Estimating Simplified Main Tour Mode Using Revealed Preference and Stated Preference Data

Final Report

September 2025

Estimating Simplified Main Tour Mode Using Revealed Preference and Stated Preference Data

[Comments]

Prepared by

Brett Smith, Doina Olaru, Tristan Reed, Hugo Nilsson and Max Davidson

Keywords: Tour based choice models, SP mode choice, household travel survey

Version control

Final (1.0)

Project No

iMOVE Project 1-081, Milestone 3

Project Steering Committee

Renlong Han, Department of Transport
Sharif Siddique, Department of Transport
Wes Soet, Main Roads WA
Sharon Biermann (Director PATREC, UWA)
Renae Leeson, iMOVE

Acknowledgments

This research is funded by PATREC and the iMOVE CRC and supported by the Cooperative Research Centres program, an Australian Government initiative.

About PATREC

The Planning and Transport Research Centre (PATREC) is a collaboration between the Government of Western Australia and local universities, constituted to conduct collaborative, applied research and teaching in support of policy in the connected spaces of transport and land use planning. The collaborating parties are: The University of Western Australia, Curtin University, Edith Cowan University, Department of Transport, Main Roads Western Australia, Western Australian Planning Commission and the Western Australian Local Government Association.

Publisher

Planning and Transport Research Centre
The University of Western Australia (M087)
35 Stirling Highway, Crawley, WA, 6009

Executive Summary

Project Context and Purpose

This project set out to develop a tour-based mode choice model for Perth, grounded in both Revealed Preference (RP) and Stated Preference (SP) data collected through the Perth Area Travel and Household Survey (PATHS, 2018–2021). The work was commissioned to update the Strategic Transport Evaluation Model (STEM) currently used by the WA Department of Transport, which had been calibrated using outdated data from between 2002 and 2006. The main objective was to provide behaviourally realistic models of travel demand that reflect Perth's unique transport context, while testing the contribution of SP data alongside RP evidence.

Summary and Context

This research aims to evaluate the Revealed Preference (RP) and Stated Preference (SP) data collected as part of the Perth Area Travel and Household Survey (PATHS), conducted between 2018 and 2021, with the objective of assisting trip chain/tour modelling. The purpose of the analysis presented in this report was to assess the extent to which the SP data can inform the WA Department of Transport's Strategic Transport Evaluation Model (STEM) in updating the mode choice component for Perth.

A current review and update of the literature has been undertaken and identified that tour-based mode choice models balance complexity, computational feasibility, and behavioural realism. Despite the low frequency of multimodal tours, the team planned to examine Conditional Trip-Level Mode Choice or Dynamic Discrete Choice Models as candidate models for complex tours. It should be noted that observed multimodal tours account for only 5.3% of all tours and complex tours themselves account for fewer than one-third of tours. One way to account for the lower shares of multimodal tours is to first model the probability of being a complex tour and then estimate mode choice models conditional on the selection probabilities.

Overview of Travel Behaviour and Tours

The PATHS survey reveals that nearly 90% of tours rely on motorised modes, with private vehicles accounting for 81%. Public transport use remains modest: Bus (3%), Train (2%), and Taxi/Rideshare (1%). Active modes are present (13% of tours), dominated by Walking (11%) and Cycling (2%). Tour distances show stark contrasts: Walking (1 km), Cycling (10 km), Car (23 km), Bus (19 km), Taxi (23 km), and Train (51 km).

Tour complexity remains limited: 67.6% are single-purpose tours, with only 5.3% being multimodal. Car dominates complex multi-activity chaining, while public transport plays a limited role in multi-activity tours. This reinforces the structural reliance on cars in Perth and highlights barriers to multimodal integration.

Modelling Insights (RP, SP, Joint, Donor)

The modelling exercise assessed RP, SP, joint RP/SP, and donor approaches. RP models derived from PATHS data produce behaviourally credible parameters, showing strong sensitivity to parking costs, transfers, and access/egress. SP models, while useful for hypothetical policy checks, are weakened by design issues: pivoting inconsistencies, unrealistic attribute values, and more than 20% of observations excluded as outliers. This hinders integration of SP with RP for reliable calibration.

RP data modelling also required imputing attributes for non-chosen alternatives (e.g., travel times and costs), which introduced potential errors, especially at the tour level. Analysing SP independently confirmed the directional importance of travel times, costs, transfers, and access, but the explanatory power was weak. Including RP information in SP models highlighted the influence of past behaviour (mode inertia and state dependence). The number of tours and total daily travel time displayed opposite effects on mode choice, suggesting heterogeneity in travel strategies.

Given limited tour complexity, a practical approach is to model the first trip of each tour. This simplified tour mode choice yielded clearer behaviour and greater reliability for strategic transport modelling. Joint RP/SP estimation added insight, but RP models remain the most reliable foundation. Local context is critical: Perth's dispersed form, strong car dependence, and evolving PT network mean that generic donor models calibrated elsewhere are unsuitable.

PATHS vs Donor Model

The 'Donor model', which used parameter restrictions and nesting structures calibrated in other cities, was found to be inappropriate for Perth. It imposed structural assumptions that did not align with local behaviour, for example, treating bus and rail as a single elemental mode and classifying Park-and-Ride (PnR) under the transit nest. PATHS data revealed that rail and bus serve distinct markets in Perth and that PnR shares more unobserved characteristics with private vehicles. Moreover, the Donor model's small time coefficients suggest low behavioural responses to travel time savings, which would underestimate the benefits of infrastructure investments such as new rail lines or frequency improvements.

The following table reports parameter estimates from PATHS operational mode and the Donor-constrained model, as reported in Section 11, Tables 28 and 29.

Table E1: Comparison of PATHS vs Donor disutility of travel time

Category	**PATHS (Model 1, Operational)**	**Donor-Constrained (Model 3)**	**Implications**
In-vehicle Time (Car)	–0.061 (HBW) / –0.058 (HBO)	–0.025 (fixed)	Donor model underestimates car time sensitivity.
In-vehicle Time (PT)	–0.052 (Bus/Rail HBW)	–0.025 (fixed)	Donor implies low responsiveness to PT travel time improvements.
Access/Egress Time (PT)	–0.054 to –0.061	–0.075 (fixed, combined with wait)	Donor specification conflates access and waiting, masking true deterrence of poor access.
Wait Time (PT)	–0.049 to –0.052	–0.075 (fixed, combined with walk)	Donor overstates waiting disutility but fails to separate waiting from walking.
Transfers (PT)	–0.606 to –0.919	–0.250 (fixed)	PATHS indicates strong aversion to transfers; donor downplays this effect.

Category	PATHS (Model 1, Operational)	Donor-Constrained (Model 3)	Implications
Bike Travel Time	–0.051 to –0.063	–0.063 (fixed)	PATHS estimates stronger gradients; donor holds coefficient constant.
Walk Travel Time	–0.110 to –0.191	–0.063 (fixed)	PATHS shows steep disutility of walking beyond short trips; donor underestimates impact.

Time Sensitivity and Values of Time

A central finding concerns the treatment of time and cost. The ‘Donor model', with small coefficients, yields low values of time (VOT), underestimating behavioural response to travel time savings. PATHS models provide larger, mode-specific VOTs consistent with international evidence and Perth’s conditions. For example, PATHS estimates are: Car IVT ($31.79/hr), Bus IVT ($14–19/hr), Rail IVT ($15–16/hr), and Taxi IVT ($51/hr). PT access/egress time and waiting were valued sharply ($18–42/hr), and transfer penalties ranged $5–9. By contrast, donor IVT averaged ~$19/hr across quartiles, with unstable ranges (e.g., PT wait $14–120/hr).

