trip purpose, i.e. travel time shifts in commuting and business trips for car, asc shifts for commuting trips carried out by bus and walking, and asc shift for business trips carried out by cycling. Other trip purpose variables, such as a cost parameter shift for business trips, were also considered; however, the coefficients were not retained as they were not in line with behavioural theory. The same approach followed by Tsoleridis, Choudhury, and Hess (2022) is used to derive the values of travel time, which uses sample enumeration on the UK National Travel Survey between 2015-2017 that are further weighted based on the distance segments observed in the UK 2011 Census work trips.

Table 6 shows the model fit, coefficients, and the value of travel times for the estimated mode choice models. It can be observed that the coefficients of the base model are intuitive with negative travel time and cost coefficients. The travel time shifts for commute and business trips carried out by cars are negative, indicating that individuals are more sensitive to the travel time for these purposes compared to other trips such as shopping and leisure trips, which is intuitive. Individuals are also more likely to carry out commuting trips by walking and bus as indicated by the positive asc shifts for these coefficients. The asc shift for business trips carried out by cycling is negative which is intuitive as individuals are

Table 6. Mode choice model estimates and values of travel time.

			Mode choice models where 3,296 trip purposes out of 4,896 are imputed by:			
Description	Base model	No trip purpose	Uniform weights	MNL	MMNL without posteriors	MMNL with posteriors
Number of individuals	525	525	525	525	525	525
Number of modelled outcomes	4,896	4,896	4,896	4,896	4,896	4,896
Estimated parameters	47	42	47	47	47	47
LL (final)	−1,689.50	−1,720.96	−1,720.47	−1,712.06	−1,715.02	−1,701.10
Values of travel time (Pounds (£)/hour)						
Car (Average)	8.97	9.85	9.97	9.44	9.41	9.33
Commute	13.43	10.28	8.68	13.69	12.82	14.61
Business	20.18	18.41	18.08	18.25	19.35	17.03
Other	7.38	9.22	9.73	7.99	8.08	7.70
Bus (Average)	3.88	3.47	3.50	3.68	3.63	3.80
Commute	4.18	3.75	3.78	3.96	3.92	4.10
Business	5.02	4.47	4.52	4.72	4.68	4.87
Other	3.67	3.29	3.32	3.49	3.45	3.61
Rail (Average)	6.71	6.32	6.41	6.46	6.47	6.69
Commute	4.66	4.39	4.45	4.49	4.50	4.64
Business	15.74	14.84	15.03	15.16	15.18	15.69
Other	6.50	6.13	6.21	6.27	6.28	6.48
MAPE (%)	–	11.00	12.09	5.08	8.15	3.54

(*continued*).

Table 6. Continued.