Willingness-to-pay analysis reinforces this: commuters place a higher value on PT accessibility and reliability than on in-vehicle time (Figure E1). Parking costs were shown to be a major lever, exerting stronger behavioural deterrence than vehicle running costs. This underpins the importance of parking management in constrained centres.

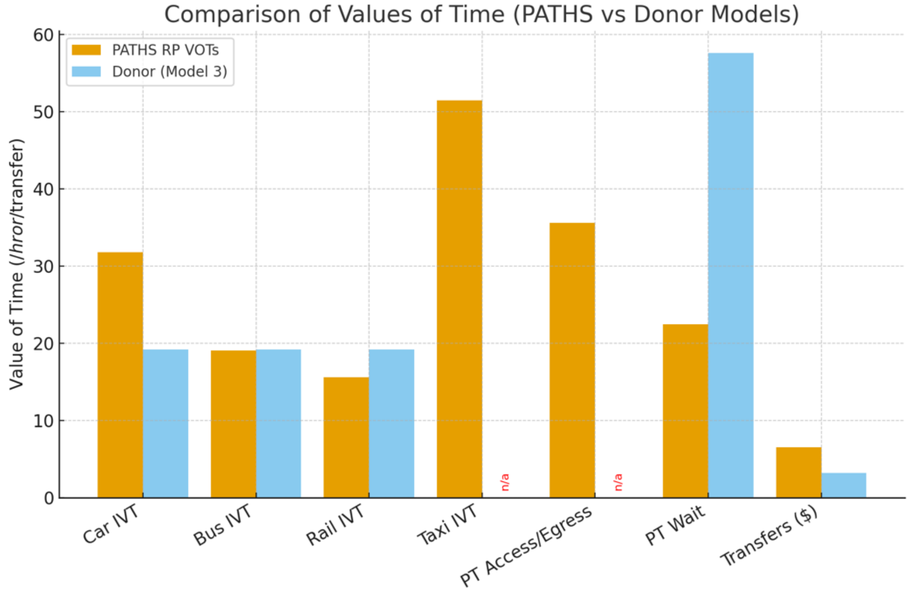


Figure E1: Comparison of PATHS vs Donor Values of Time

Policy Implications

- Car time is highly valued, meaning congestion relief and efficiency projects deliver significant benefits.
- Public transport access and transfers are strongly penalised, justifying investment in feeder services, station access, and direct services.
- Parking costs are a major lever for mode shift, more influential than vehicle running costs.
- SP provides directional checks but cannot replace RP evidence for calibration.
- Donor models underestimate behavioural responses and misrepresent Perth's mode structures.
- Tour-based modelling in this context is complicated by low multimodality; simplified tour models focusing on first trips are more robust.

Recommendations

Consider models estimated on PATHS RP-based as the foundation for STEM calibration and policy evaluation.
Models built on locally observed RP data provide the most behaviourally realistic basis for STEM. They capture the actual trade-offs Perth travellers make between cost, time, access, and reliability. Using these models for calibration ensures that policy evaluation and demand forecasting reflect Perth's unique context, rather than relying on assumptions imported from other cities.

Question the donor-constrained models for Perth, given their structural and behavioural implausibility.
Donor models, while attractive for their parsimony, impose structural assumptions that do not fit Perth's network or traveller behaviour. By grouping rail and bus as a single mode, misclassifying PnR, and underestimating time sensitivity, they risk misrepresenting travel choices and undervaluing infrastructure investments. Policymakers should be cautious about their use and instead prioritise Perth-specific evidence.

Incorporate mode-specific values of time (Car, Bus, Rail, Taxi, PT access, waiting, transfers) into appraisal frameworks.
The PATHS analysis highlights that not all minutes are valued equally: waiting, transfers, and access time carry far higher penalties than in-vehicle time. Appraisal frameworks should reflect these distinctions to ensure benefits from service reliability, direct routing, and improved access are properly valued alongside traditional in-vehicle time savings.

Reclassify PnR as Car-like rather than Transit-like in nesting structures.
Empirical results show that PnR shares more unobserved characteristics with car travel than with transit. Retaining it within the car nest better reflects its role as a hybrid access mode. This adjustment will improve model structure, ensuring that PnR demand forecasts are more consistent with observed Perth behaviour.

Consider parking pricing as a central policy lever to manage demand.
The modelling results demonstrate that parking costs have a stronger deterrent effect than vehicle operating costs. This underscores parking management as a highly effective lever for influencing mode choice. Policy measures such as differentiated parking charges in activity centres could reduce car dominance and shift demand towards public and active transport.

Integrate socio-demographic variables (e.g., household vehicles, licence holding, age, gender) into operational STEM models to reflect traveller heterogeneity.
Mode choice behaviour varies significantly across household and personal characteristics. Incorporating socio-demographics into STEM will better capture heterogeneity in mode preferences and constraints, improving both forecasting accuracy and the capacity to evaluate equity impacts of transport policies.

Pay more attention to customising SP design in future to provide more realistic sensitivity estimates.
The SP experiment design used in PATHS excluded over 20% of observations due to unrealistic attribute combinations. Future SP surveys should be carefully customised to Perth's travel context, ensuring attribute ranges are realistic and choice tasks remain credible to respondents. This will provide more reliable sensitivity estimates that complement RP data.

Conclusion

The PATHS dataset offers a robust empirical foundation for building a mode choice model for use in STEM. RP models outperform SP and donor approaches in explanatory power, behavioural realism, and policy relevance. Donor models not only misrepresent Perth's mode structures (e.g., conflating bus and rail, misclassifying PnR), but also produce low values of time for private vehicles. SP models are useful for scenario checks, but unsuitable for calibration due to design issues. Ultimately, transport investment in Perth should be guided by PATHS-derived RP models, which capture the realities of Perth's car dependence, PT challenges, and urban form, while offering clear evidence on the behavioural levers most relevant for policy.

Table of Contents

1 Introduction

The Perth Area Travel and Household Survey (PATHS) provides valuable data for estimating a tour-based mode choice model. Compared to a trip-based transport model, a tour-based (or activity-based) transport model is behaviourally more realistic and is expected to provide better forecasts for policy or infrastructure initiatives. Recognising this, the Department of Transport (DoT) of Western Australia has decided to develop and implement a tour-based strategic transport model. This project aims to estimate a tour-based mode choice model on the revealed and stated preference data collected by PATHS.

This report details the work undertaken towards this aim; namely a literature review of existing studies using Revealed Preference (RP) and Stated Preference (SP) data to estimate a tour-based mode choice model and related work, alongside exploration of the RP and then the SP datasets to understand their nature before using them as inputs to choice models for trips and tours.

1.1 Project significance

The economic contribution of transport infrastructure projects and travel demand management policies are entirely dependent on the demand response induced by the project. It is expected that strategic transport models can accurately forecast the demand response to government expenditure and policy.

The WA Department of Transport (DoT)'s existing Strategic Transport Evaluation Model (STEM) is calibrated to outdated data from the 2002-2006 Perth and Regions Travel Survey household travel survey. The current model is trip-based and predates the modelling review undertaken by PATREC in 2014, recommending a tour-based model to be implemented, with the objective of providing a more realistic representation of household travel choices. To address this, a pressing need was seen to update the mode choice component using Perth's most recent household and travel survey. New estimates based on this up-to-date dataset are expected to capture changes in households and offer valuable insights into any changes in travel behaviour and mode choices over time, such as increased Working-from-Home (WFH) instigated by COVID-19 and new modes such as e-rideables. Utilising a tour-based model with this dataset would provide a richer consideration of travel behaviour patterns, compared to the existing tour-based model.

The accuracy and currency of travel behaviour estimation are crucial for establishing the credibility of a transport model. The use of the PATHS Household Travel Survey, including revealed preference (RP) and stated preference (SP) data, is particularly valuable for transport modelling, including mode choice modelling. The SP survey conducted by the State Government for the entire Perth metropolitan area was seen as a unique and highly useful resource for developing mode choice models. Considering the recency of the data, this project commenced in 2023, to ensure that the insights reflect the most relevant and reliable information about travel behaviour in Perth.

We recognise that local context plays a critical role in shaping the utility structure and parameter estimates of travel behaviour models. Reliable mode choice models must be grounded in the observed behaviour of Perth's residents. The city's distinctive car dependence, dispersed urban form, and evolving public transport network mean that generic models, or those developed for other cities, often fail to capture important local distinctions. Two defining features of Perth's rail network are the length and speed of the Butler and Mandurah lines, which, when coupled with extensive park-and-ride facilities, make rail a viable option for travel over longer distances.

This report interrogates the PATHS data with the primary aim of developing robust tour-based mode choice models, enabling the transport agency to compare our results with existing models currently applied in Perth's strategic transport modelling. A secondary objective is to critically assess the usefulness of the stated choice data and its contribution to enhancing the estimation of time and cost parameters, which are essential for the economic evaluation of transport policy and infrastructure investments.

1.2 Project objectives

The overall objective of this project was to be able to provide more realistic and better travel choice modelling to inform and influence better policy and infrastructure decisions. This was achieved through the overarching project aim, which is to estimate a tour-based mode choice model using the RP and SP data collected by the PATHS. To achieve this aim, the project has:

- profiled household travel patterns and identified the common home-based and non-home-based tours undertaken by Perth households;
- examined the quality and usability of the SP data, by estimating stand-alone SP choice models and assessing the reasonableness of behavioural outputs: the value of travel time savings (in-vehicle and out-of-vehicle); value of transfers; perception of parking cost compared to perception of expenditure on fuel or fares;
- estimated a tour-based mode choice model on SP data, augmented by RP, including information on the daily tours;
- estimated tour-based RP mode choice models, considering imputed data for the choice set and attributes; and finally,
- estimated a joint RP and SP mode choice model.

2 Literature review

The way people travel is complex and travel patterns vary greatly over households. Activity-based modelling offers a more accurate representation by capturing the motivations behind trips and the intricate connections between daily activities, but like real behaviour, these models are complex and may prove to be impractical (Bowman & Ben-Akiva, 2001; Buliung & Kanaroglu, 2007; Rasouli & Timmermans, 2013; Tajaddini et al., 2020). Over the past 60 years transport academics and practitioners, alike, have sought to capture the essence of household travel by segmenting travel patterns into components, such as trips. Yet, despite the need to simplify household decisions, there has been significant progress towards the specification of functioning activity-based transport models, and one such “compromising” solution is the tour-based model (Hasnine & Habib, 2020; Miller et al., 2005).

The overarching purpose of this literature review is to examine the advances in transport modelling and its progress towards closer representations of household activity decisions and resulting travel patterns; paying close attention to where the models have been used in practice for strategic decision-making and policy evaluation. The tight scope is to position tour-based models along the spectrum of ‘practicality’ and ‘representativeness’.

2.1 Activity models for transport

Tracing the origins of activity-based models can be somewhat ambitious, but three key research streams are notably relevant. Firstly, time geography, introduced by Hägerstrand (1970), conceptualises travel as a derived demand resulting from the need to engage in activities across different locations. The time required to travel imposes constraints, limiting accessible destinations. Hägerstrand's space-time prisms have notably influenced the measurement of accessibility and provided a foundational framework for both computational process models and economic models in activity-based travel analysis (Rai et al., 2007).

2.1.1 Computational process models

Hägerstrand’s (1970) space-time prisms play a critical role in understanding travel by highlighting the interdependencies between activities. The choice of location and travel for one activity can create constraints for subsequent activities. Additionally, the interdependencies among household members and their joint decision-making processes are crucial. Understanding these interdependencies profoundly affects our comprehension of household travel choices, as emphasised by Jones (1979) and Jones et al. (1983). Modelling these rules or household heuristics has been influential in transport research, leading to the computational process model family of activity models. These models combine pattern recognition and multi-objective optimisation with discrete choice models to simulate travel patterns based on demands for activities (Recker et al., 1986a and 1986b). The application of heuristics was motivated by the recognition that decision makers could not possibly form preferences over many thousands of alternatives that represent the ‘when’, ‘where’, ‘with whom’ and ‘how’ to undertake a schedule of activities and that it is most likely that decisions would be made using information processing to limit the complexity of the decision. Gärling et al. (1994) argue that utility maximisation may be useful when focusing on a single trip choice and its relation to various factors like mode, departure time, and route choices. However, these models fall short when it comes to modelling the interdependency trips forming an activity schedule.

Advances in computational process models (CPMs) have incorporated trade-offs between activities (e.g., time constraints and durations), overall schedules (e.g., total time allocated to activities and travel), and the scheduling process itself (Auld & Mohammadian, 2009). Due to the complexity of these schedules and the large combinatorial problem they present, heuristics are often applied to simplify the modelling process. Rather than building individual schedules from scratch, groups of individuals with similar time-use activity patterns can be identified and modelled collectively. Hesam Hafezi et al. (2022) employed a Fuzzy C-Means (FCM) machine learning approach to cluster individuals based on the similarity of their daily activity patterns. While the primary focus of these models is on scheduling and processing activities, they also integrate utility-based components to capture choice behaviour.

2.1.2 Economic models

Economic models are underpinned by utility theory and use regression methods to quantify the attractiveness of different travel choices. These functions consider various travel attributes (e.g., time, cost) and individual characteristics (e.g., income, car ownership) to assess the relative desirability of alternatives like different modes of transport, destinations, or departure times (Bowman and Ben-Akiva, 2001). Unlike CPMs that simulate the step-by-step process of decision-making, economic models primarily focus on predicting the choices individuals make and the probabilities associated with those choices. This emphasis reflects the understanding that travel behaviour is inherently variable, and individuals may not always choose the objectively "best" option due to factors like imperfect information, time constraints, or personal preferences. By estimating choice probabilities, economic models can account for this variability and provide a more realistic representation of travel patterns.

Traditionally, economic models focus on individual trips as the unit of analysis. While representing a step forward from traditional aggregate approaches, these models often failed to adequately represent the sequential and interconnected nature of travel decisions (Bastarianto et al., 2023). These models do not capture the influence of previous trip choices or changing circumstances on subsequent travel behaviour, and hybrid models (activity simulators with choice models at their core) often use the tour as the basis of travel choice (Hasnine & Habib, 2020).

Table 1 summarises the key differences between economic and CPM models. Economic models are grounded in utility theory, but often rely on simplifying assumptions to ensure the choice models are both feasible and estimable. In contrast, CPM models offer greater flexibility but can be computationally intensive and more challenging to validate.