Coefficients	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio
asc car	0	NA	0	NA	0	NA	0	NA	0	NA	0	0
asc bus	0.26	0.56	0.31	0.63	0.30	0.72	0.23	0.54	0.26	0.62	0.29	0.68
Δ *bus income* > 50k	−1.55	−2.62	−1.64	−2.76	−1.65	−3.47	−1.68	−3.58	−1.68	−3.57	−1.70	−3.61
Δ *bus weekend*	−0.49	−2.24	−0.71	−3.37	−0.71	−3.9	−0.53	−2.72	−0.56	−2.94	−0.54	−2.84
Δ *bus unemployed*	0.08	0.11	0.06	0.09	0.06	0.16	0.05	0.13	0.06	0.15	0.05	0.14
Δ *bus commute*	**0.74**	**3.27**	**0**	**NA**	**0.05**	**0.15**	**0.52**	**2.06**	**0.52**	**1.95**	**0.54**	**2.26**
asc rail	3.88	2.33	3.79	2.22	3.77	2.75	3.77	2.75	3.83	2.80	3.82	2.81
asc taxi	−2.89	−6.12	−2.91	−6.21	−2.92	−7.25	−2.89	−6.43	−2.87	−6.56	−2.87	−6.60
Δ *taxi male*	−0.44	−1.31	−0.44	−1.32	−0.44	−1.60	−0.44	−1.60	−0.44	−1.60	−0.44	−1.59
Δ *taxi 18to24 age*	1.88	5.21	1.92	5.38	1.92	7.35	1.88	7.21	1.88	7.21	1.88	7.23
Δ *taxi 25to29 age*	1.16	2.53	1.19	2.56	1.19	3.23	1.16	3.16	1.16	3.15	1.16	3.16
Δ *taxi lower education*	−3.02	−2.30	−3.11	−2.29	−3.11	−2.54	−3.07	−2.52	−3.06	−2.51	−3.07	−2.51
Δ *taxi income 40to50k*	−0.92	−1.96	−0.88	−1.85	−0.88	−2.26	−0.89	−2.30	−0.89	−2.29	−0.89	−2.30
asc cycling	−4.80	−7.17	−4.90	−7.57	−4.92	−12.11	−4.85	−11.17	−4.83	−11.33	−4.83	−11.27
Δ *cycling male*	1.73	3.52	1.61	3.36	1.61	5.78	1.64	5.77	1.64	5.75	1.66	5.82
Δ *cycling business*	**−1.52**	**−2.37**	**0**	**NA**	**0.03**	**0.05**	**−0.72**	**−1.11**	**−0.61**	**−0.95**	**−0.91**	**−1.34**
Δ *cycling income 10to20k*	0.91	1.98	0.94	2.05	0.94	3.78	0.91	3.58	0.91	3.59	0.91	3.58
Δ *cycling income 75to100k*	3.75	3.44	3.57	3.27	3.58	7.22	3.62	6.82	3.60	6.85	3.62	6.90
Δ *cycling weekend*	−0.94	−2.04	−0.94	−2.14	−0.94	−2.54	−0.88	−2.26	−0.90	−2.34	−0.90	−2.35
Δ *cycling student*	1.03	1.89	1.10	2.07	1.09	3.17	1.05	2.96	1.06	3.00	1.06	2.99
Δ *cycling unemployed*	0.05	0.04	−0.24	−0.24	−0.24	−0.29	−0.28	−0.32	−0.26	−0.31	−0.29	−0.34
asc walking	2.29	3.75	2.41	4.41	2.38	5.16	2.36	4.21	2.35	4.42	2.37	4.39
Δ *walking 18to29 age*	0.71	2.61	0.70	2.56	0.70	3.77	0.72	3.84	0.72	3.82	0.72	3.83
Δ *walking lower education*	−0.71	−2.86	−0.71	−2.82	−0.71	−4.08	−0.70	−4.00	−0.70	−4.04	−0.70	−4.03
Δ *walking commute*	**0.81**	**2.96**	**0**	**NA**	**0.10**	**0.22**	**0.74**	**1.99**	**0.73**	**1.82**	**0.70**	**1.98**
Δ *walking weekend*	−0.23	−0.84	−0.42	−1.55	−0.42	−1.80	−0.22	−0.88	−0.26	−1.04	−0.25	−1.05
Δ *walking student*	0.66	1.98	0.60	1.78	0.60	2.65	0.65	2.84	0.63	2.78	0.64	2.81
β *car time*	−0.16	−2.09	−0.21	−2.37	−0.21	−2.88	−0.19	−2.64	−0.18	−2.69	−0.18	−2.70
Δ *car time commute*	**−0.09**	**−1.77**	**0**	**NA**	**0.04**	**0.76**	**−0.09**	**−1.34**	**−0.07**	**−1.16**	**−0.07**	**−1.36**
Δ *car time business*	**−0.05**	**−1.26**	**0**	**NA**	**0.01**	**0.26**	**−0.03**	**−0.42**	**−0.04**	**−0.60**	**−0.01**	**−0.17**
Δ *car time am peak*	−0.06	−1.34	−0.11	−1.99	−0.11	−2.59	−0.08	−2.06	−0.08	−2.23	−0.07	−1.96

(*continued*).

Table 6. Continued.

Coefficients	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio	Estimate	t-ratio
Δ *car time pm peak*	0.14	2.49	0.17	2.84	0.17	2.69	0.16	2.48	0.16	2.57	0.16	2.48
β *bus time*	−0.09	−2.24	−0.09	−2.19	−0.09	−2.72	−0.09	−2.57	−0.09	−2.60	−0.09	−2.58
Δ *bus time am pm peak*	−0.04	−1.83	−0.04	−1.78	−0.04	−2.28	−0.04	−2.32	−0.04	−2.32	−0.04	−2.33
β *train time*	−0.02	−1.55	−0.02	−1.66	−0.02	−2.34	−0.02	−2.22	−0.02	−2.28	−0.02	−2.25
β *taxi time*	−0.26	−2.28	−0.27	−2.25	−0.27	−2.86	−0.29	−2.65	−0.27	−2.71	−0.27	−2.69
Δ *taxi time pm peak*	−0.03	−0.45	−0.05	−0.71	−0.05	−0.93	−0.04	−0.75	−0.04	−0.77	−0.04	−0.72
Δ *taxi time am peak*	−0.10	−1.73	−0.09	−1.58	−0.09	−1.80	−0.10	−1.89	−0.10	−1.87	−0.10	−1.89
β *cycling time*	−0.19	−2.22	−0.20	−2.22	−0.20	−3.03	−0.21	−2.77	−0.21	−2.84	−0.20	−2.78
β *walking time*	−0.43	−2.35	−0.44	−2.43	−0.44	−3.10	−0.47	−2.78	−0.45	−2.87	−0.45	−2.82
λ *travel time*	0.72	5.21	0.71	5.26	0.71	6.85	0.69	5.95	0.70	6.27	0.71	6.17
β *bus out of vehicle time (OVT)*	−1.54	−4.09	−1.50	−3.95	−1.49	−4.44	−1.52	−4.49	−1.51	−4.50	−1.52	−4.50
Δ *bus OVT noIncomeResponse*	−1.75	−4.01	−1.73	−3.88	−1.74	−4.25	−1.75	−4.30	−1.74	−4.30	−1.75	−4.31
β *train OVT*	−2.55	−2.62	−2.50	−2.52	−2.49	−2.99	−2.52	−3.05	−2.51	−3.04	−2.51	−3.05
Δ *train OVT noIncomeResponse*	−2.29	−2.38	−2.27	−2.30	−2.27	−2.77	−2.28	−2.81	−2.27	−2.81	−2.27	−2.82
λ *OVT*	0.01	0.03	0.01	0.09	0.01	0.12	0.01	0.08	0.01	0.09	0.01	0.08
OVT income elasticity	0.07	2.04	0.07	1.97	0.07	2.49	0.07	2.54	0.07	2.52	0.07	2.54
β *log cost*	−0.77	−8.32	−0.81	−8.76	−0.81	−13.55	−0.79	−12.85	−0.80	−12.95	−0.79	−12.87