Table 1: Differences between Economic and Computational Process Models

Feature	**Economic Models**	**Computational Process Models**
Theoretical Basis	Random Utility Theory	Context-dependent Choice Preferences, Cognitive Processes
Focus	Predicting choices	Simulating decision-making process
Methods	Statistical analysis (e.g., discrete choice models)	Heuristic rules, simulations
Strengths	Can predict aggregate behaviour, statistically robust	Can model complex processes, flexible
Weaknesses	May oversimplify decision-making	Can be computationally demanding, harder to validate

2.2 Tour-based models

Tour-based transport models have gained prominence in activity-based travel demand modelling in the past two decades. These models focus on interconnected trips, considering factors such as tour type and mode choice (Bastarianto et al., 2019). Tour-based approaches are considered more relevant than trip-based models for activity-based frameworks (Hasnine & Habib 2020) and various methodologies have been developed, ranging from simplified main tour mode to complex dynamic discrete choice models (Hasnine & Habib, 2018). Instead of treating trips as isolated events, tour-based models account for the interdependence between trips, recognising that mode choice for any given trip may be influenced not only by its attributes, but also by those of preceding or subsequent trips (Recker et al., 1986). Hence, the importance of a tour-based modelling perspective is for:

- **Capturing trip interdependencies:** Recognising tours acknowledges the interrelationships between trips taken within a day. The choice of mode for one trip can influence subsequent trip decisions, and the overall pattern of travel is shaped by the desired activity sequence (Hasnine & Habib, 2018; Hasnine & Habib, 2020).

- **Understanding activity participation:** Tour-based models link travel choices to the broader context of daily activity patterns. This connection allows researchers to explore how factors like work schedules, childcare responsibilities, and access to amenities influence both travel decisions and overall well-being (Bastarianto et al., 2023).

- **More realistic policy evaluation:** By considering tours, transport planners can better assess the impact of policy interventions on travel behaviour (Hasnine & Habib, 2018; Kagho et al., 2020). For example, a new public transport service may not only affect the mode choice for that trip, but also influence the entire day's travel patterns.

Tour-based modelling requires analysts to make decisions about how to classify and construct tours from travel diaries. This is partly because tours often comprise multiple trips to various activities, raising questions about which activity defines the tour and what constitutes the choice set when each trip component may involve different modes, timings, and destinations. The stages required to be undertaken to achieve this are:

- **Tour definition and classification**: This involves identifying the types of tours individuals undertake, such as home-based work tours, home-based shopping tours, or tours for passenger servicing (Ben-Akiva & Bowman, 1998). Tours are typically defined by the primary activity, but the number of intermediate stops can also indicate more complex tour patterns (Hesam Hafezi et al., 2022).

- **Tour creation**: Individual trips reported in travel surveys are integrated into tours based on their origin, destination, purpose, and temporal sequence. For more complex tours involving multiple stops and modes, a hierarchy of purposes and modes can be used to identify the primary purpose and mode of the tour (Frank et al., 2008).

- **Choice set generation**: Identifying the set of feasible travel options, or alternatives, available to an individual for a given tour is critical for modelling travel choices. This process involves determining potential modes, routes, departure times, and combinations of modes for different legs of the tour. Capturing the full range of realistic alternatives can be computationally intensive, often requiring simplifications in model implementations (Miller et al., 2005).

Hasnine and Habib (2020) offered an insightful summary of tour-based mode choice models, evaluating the strengths and limitations of each approach. They provided a valuable categorisation of these models and detailed the mode choice frameworks implemented in some of the most prominent activity-based and agent-based models in transport research. Their work highlights how different approaches have been tailored to address specific challenges, such as capturing behavioural realism, computational feasibility, and multimodal travel behaviour. In the following section, we summarise the categories of tour-based mode choice models as outlined in Hasnine and Habib, with references to the original works. However, for a comprehensive understanding and detailed comparisons, readers are encouraged to consult their review paper directly. These categories are:

- **Simplified main tour mode model** is adopted in several operational activity-based models, addressing the complexity of transport modelling by identifying a primary mode for an entire tour and using it as the basis for the choice model. Fundamentally, this approach aligns with a trip-based mode choice framework and disregards the interactions and dependencies between multiple trips that constitute a tour (Hasnine & Habib, 2018). While the simplification facilitates computational efficiency and model tractability, it potentially overlooks critical behavioural dynamics inherent in multi-modal travel patterns. Early economic models such as AMOS (Pendyala et al., 1997) and Bowman and Ben-Akiva's (2001) activity-based model applied this main-mode specification. The approach has also been integrated as a sub-model within hybrid activity-based and econometric frameworks, including ALBATROSS (Arentze & Timmermans, 2004) and within ADAPTS (Auld & Mohammadian, 2012).

- **Two-tier nested logit model** uses a hierarchical structure to model travel behaviour, with the upper tier representing location choice and the lower tier capturing mode choice conditional on the selected destination. This framework handles the interdependencies between destination and mode decisions. An early implementation of this approach is the Prism-Constrained Activity Travel Simulator (PCATS), developed by Kitamura and Fujii (1998). PCATS reflects realistic travel patterns by integrating hierarchical decision-making processes constrained by spatial and temporal feasibility. Unlike PCATS, Ho and Mulley (2013) applied a two-tier nested logit model with a distinct focus on joint household travel patterns in Sydney. The model extended beyond destination choice to capture other decision factors. The upper tier addressed joint travel activity patterns within households, reflecting the coordination of activities among members, while the lower tier modelled individual mode choices between public transport, car, and walking.

- **Simplified main tour mode and conditional trip-level mode choice model** is a two-stage process commonly used in activity-based models to address the complexities of tour-based mode choice while maintaining computational feasibility (Hasnine & Nurul Habib, 2021). In the first stage, Simplified Main Tour Mode, a single primary mode is assigned to the entire tour, such as categorising a multimodal trip (e.g., bus and walking) as a "bus" tour if the bus is the dominant mode. In the second stage, trip-based mode choice models are conditioned on the modelled upper-level tour choice. DaySim (Bradley et al., 2010) approaches the two stages as an economic model with log-sums passing from the second stage (trips) to the tour-based model. CT-RAMP (Davidson et al., 2010) and SimMobility (Azevedo et al., 2017) use variations of the two-stage approach within the mode choice sub-models.

- **Simulation-based tour-based mode choice models** address the issue of choice set explosion when a tour has many trips, which can be difficult to handle with closed-form models. Simulation-based approaches take a ‘draw’ of possible choice sets. This method is adopted in the Travel/Activity Scheduler for Household Agents (TASHA) model and can be used to estimate tour-based mode choice (Miller et al., 2005). This approach simulates the mode choice process using Monte Carlo simulation and uses deterministic rules for activity scheduling.

- **Combinatorial tour-based mode choice models** use advanced discrete choice models, such as recursive logit models, to capture all feasible trip mode combinations within a tour. This approach is more complex but allows for a more detailed analysis of mode choice behaviour. The idea stems route choice modelling application proposed by Mai et al. (2015), whereby the enumeration of all possible routes leads to an impossibly large choice set and the problem is broken down to focus on the ‘next move’. Vovsha et al. (2017) make use of the recursive logit model, for a stand-alone tour-mode choice model.

- **Dynamic discrete choice models** account for the sequential nature of decision-making, recognising that choices made at one stage influence subsequent decisions. For example, opting to drive for one trip increases the likelihood of driving for the next trip, as the car remains available. These models leverage state dependence, where earlier decisions strongly affect current choice tasks. Hasnine and Habib (2020) argue that dynamic models will play an increasingly critical role in future transport networks, particularly as higher levels of ridesharing and multimodal travel patterns emerge in future networks with autonomous vehicle technologies.

In summary, tour-based mode choice models trade-off complexity, computational feasibility and behavioural realism. The Simplified Main Tour Mode is most appropriate in contexts with lower levels of multimodal trips. This method identifies a primary mode for the entire tour and simplifies the model structure. While computationally efficient and easier to implement, it ignores multimodal tours that may be present in the data. A potential enhancement involves integrating Conditional Trip-Level Mode Choice, where trip-level models are conditioned on a higher-level tour-based choice, offering a balance between tractability and behavioural detail.

Where high levels of multimodal trips are observed or where future mobility expectations are that there will be an increased level of ridesharing or the advent of autonomous vehicles, more advanced methods may be needed. Simulation-Based Models, such as TASHA, can address choice set complexity by simulating mode choice using Monte Carlo draws. Combinatorial Models, such as recursive logit frameworks, allow detailed analyses of mode choice combinations but demand significant computational power and are complex to implement. Finally, Dynamic Discrete Choice Models capture the sequential nature of decisions, recognising the influence of prior choices on current and future behaviours. These models are particularly well-suited for contexts expecting higher adoption of multimodal travel or shared mobility solutions, though their complexity and data requirements remain significant barriers. Ultimately, the appropriate model choice hinges on the prevalence of multimodal travel and future transport priorities.

3 Travel profiling (data analysis)

The PATHS survey was undertaken between 2018 and 2022 by Ipsos on behalf of Main Roads Western Australia (MRWA, 2021). This household travel survey gathered data on randomly selected households and their eligible occupants utilising a GPS device to record travel movements for five days and utilising a 15-minute survey beforehand.

After a thorough data cleaning and validation process, a variety of artefacts were produced by this process, which can be used for transport planning and modelling purposes. Primarily, these were datasets that were derived and visualised from the initial dataset to better understand the properties of the dataset. Alongside this, a dataset consisting of a set of tours that can be utilised for the tour-based mode choice model was also produced.

The process and artefacts produced by Bentley Systems were first examined for use within the project. Unfortunately, this process generated tours that only contained geographic information (i.e. where the tour started, where the tour ended and any stops in between). While useful for network planning, additional information regarding the existing choice of mode and tour purpose was required as these insights will help develop the choice model. As such, the process undertaken by Bentley Systems was studied and adapted to generate additional tour-based datasets, which have been used in the second part of the project.

3.1 Household level

PATHS surveyed 6,444 households, of which 4,895 were family units with or without children, 1,404 were sole occupants, and 145 were shared or multiple-family households. Mostly, households resided in separate/detached houses, with 20.7% of surveyed households residing in attached dwellings such as townhouses or apartments (Table 2).

Table 2: Dwelling types for surveyed households

Dwelling	**Survey Frequency**	**Rel. Frequency**
Separate House	5,099	79.1%
Townhouse / Duplex / Villa	815	12.7%
Flat / Unit / Apartment	518	8.0%
Other	12	0.2%

3.1.1 Household income

Figure 1 displays total annual household income (HHID) aggregated from personal income data per week (from PATHS person data). Each individual's weekly income was summed within their respective household to reflect a combined household income, and this was then multiplied by 52 weeks to represent the annual income. Almost a quarter of the households reported zero income (1,501, representing 23.29%) and were removed from the reported statistics.

The percentage distribution of household incomes varies across various annual income brackets. The significant income brackets include those earning between $120k-$160k (1,017 households with 20.57%) and those earning between $40k-$80k (1,003 households with 20.29%). Additionally, the distribution shows that 18.39% (909 households) earn only between $1k-$40k annually. Households with an income between $80k-$120k represent 17.46% (863 households). Fewer households have their income over $160k annually, with those in the bracket $160k-$200k, encompassing 9.45% (467 households), and those over $200k representing 13.84% (684 households).

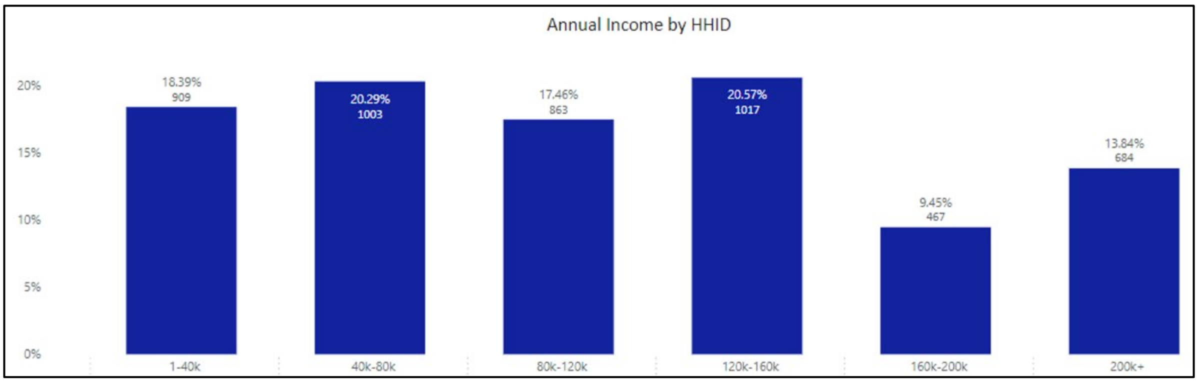


Figure 1: Household Annual Income, by range and percentage of sample

3.1.2 Vehicle ownership

The PATHS data displays slightly higher vehicle ownership in Perth Metro area compared to the Census data (City of Perth, 2021) indicating that 90.9% of the households owned at least one car[1]. The PATHS sample also included more households with two or more vehicles (60.8%), compared to 57.2% recorded in the most recent Census.

Table 3: Vehicle ownership

Number of vehicles	Survey Frequency	Relative Frequency
None	243	3.8%
One Vehicles	2,284	35.4%
Two Vehicles	2,852	44.3%
Three Vehicles	774	12.0%
Four or More Vehicles	291	4.5%

Figure 2 provides the distribution of car ownership by structure of the household. Unsurprisingly, single-person households account for the largest part of households without a car (176 out of 243). Yet, there are also 67 families without car access. This car availability is expected to be influential in the tour mode choice.