less likely to carry out business trips by cycling. It can be observed that there is a statistically significant drop in the mode choice model from a log-likelihood (LL) of -1,689.50 to -1,720.96 once the trip purposes are removed in the mode choice model, indicating that trip purpose is a significant explanator. The models where nearly two-third of the trip purposes are imputed are better in LL compared to the model with no trip purposes. There is a more substantial improvement in the log-likelihood where the trip purposes are imputed using the mixed multinomial logit with posteriors compared with the model where the trip purposes are not present (increase from −1,720.96 to −1,701.10). The model fit of the imputed mode choice models are in-line with the previous section as the best-performing imputation model is the mixed multinomial logit with posteriors followed by the multinomial logit model and mixed multinomial logit model without posteriors. This is intuitive, as better imputation of the trip purposes leads to larger improvements in the mode choice model estimation. Further, we observe that imputing trip purposes with uniform weights is not a good strategy as ideally the trip purposes which are *actually* business or commuting should have a higher probability of being correctly inferred rather than replicating the market shares. The coefficients of the models where the trip purposes are imputed using MNL, MMNL with and without posteriors, are mostly similar in sign and statistical significance compared to the model where the actual trip purposes are observed, indicating no potential biases of imputing trip purposes in a mode choice model.

Table 6 also shows the values of travel time for the different models. It can be observed that in the base model, the values of travel time are intuitive since business trips mostly have a higher VTT followed by commuting and other trips. The values are slightly different than the ones derived using the model developed on the complete dataset (Tsoleridis, Choudhury, and Hess 2022), which is expected as a smaller subset of the data has been used. It can be observed in Table 6 that the most deviation in the VTTs from the base model is in the case of the model where imputation is carried out by uniform weights followed closely by the model with no trip purposes, with MAPE of 12.09% and 11.00%, respectively. The model with the least deviation with respect to the base VTTs is when the trip purposes are inferred using MMNL with posteriors with a MAPE of 3.54%, which is less than half compared to the model with no trip purposes. The model where the trip purposes are imputed using the MNL model also has a lower deviation in VTTs with a MAPE of 5.08%. Based on these results, it can be observed that if the predictive performance of the trip purpose is satisfactory, i.e. in the cases when trip purposes are imputed using the MNL and MMNL with posteriors, it may be beneficial to use the imputed trip purposes rather than to not consider trip purposes at all in the mode choice model. Further, our analysis in this context indicates that it may be possible to achieve reliable values of travel time in a mode choice model where a large proportion of trip purposes are uncertain and inferred.

6. Conclusion

In this study, we propose the use of posterior analysis within the mixed multinomial logit (MMNL) modelling framework to better predict trip purposes of individuals who have tagged some part of their trip purposes but not all. Additionally, we evaluate whether imputing these missing trip purposes enhances the performance of subsequent analyses, such as mode choice modelling. Using a 2-week travel survey diary collected in the UK by a smartphone-based travel survey application, we first observe that only about 15% of the respondents fully tag all of their trips, 10% of respondents do not tag any of their trips at all, while the remaining individuals (about 75% of respondents) do not tag some parts of their trip purposes.

Our results show that the MMNL model with posterior analysis achieves an average probability of 0.537 for correctly inferring the chosen alternative in an out-of-sample testing dataset. This outperforms both the standard MNL model and the MMNL without posterior analysis, which achieve average probabilities of 0.413 and 0.408, respectively. Additionally, MMNL with posteriors also has the least prediction error in predicting aggregate market shares for trip purposes compared to MNL and MMNL without posteriors. Therefore, to predict untagged trip purposes for individuals who have previously tagged some of their trip purposes, researchers are advised to use MMNL with posterior analysis. The results strengthen the findings of Danaf et al. (2019) and Song et al. (2018), who recommend that smartphone-based travel survey applications must make use of posteriors to generate personalised recommendations to infer information about trip details. We further recommend that to gather long-duration data from users, it is substantially better that individuals tag or verify some of their trips rather than skipping verification altogether, as this would enable the use of posterior analysis to better predict untagged trip purposes. We also find that using posterior analysis on a larger number of tagged observations leads to improved trip purpose prediction, indicating that the proposed framework is particularly advantageous for datasets with higher availability of tagged data.

In this study, we also check the added benefit of utilising imputed trip purposes in a subsequent mode choice model, which has not been explored in previous studies. It is observed that the mode choice model where the trip purposes are imputed from the MMNL with posteriors has a better model fit and lower deviation from the true values of travel time compared to the models where the trip purposes are not added as explanatory variables and other methods of imputing trip purposes, i.e. MNL and MMNL without posteriors. In fact, the mean absolute percentage error in the values of travel time where the trip purposes are imputed using MMNL with posteriors is less than half compared to the model where the trip purposes are not added as explanatory variables (3.54% vs 11.00%). Our study therefore demonstrates that it is better to impute trip purposes in a mode choice model to generate reliable values of travel time rather than not to use trip purposes as explanatory variables at all. This finding is particularly relevant for researchers who work with smartphone-based travel surveys or even other passively collected datasets, as there may be benefits in utilising imputed trip purposes in downstream models, such as mode choice models. While these findings are case-specific and difficult to generalise across other cases, future research should examine the robustness of this finding in additional case studies. Nonetheless, the proposed framework for estimating a mode choice model with uncertain trip purposes by weighting the likelihood will be useful to researchers who encounter uncertain trip purposes or other uncertain discrete independent explanatory variables in their models.

There are several directions in which this research can be extended. Currently, we model and predict missing trip purposes independently, i.e. do not consider the impact of the preceding trip purpose. For future work, we want to consider a dynamic approach to model trip purposes that could better incorporate the structure of daily activity–travel chains. Further improvements can be achieved in the trip purpose model by a more detailed specification of trip start times and the inclusion of additional explanatory variables such as time spent between the trips, workplace location. In this study, we evaluate our approach on a single dataset and assess robustness by varying the proportion of tagged data. Nonetheless, applying the method to additional datasets would further strengthen the findings. While we employ an inverse Wishart distribution as a prior for the covariance matrix for the mixed multinomial logit model, alternative specifications, such as the LKJ or scaled-inverse priors, could be explored for potentially improved estimation (Akinc and Vandebroek 2018). In our study, we examine the impact of imputing trip purposes in a mode choice model using a weighted likelihood approach. Future research could extend the analysis to cases where both modes and purposes are untagged and assess the impact on different downstream models, such as route choice, activity generation, and gravitational flow models.

Note

1. The results for the multinomial logit model are also similar and presented in Appendix - Table A1

Acknowledgments

The authors thank Dr. Panagiotis Tsoleridis for sharing the cleaned and processed DECISIONS data, along with mode choice and value of travel time enumeration scripts.

Author contributions

Azam Ali: Conceptualisation, Methodology, Formal analysis, Software, Visualisation, Writing–original draft; **Stephane Hess**: Supervision, Conceptualisation, Methodology, Funding acquisition, Software, Project administration, Writing–review & editing; **Thijs Dekker**: Conceptualisation, Methodology, Writing–review & editing, **Charisma F. Choudhury**: Supervision, Conceptualisation, Methodology, Funding acquisition, Project administration, Writing–review & editing.

Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This work was supported by the UK Economic Social Research Council AQM Grant Reference No ES/P000746/1. Charisma Choudhury's time was supported by the UKRI Future Leaders Fellowship Next Generation Travel Behaviour Models' [Grant Number MR/T020423/1] and Stephane Hess acknowledges support from the European Research Council through advanced Grant 101020940-SYNERGY.

ORCID

Azam Ali http://orcid.org/0009-0009-6747-5012
Thijs Dekker https://orcid.org/0000-0003-2313-8419
Charisma F. Choudhury https://orcid.org/0000-0002-8886-8976

References

Akinc, D., and M. Vandebroek. 2018. “Bayesian Estimation of Mixed Logit Models: Selecting an Appropriate Prior for the Covariance Matrix.” *Journal of Choice Modelling* 29:133–151. https://doi.org/10.1016/j.jocm.2017.11.004.

Ali, A., E. Flaata, T. Torset, S. Hess, and C. Choudhury. 2026. “Investigating the impact of inferring trip purposes in a daily trip generation model.” *Transportation Research Part C: Emerging Technologies* 186: 105619. https://doi.org/10.1016/j.trc.2026.105619.

Ali, A., A. Kalatian, and C. F. Choudhury. 2023. “Comparing and Contrasting Choice Model and Machine Learning Techniques in the Context of Vehicle Ownership Decisions..” *Transportation Research Part A: Policy and Practice* 173: 103727. https://doi.org/10.1016/j.tra.2023.103727.

Amaya, M., R. Cruzat, and M. A. Munizaga. 2018. “Estimating the Residence Zone of Frequent Public Transport Users to Make Travel Pattern and Time Use Analysis.” *Journal of Transport Geography*66:330–339. https://doi.org/10.1016/j.jtrangeo.2017.10.017.

Andersson, A., L. Engelson, M. Börjesson, A. Daly, and I. Kristoffersson. 2022. “Long-Distance Mode Choice Model Estimation Using Mobile Phone Network Data.” *Journal of Choice Modelling*42:100337. https://doi.org/10.1016/j.jocm.2021.100337.

Arentze, T. A., D. Ettema, and H. J. Timmermans. 2013. “Location Choice in the Context of Multi-day Activity-Travel Patterns: Model Development and Empirical Results.” *Transportmetrica A: Transport Science* 9 (2): 107–123. https://doi.org/10.1080/18128602.2010.538870.

Bierlaire, M. 2003. “Biogeme: A Free Package for the Estimation of Discrete Choice Models.” In *Proceedings of the 3rd Swiss Transportation Research Conference*, 19–21. Ascona, Switzerland.