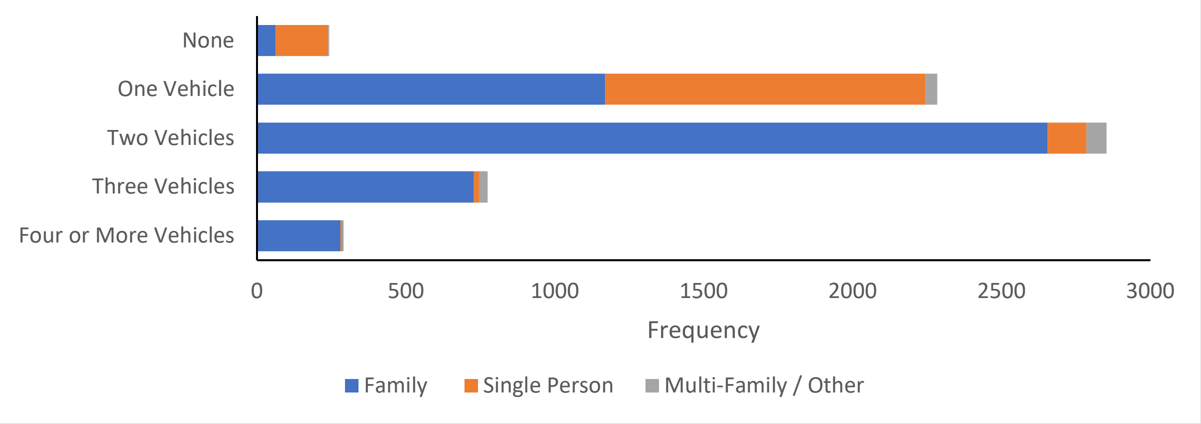


Figure 2: Vehicle Ownership by Household Structure

---

[1] 4.2% did not state their car ownership.

3.2 Person level

A summary of age and gender for the sample is given in Table 4. In the age group under 15, males slightly outnumber females (51.31% vs. 47.92%, Other: 0.77%). This continues in the age group 15-29 (Males: 51.22%, Females: 48.74%, Other: 0.04%).

From 30-44, females begin to slightly outnumber males (Females: 50.95%, Males: 48.89%, Other: 0.16%) and this trend becomes clearer in the age group 45-59 (Females: 52.06%, Males: 47.90%, Other: 0.03%) and continues in 60-74 (Females: 51.46%, Males: 48.47%, Other: 0.07%). In the 75-89 age group, the proportion of females is higher (Females: 53.12%, Males: 46.88%), however, the reverse occurs in the 90+ group (Males: 53.57%, Females: 46.43%).

Overall, females constitute 50.36%, males 49.42%, and the "Other" gender remains a very small percentage across all age groups (0.22%).

When examining the age distribution, similar proportions of individuals under 15 and over 60 were recorded, with over 57% of the sample representing working-age individuals. The age distribution is relevant when accounting for the purpose of the trips within the trip chains.

Table 4: Sex and age distribution

Age group	N	% Females	% Males	% Other
< 15 years	3,516	47.92	51.31	0.77
15-29 years	2,780	48.74	51.22	0.04
30-44 years	3,649	50.95	48.89	0.16
45-59 years	3,198	52.06	47.90	0.03
60-74 years	2,676	51.46	48.47	0.07
75-89 years	977	53.12	46.88	0
> 89 years	56	46.43	53.57	0

3.2.1 Driver licenses

Excluding the age categories too young to hold a licence, 93.8% (n = 11,505) hold a Class C drivers' license, 3.21% (n = 394) hold a learner permit, and 2.70% (n = 331) are driving under a provisional license. Motorcycle licences are much less common with 8% (n = 964) holding a full licence and 2% (n = 294) holding a licence to operate a motorcycle of up to 250cc.

3.2.2 Taxi and rideshare usage

Excluding respondents under 15 years old, 54.5% (n = 6,080) reported never having ridden in a taxi or used rideshare services. Additionally, 6.8% (n = 773) stated they had taken a taxi, but had never used rideshare services. Among respondents who reported some level of rideshare use, 26.4% (n = 3,208) reported having never taken a taxi. For the remaining respondents, Figure 3 illustrates their levels of rideshare and taxi use. Rideshare services not only have a higher level of reporting, but the frequency of use is also relatively higher, with 24% (n = 1,070) of respondents reporting at least fortnightly use, compared to only 9% (n = 205) for taxi use.

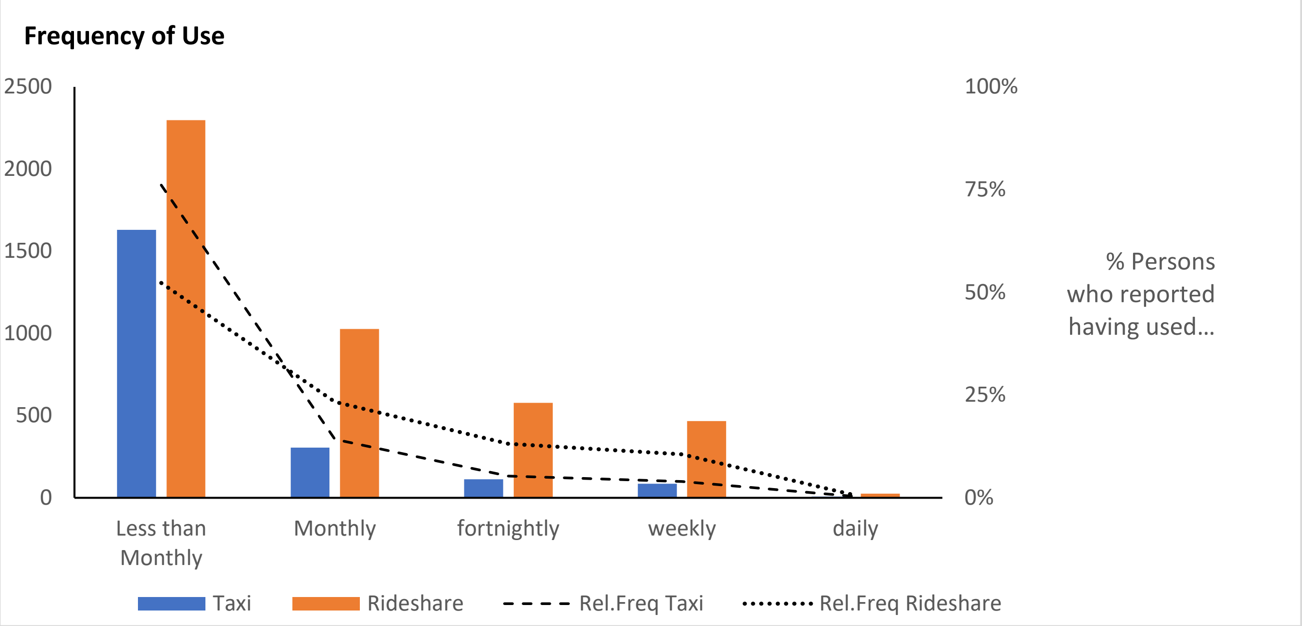


Figure 3: Taxi and Rideshare Frequency of Use

This reduced frequency of use is consistent with the frequency of modes present in the SP scenarios (section 6.1). Examining the trip-level data from the PATHS RP dataset, very few trips were made by taxi or rideshare services. Of the 60,781 trips within the dataset, only 0.11% (n = 67) were by taxi with 0.26% (n = 155) taken by rideshare services. The vast majority of trips (81.2%) were made by private vehicle (n = 49,349).

3.2.3 Public transport ridership

Continuing to examine the trips taken in the PATHS RP dataset, public transport (PT) was used for substantially more trips (as the main mode) than rideshare and taxi, but still a much smaller proportion than private vehicles. Buses were used as the main mode in 1.76% of trips (n = 1,072) and trains were used for 1.67% of trips (n = 1,013). Smaller proportions of trips were taken with school buses (n = 212), private buses (n = 55) and ferries (n = 14).

3.2.4 Active travel

Active travel modes were well represented within the PATHS RP dataset, with walking being the most common main mode taken for trips other than private vehicle, with 11.82% of trips being walked (n = 7,184). Cycling was less popular but still used for 1.42% of trips (n = 864). A smaller proportion was seen for running (n = 73).