Biswas, M., C. R. Bhat, S. Ghosh, and A. R. Pinjari. 2024. “Choice Models with Stochastic Variables and Random Coefficients.” *Journal of Choice Modelling* 51:100488. https://doi.org/10.1016/j.jocm.2024.100488.

Bohte, W., and K. Maat. 2009. “Deriving and Validating Trip Purposes and Travel Modes for Multi-day GPS-based Travel Surveys: A Large-Scale Application in the Netherlands.” *Transportation Research Part C: Emerging Technologies* 17 (3): 285–297. https://doi.org/10.1016/j.trc.2008.11.004.

Bwambale, A., C. F. Choudhury, and S. Hess. 2019. “Modelling Trip Generation Using Mobile Phone Data: A Latent Demographics Approach.” *Journal of Transport Geography* 76:276–286. https://doi.org/10.1016/j.jtrangeo.2017.08.020.

Calastri, C., R. Crastes dit Sourd, and S. Hess. 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 (1): 175–201. https://doi.org/10.1007/s11116-018-9858-7.

Campbell, D., W. G. Hutchinson, and R. Scarpa. 2006. *Using Discrete Choice Experiments to Derive Individual-Specific Wtp Estimates for Landscape Improvements under Agri-Environmental Schemes: Evidence from the Rural Environment Protection Scheme in Ireland*. Technical Report 12220, Fondazione Eni Enrico Mattei (FEEM).

Cottrill, C. D., F. C. Pereira, F. Zhao, I. F. Dias, H. B. Lim, M. E. Ben-Akiva, and P. C. Zegras. 2013. “Future Mobility Survey: Experience in Developing a Smartphone-Based Travel Survey in Singapore.” *Transportation Research Record: Journal of the Transportation Research Board* 2354 (1): 59–67. https://doi.org/10.3141/2354-07.

Cui, Y., C. Meng, Q. He, and J. Gao. 2018. “Forecasting Current and Next Trip Purpose with Social Media Data and Google Places.” *Transportation Research Part C: Emerging Technologies* 97:159–174. https://doi.org/10.1016/j.trc.2018.10.017.

Danaf, M., F. Becker, X. Song, B. Atasoy, and M. Ben-Akiva. 2019. “Online Discrete Choice Models: Applications in Personalized Recommendations.” *Decision Support Systems* 119:35–45. https://doi.org/10.1016/j.dss.2019.02.003.

Deng, Z., and M. Ji. 2010. “Deriving Rules for Trip Purpose Identification from GPS Travel Survey Data and Land Use Data: A Machine Learning Approach.” In *Traffic and Transportation Studies 2010*, 768–777. Kunming, China: American Society of Civil Engineers.

Dumont, J., M. Giergiczny, and S. Hess. 2015. “Individual Level Models Vs. Sample Level Models: Contrasts and Mutual Benefits.” *Transportmetrica A: Transport Science* 11 (6): 465–483. https://doi.org/10.1080/23249935.2015.1018681.

Ectors, W., S. Reumers, W. D. Lee, K. Choi, B. Kochan, D. Janssens, T. Bellemans, and G. Wets. 2017. “Developing an Optimised Activity Type Annotation Method Based on Classification Accuracy and

Entropy Indices." *Transportmetrica A: Transport Science* 13 (8): 742–766. https://doi.org/10.1080/23249935.2017.1331275.

Ermagun, A., Y. Fan, J. Wolfson, G. Adomavicius, and K. Das. 2017. "Real-Time Trip Purpose Prediction Using Online Location-Based Search and Discovery Services." *Transportation Research Part C: Emerging Technologies* 77:96–112. https://doi.org/10.1016/j.trc.2017.01.020.

Faghih Imani, A., C. Harding, S. Srikukenthiran, E. J. Miller, and K. Nurul Habib. 2020. "Lessons from a Large-Scale Experiment on the Use of Smartphone Apps to Collect Travel Diary Data: The 'City Logger' for the Greater Golden Horseshoe Area." *Transportation Research Record: Journal of the Transportation Research Board* 2674 (7): 299–311. https://doi.org/10.1177/0361198120921860.

Fu, X., Y. Zhang, J. d. D. Ortúzar, and G. Lü. 2025. "Activity-Travel Pattern Inference Based on Multi-source Big Data." *Transport Reviews* 45 (1): 26–48. https://doi.org/10.1080/01441647.2024.2400341.

Gao, Q., J. Molloy, and K. W. Axhausen. 2021. "Trip Purpose Imputation Using GPS Trajectories with Machine Learning." *ISPRS International Journal of Geo-Information* 10 (11): 775. https://doi.org/10.3390/ijgi10110775.

Gong, L., R. Kanamori, and T. Yamamoto. 2018. "Data Selection in Machine Learning for Identifying Trip Purposes and Travel Modes from Longitudinal GPS Data Collection Lasting for Seasons." *Travel Behaviour and Society* 11:131–140. https://doi.org/10.1016/j.tbs.2017.03.004.

Harding, C., A. Faghih Imani, S. Srikukenthiran, E. J. Miller, and K. Nurul Habib. 2021. "Are We There Yet? Assessing Smartphone Apps as Full-Fledged Tools for Activity-Travel Surveys." *Transportation* 48 (5): 2433–2460. https://doi.org/10.1007/s11116-020-10135-7.