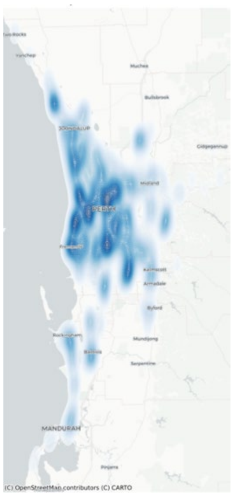
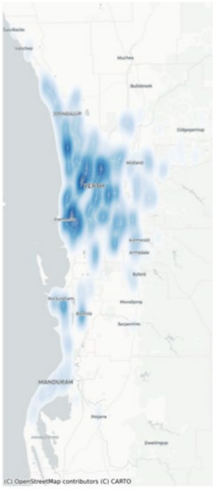
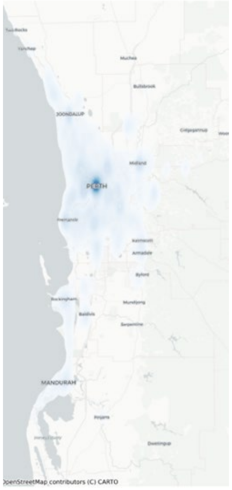
4 Exposition of trips

To better understand where people began and ended their trips, based upon the trip purpose (such as education, work or home), heatmaps were produced to visualise the density and geographic spread of these journeys throughout Perth, using the location data embedded within the trip dataset.

The kernel density estimation (KDE) process was used to generate the heatmaps, which are visualised in Figure 4 below. This process allows the summarisation of data through smoothing over space (in this instance), such that areas within the heatmap where there are many trips that have the location as a destination are shown as a deeper colour to those where less occurred and where none occurred within the sample.

It should be noted that this process only visualises trips within the sample of Perth; hence while it should show a representative view of the general "common" trips taken by the sampled residents of Perth, it should not be expected to show every trip taken by each person. The intensity of colour is also relative to each individual map; hence an intensity of colour in one location of one map should not be taken to mean there is an equivalent number of trips at a location in another map with the same intensity of colour. Rather, the most intense colour in each map shows the location(s) with the most trips for the chosen location type, with the other locations in that map being coloured in terms of intensity relative to this 'peak' location with the most intense colour.

The set of heatmaps in Figure 4 shows trips by destination location and trips by destination location for the following trip purposes respectively and in a clockwise fashion: education, home, not home, retail, services and work.

		
Education	Home	Not-home

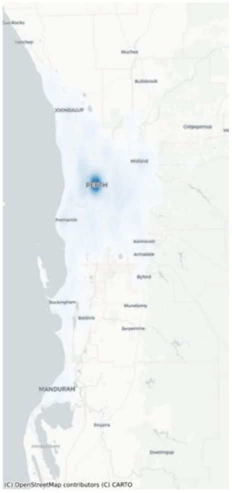
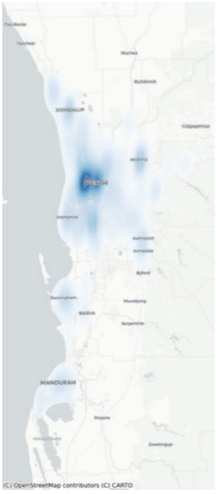
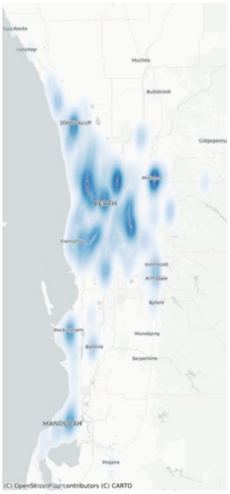
		
Work	Services	Retail

Figure 4: Heatmaps of trip destinations for education, home, not-home, retail, services purposes and work purposes (clockwise from top-left)

For trips that have ‘home’ as a destination, the density and location of destinations roughly correspond to the population density and locations of major urban settlements throughout Perth. On a higher level, the heatmap’s extent (excluding very low-density areas) corresponds to the extent of the Perth metropolitan area.

For other destination types, the density and locations also mirror the major relevant areas of interest for Perth – education is most concentrated around the universities, work purposes around the CBD, services around the hospitals (notably the QEII Medical Centre just west of the CBD) and retail around major shopping centres, including the Perth CBD. Examining all destinations other than ‘home’, the CBD dominates and hence so do work-based trips.

Regarding origins, as most trips formed part of a tour with two legs (that is, pendular trips, from home to a destination such as work or shopping and back home), the heatmaps are not substantially different as the trips are close to evenly split between ‘there’ and ‘back’. The split of these tours by the number of legs is further explored in Section 5.2.

5 Tours

5.1 Tour construction

To generate tours, we combined relevant data files containing information about stops, trips, persons, and households. This process involved merging the files, grouping stops by trip ID, and reconstructing the tours by retaining the origin of the first stop and the destination of the last stop for each trip. The resulting file was organised by trip ID and sorted within each group by stop ID in ascending order, ensuring the correct sequence of stops for each trip. For trips with multiple stops, the origin was identified as the starting point of the first stop, and the destination was the endpoint of the last stop. Thus, for each trip, we extracted the origin-related data of the first stop and the destination-related data of the last stop.

The tour file is merged with the PATHS person file to extract the individual's demographic information. For every trip within a tour, the data is merged with the PATHS trip file to extract the higher-level destination purpose and lower-level destination place using this information we were able to allocate groups of trips to the following tour classifications (Table 5):

Table 5: Tour coding, explanation and derivation

Code	Explanation
HBW	**Home-based work:** the trip consists of a person (worker) travelling to or from their workplace from/to home.
HBU	**Home-based university:** the trip consists of a person (student) travelling to or from their university from/too home.
SCH	**School:** the trip consists of a person (student) travelling to or from their school from/to home.
HBS	**Home-based shopping:** the trip consists of a person travelling to or from a retail shop, (where they intend to undertake shopping) from/to home.
SPS	**Service passenger:** the trip was undertaken in support of transporting another person from home.
HBO	**Home-based other:** the trip was undertaken to or from home for some other purpose.

After assigning trip codes to each trip within a tour, the tour code is identified based on a home-based priority order: SCH > HBU > HBW > HBS > SPS > HBO.

To identify IX/XI tours, tour geo-codes were overlayed on greater Perth. Tours having all destination are fully within Perth region were allocated one of the tour codes as given in Table 5. The remaining tours were classified as IX (internal-external) if the tour originated within the boundaries of greater Perth, or XI (external-internal) should the tour include a stop outside of Perth.

5.2 Tour statistics

The observed tours and their frequencies by classification (e.g. HBW or SPS) are illustrated in Figure 5. Most tours originate from home (21,815 of 24,344 or 89.7%), of these, the following pattern emerges: a third have other purposes (HBO), a quarter for work (HBW), followed by shopping (HBS, 19%), servicing a passenger SPS (12%) and school (SCH, 11%). Tours involving travel for tertiary studies University (HBU) are less than 1% of all tours.