Heinonen, S., L. M. de Freitas, A. Meister, L. Schwab, J. Roth, B. Hintermann, T. Götschi, and K. Axhausen. 2024. "The E-Biking in Switzerland (ebis) Study: Methods and Dataset." *Transportation* 1–24. https://doi.org/10.1007/s11116-024-10552-y.

Hensher, D. A., W. H. Greene, and J. M. Rose. 2006. "Deriving Willingness-to-pay Estimates of Travel-Time Savings from Individual-Based Parameters." *Environment and Planning A* 38 (12): 2365–2376. https://doi.org/10.1068/a37395.

Hess, S. 2007. "Posterior Analysis of Random Taste Coefficients in Air Travel Behaviour Modelling." *Journal of Air Transport Management*13 (4): 203–212. https://doi.org/10.1016/j.jairtraman.2007.03.004.

Hess, S., A. Daly, T. Dekker, M. O. Cabral, and R. Batley. 2017. "A Framework for Capturing Heterogeneity, Heteroskedasticity, Non-linearity, Reference Dependence and Design Artefacts in Value of Time Research." *Transportation Research Part B: Methodological* 96:126–149. https://doi.org/10.1016/j.trb.2016.11.002.

Hess, S., and D. Palma. 2019. "Apollo: A Flexible, Powerful and Customisable Freeware Package for Choice Model Estimation and Application." *Journal of Choice Modelling* 32:100170. https://doi.org/10.1016/j.jocm.2019.100170.

Hillel, T. 2019. "Understanding Travel Mode Choice: A New Approach for City Scale Simulation." PhD Thesis, University of Cambridge.

Hossain, S., and K. N. Habib. 2021. "Inferring the Purposes of Using Ride-Hailing Services through Data Fusion of Trip Trajectories, Secondary Travel Surveys, and Land Use Data." *Transportation Research Record: Journal of the Transportation Research Board* 2675 (9): 558–573. https://doi.org/10.1177/03611981211003593.

Krueger, R., M. Bierlaire, R. A. Daziano, T. H. Rashidi, and P. Bansal. 2021. "Evaluating the Predictive Abilities of Mixed Logit Models with Unobserved Inter- and Intra-individual Heterogeneity." *Journal of Choice Modelling* 41:100323. https://doi.org/10.1016/j.jocm.2021.100323.

Liao, C., C. Chen, S. Guo, Z. Wang, Y. Liu, K. Xu, and D. Zhang. 2022. "Wheels Know why You Travel: Predicting Trip Purpose via a Dual-Attention Graph Embedding Network." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6 (1): 1–22. https://doi.org/10.1145/3517239.

Liu, Y., E. J. Miller, and K. N. Habib. 2023. "Inferring Trip Destination Purposes for Trip Records Collected through Smartphone Apps." *Journal of Transportation Engineering, Part A: Systems* 149 (2):04022145. https://doi.org/10.1061/JTEPBS.TEENG-7350.

Liu, Z., X. Liu, X. Yan, F. Li, H. Zhong, and Y. Wang. 2026. "Influences of Built Environment on Intercity Travel Demand of Urban Agglomerations: The Discrepancies across Inferred Trip Purposes." *Cities (London, England)* 168:106404.

Lizana, M., C. Choudhury, and D. Watling. 2023. "Using Smart Card Data to Model Public Transport User Profiles in Light of the Covid-19 Pandemic." *Travel Behaviour and Society* 33:100620. https://doi.org/10.1016/j.tbs.2023.100620.

Lu, Y., S. Zhu, and L. Zhang. 2012. "A Machine Learning Approach to Trip Purpose Imputation in GPS-Based Travel Surveys." In *4th Conference on Innovations in Travel Modeling*, Tampa.

Marra, A. D., H. Becker, K. W. Axhausen, and F. Corman. 2019. "Developing a Passive GPS Tracking System to Study Long-Term Travel Behavior." *Transportation Research Part C: Emerging Technologies*104:348–368. https://doi.org/10.1016/j.trc.2019.05.006.

McCool, D., P. Lugtig, O. Mussmann, and B. Schouten. 2021. "An App-Assisted Travel Survey in Official Statistics: Possibilities and Challenges." *Journal of Official Statistics* 37 (1): 149–170. https://doi.org/10.2478/jos-2021-0007.

Meister, A., N. F. Bashan, R. Basu, X. Shen, R. Q. Wang, and A. Sevtsuk. 2025. "The Bostonwalks Study: A Longitudinal Travel Survey Using Smartphone Tracking." *Transportation* 52: 2249–2279. https://doi.org/10.1007/s11116-025-10637-2.

Molloy, J., A. Castro, T. Götschi, B. Schoeman, C. Tchervenkov, U. Tomic, B. Hintermann, and K. W. Axhausen. 2022. "The MOBIS Dataset: A Large GPS Dataset of Mobility Behaviour in Switzerland." *Transportation* 50: 1983–2007. https://doi.org/10.1007/s11116-022-10299-4.

Montini, L., N. Rieser-Schüssler, A. Horni, and K. W. Axhausen. 2014. "Trip Purpose Identification from GPS Tracks." *Transportation Research Record: Journal of the Transportation Research Board*2405 (1): 16–23. https://doi.org/10.3141/2405-03.

Nair, S., K. Javkar, J. Wu, and V. Frias-Martinez. 2019. "Understanding Cycling Trip Purpose and Route Choice Using GPS Traces and Open Data." *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3 (1): 1–26. https://doi.org/10.1145/3314407.

Nurul Habib, K. M., and E. J. Miller. 2009. "Modelling Activity Generation: A Utility-Based Model for Activity-Agenda Formation." *Transportmetrica* 5 (1): 3–23. https://doi.org/10.1080/18128600802591384.

Oliveira, M. G. S., P. Vovsha, J. Wolf, and M. Mitchell. 2014. "Evaluation of Two Methods for Identifying Trip Purpose in GPS-Based Household Travel Surveys." *Transportation Research Record: Journal of the Transportation Research Board* 2405 (1): 33–41. https://doi.org/10.3141/2405-05.

Ordnance Survey. 2023. "Points of Interest Data." Accessed March 29, 2024. https://www.ordnancesurvey.co.uk/products/points-of-interest.

Revelt, D., and K. Train. 1999. "Customer-Specific Taste Parameters and Mixed Logit." Working Paper No E00-274, Economics Working Papers, University of California at Berkeley.

Revelt, D. A. 1999. "Three Discrete Choice Random Coefficients Papers and One Police-Crime Study." PhD Thesis. Berkeley: University of California.

Richter, L.-L., and M. G. Pollitt. 2018. "Which Smart Electricity Service Contracts Will Consumers Accept? The Demand for Compensation in a Platform Market." *Energy Economics* 72:436–450. https://doi.org/10.1016/j.eneco.2018.04.004.

Rossi, P. E., and G. M. Allenby. 2003. "Bayesian Statistics and Marketing." *Marketing Science* 22 (3): 304–328. https://doi.org/10.1287/mksc.22.3.304.17739.

Sanko, N., S. Hess, J. Dumont, and A. Daly. 2014. "Contrasting Imputation with a Latent Variable Approach to Dealing with Missing Income in Choice Models." *Journal of Choice Modelling* 12:47–57. https://doi.org/10.1016/j.jocm.2014.10.001.

Servizi, V., F. C. Pereira, M. K. Anderson, and O. A. Nielsen. 2021. "Mining User Behaviour from Smartphone Data: A Literature Review." *European Transport Research Review* 13 (1): 57. https://doi.org/10.1186/s12544-021-00516-z.

Shen, L., and P. R. Stopher. 2013. "A Process for Trip Purpose Imputation from Global Positioning System Data." *Transportation Research Part C: Emerging Technologies* 36:261–267. https://doi.org/10.1016/j.trc.2013.09.004.

Sillano, M., and J. de Dios Ortúzar. 2005. "Willingness-to-pay Estimation with Mixed Logit Models: Some New Evidence." *Environment and Planning A* 37 (3): 525–550. https://doi.org/10.1068/a36137.

Song, X., M. Danaf, B. Atasoy, and M. Ben-Akiva. 2018. "Personalized Menu Optimization with Preference Updater: A Boston Case Study." *Transportation Research Record: Journal of the Transportation Research Board* 2672 (8): 599–607. https://doi.org/10.1177/0361198118758674.

Tabasi, M., A. Siripanich, N. A. Khan, J. Auld, and T. H. Rashidi. 2024. "Gps-Supported Smartphone App-Based Integrated Travel Diary and Time-Use Data Collection: Challenges and Lessons Learned." *Transportation* 1–26. https://doi.org/10.1007/s11116-024-10546-w.

Tørset, T., and G. B. A. Svaboe. 2020. "Show Case – Using TravelVu and TravelViewer in Trondheim." https://www.ntnu.no/documents/1283830479/1283895465/Show+case+Trondheim.pdf/6107fca9-55b5-ef02-9fb4-2f2dbe1ed994?t=1663068717725.

Train, K. E. 2009. *Discrete Choice Methods with Simulation*. 2nd ed. Cambridge: Cambridge University Press.

Tsoleridis, P., C. F. Choudhury, and S. Hess. 2022. "Deriving Transport Appraisal Values from Emerging Revealed Preference Data." *Transportation Research Part A: Policy and Practice* 165:225–245.

Vij, A., and K. Shankari. 2015. "When Is Big Data Big Enough? Implications of Using GPS-based Surveys for Travel Demand Analysis." *Transportation Research Part C: Emerging Technologies*56:446–462. https://doi.org/10.1016/j.trc.2015.04.025.

Wang, L., W. Ma, Y. Fan, and Z. Zuo. 2019. "Trip Chain Extraction Using Smartphone-Collected Trajectory Data." *Transportmetrica B: Transport Dynamics* 7 (1): 255–274. https://doi.org/10.1080/21680566.2017.1386599.

Wardman, M., V. P. K. Chintakayala, and G. de Jong. 2016. "Values of Travel Time in Europe: Review and Meta-analysis." *Transportation Research Part A: Policy and Practice* 94:93–111.

Winkler, C., A. Meister, and K. W. Axhausen. 2024. "The Timeuse+ Data Set: 4 Weeks of Time Use and Expenditure Data Based on Gps Tracks." *Transportation* 53: 909–935. https://doi.org/10.1007/s11116-024-10517-1.

Xiao, G., Z. Juan, and C. Zhang. 2016. "Detecting Trip Purposes from Smartphone-Based Travel Surveys with Artificial Neural Networks and Particle Swarm Optimization." *Transportation Research Part C: Emerging Technologies* 71:447–463. https://doi.org/10.1016/j.trc.2016.08.008.

Yazdizadeh, A., Z. Patterson, and B. Farooq. 2019. "An Automated Approach from GPS Traces to Complete Trip Information." *International Journal of Transportation Science and Technology* 8 (1): 82–100. https://doi.org/10.1016/j.ijtst.2018.08.003.

Zahroh, S., P. Lugtig, Y. Gootzen, J. Klingwort, and B. Schouten. 2025. *Predicting Trip Purpose in a Smartphone-Based Travel Survey*. Technical Report, Statistics Netherlands (CBS), Henri Faasdreef 312, 2492 JP The Hague, The Netherlands.

Zhao, X., X. Yan, A. Yu, and P. Van Hentenryck. 2020. "Prediction and Behavioral Analysis of Travel Mode Choice: A Comparison of Machine Learning and Logit Models." *Travel Behaviour and Society*20:22–35. https://doi.org/10.1016/j.tbs.2020.02.003.

Appendix

A.1 Sensitivity analysis

In the study, we had assumed a 70–30% split for selecting the training and testing dataset respectively for the trip purpose prediction model. However, to better generalise the findings, a 50-50% and a 60–40% split were also considered, which led the 50–50% split having 12,048 and 12,068 trips in the training and testing data, respectively. Whereas the 60–40% split contained 14,486 and 9,630 trips in the training and testing datasets, respectively. Table A2 shows the predictive performance of the models in the testing dataset for the different data splits. It can be observed that the best predictive performance for both data splits is for the mixed multinomial logit model with posteriors followed by the MNL and MMNL without posteriors in terms of the average probability of the chosen alternative and accuracy. It can be observed that with an increase in the number of observations in the training dataset, the MMNL model with posterior performs better in terms of APCA and lower prediction error for market counts as there is more information (i.e. past observations) to compute the posterior distributions. This suggests that increasing the size of the tagged portion of the dataset can further enhance the benefits of utilising posterior analysis. However, there is no difference in the predictive performance of MNL and MMNL without posteriors in the data splits. The aggregate market counts are slightly better captured in the MNL model when a 50–50% split is assumed compared to the MMNL model with posteriors; however, the MMNL with posteriors has a significantly higher average probability of correctly inferring the correct alternative.

Table A1. Estimates for multinomial logit model for trip purpose prediction (home set as base).

	Commute		Business		Leisure		Shopping	
Parameter	Estimate	Rob t ratio	Estimate	Rob t ratio	Estimate	Rob t ratio	Estimate	Rob t ratio
asc	−7.35	−89.53	−6.37	−76.58	−0.46	−1.39	−1.59	−13.39
$\beta_{\texttt{weekday}}$	2.10	16.14	1.78	10.00	−0.43	−6.66	0.05	0.08
$\beta_{\texttt{fully employed}}$	1.23	8.63	1.22	6.15	–	–	–	–
$\beta_{\texttt{distance}}$	0.65	14.81	0.59	12.44	0.42	15.76	0.36	16.99
$\beta_{\texttt{time 0 to 9hrs}}$	4.29	36.78	3.36	29.40	0.56	3.68	1.91	20.19
$\beta_{\texttt{time 9 to 16hrs}}$	2.38	15.36	2.96	18.08	0.38	3.70	1.37	18.21
$\beta_{\texttt{time 16 to 19hrs}}$	0.22	0.08	0.91	3.00	−0.07	−0.67	0.51	4.86
$\beta_{\texttt{restaurants}}$	0.10	0.36	0.08	0.27	0.15	0.70	0.10	0.52
$\beta_{\texttt{services}}$	0.05	0.24	0.01	0.01	−0.03	−0.09	0.03	0.12
$\beta_{\texttt{transportation}}$	–	–	–	–	0.14	0.01	–	–
$\beta_{\texttt{shops offices}}$	0.07	0.16	−0.01	−0.03	0.11	0.47	0.20	9.65
$\beta_{\texttt{entertainment}}$	0.26	5.11	0.09	1.78	0.17	4.56	0.03	0.04
$\beta_{\texttt{industrial}}$	0.15	3.77	0.05	0.13	–	–	–	–
$\beta_{\texttt{green}}$	–	–	–	–	0.32	2.78	0.39	3.48
$\beta_{\texttt{education}}$	–	–	–	–	–	–	0.11	2.84

Table A2. Predictive performance for trip purpose prediction model with different data splits in testing data.

		Mixed multinomial logit	
Metric	Multinomial logit	Without posteriors	With posteriors
	50–50% split		
APCA	0.413	0.406	0.526
Accuracy (%)	54.27	44.15	60.89
MAPE for market counts (%)	2.06	6.61	2.46
	60–40% split		
APCA	0.415	0.406	0.532
Accuracy (%)	54.94	44.23	62.05
MAPE for market counts (%)	2.43	5.91	1.92