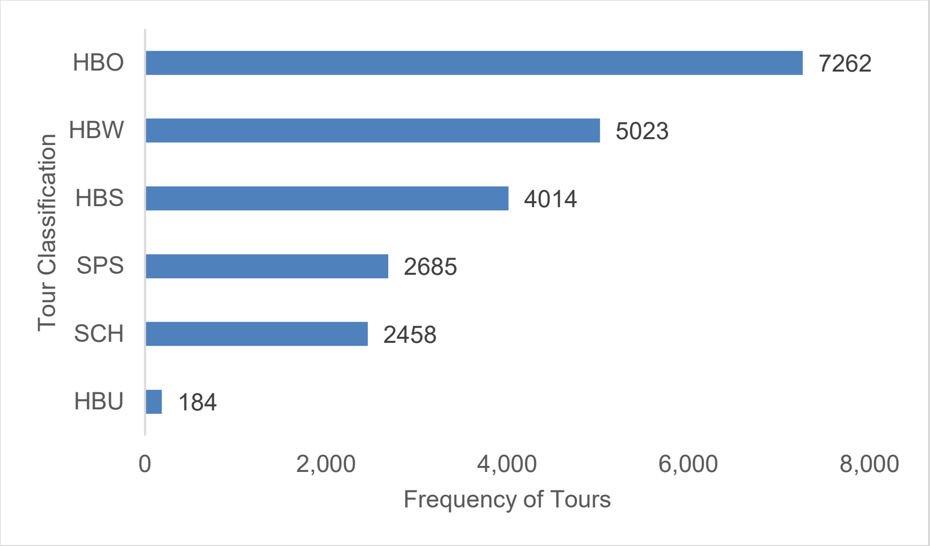


Figure 5: Distribution of Tours by Type

Considering the complexity of tours, it is noted that 67% are ‘simple’ or involve a single activity before returning home. The remainder (33%) represents ‘complex’ tours and involve trip chaining over multiple activities. Figure 6 clearly shows a declining frequency of observed tours as the complexity increases, with about 12% including three or more activities.

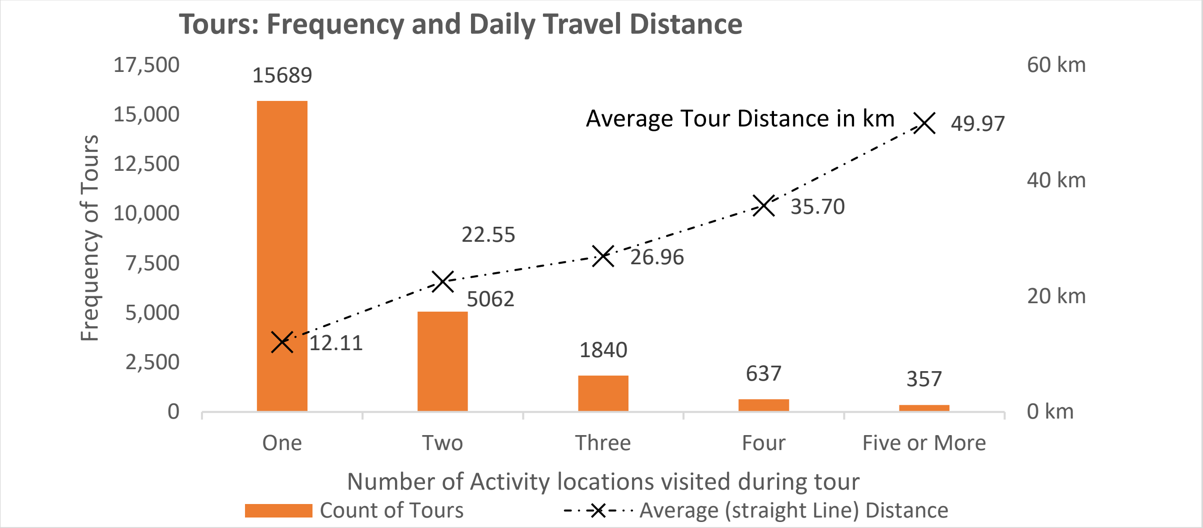


Figure 6: Frequency of Tours and Average Tour Distance

Tour distances increase with additional activities. The average distance for simple tours is 12.1 km and this extends to approximately 50km for tours with five or more activities (Figure 6). Overall, the survey data suggests that, while multi-location tours are less common, they are associated with longer travel distances.

5.2.1 Main tour mode

For each tour, a main mode was allocated. In most cases, this was straightforward, as 94. of tours were unimodal. For the small proportion of multimodal tours, the main mode was determined using a hierarchical classification based on modal dominance: Train > Bus > Taxi or Rideshare > Cycle > Private Vehicle > Walk. This approach prioritises modes typically associated with greater capacity, longer distances, and higher strategic importance within the transport system.

Figure 7 shows the distribution of main modes for tours, separating motorised and active travel.

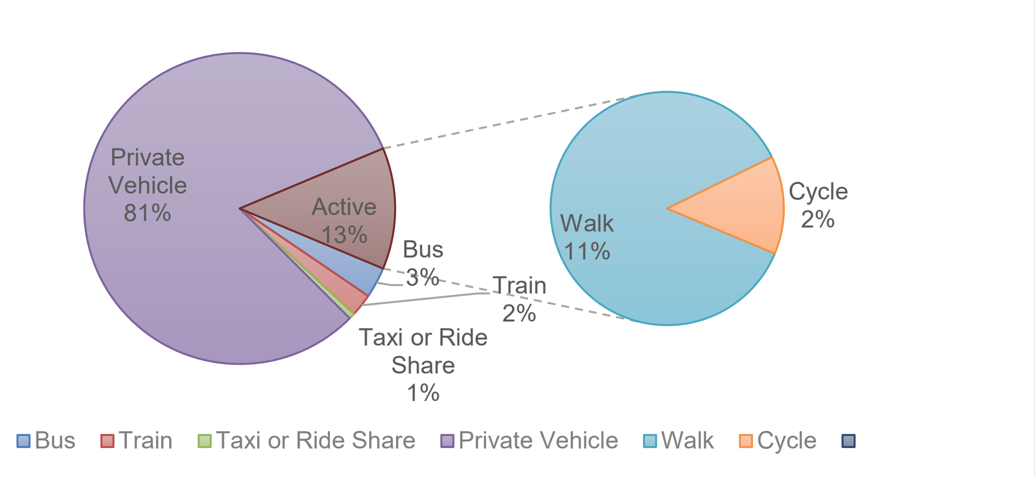


Figure 7: Distribution of Tours by Travel Modes

Nearly 90% of tours rely on motorised modes, with private vehicles dominating at 81%, while other options, such as bus (3%), train (2%) and taxi or ride share (1%), accounted for smaller shares. The active travel (AT) was dominated by walking, comprising 11% of the tours, and cycling accounted for 2% of all tours.

Figure 8 shows the average tour distances by the most frequently used modes. AT modes are associated with the shortest tour distances, with walking tours averaging a little over 1 km (not including very short <100m walk legs) and cycling tours just under 10 km. In contrast, motorised modes cover substantially greater distances. Train-based tours have the longest average distance at approximately 51 km, reflecting the structure and role of Perth's rail network. The Mandurah line (extending over 70 km) and then Butler line (around 40 kilometres) provide efficient links to the Perth CBD and are used for both commuting and social trips. Importantly, the presence of Park-and-Ride (PnR) facilities at key stations on these lines enhances accessibility for residents living further from the rail corridors, making train travel a viable option even in outer suburbs. Tours with bus as the main mode (19 km) are comparable in length to those taken by taxi or rideshare (23 km) and private vehicle (also 23 km), suggesting that buses support a broad mix of medium-distance travel.

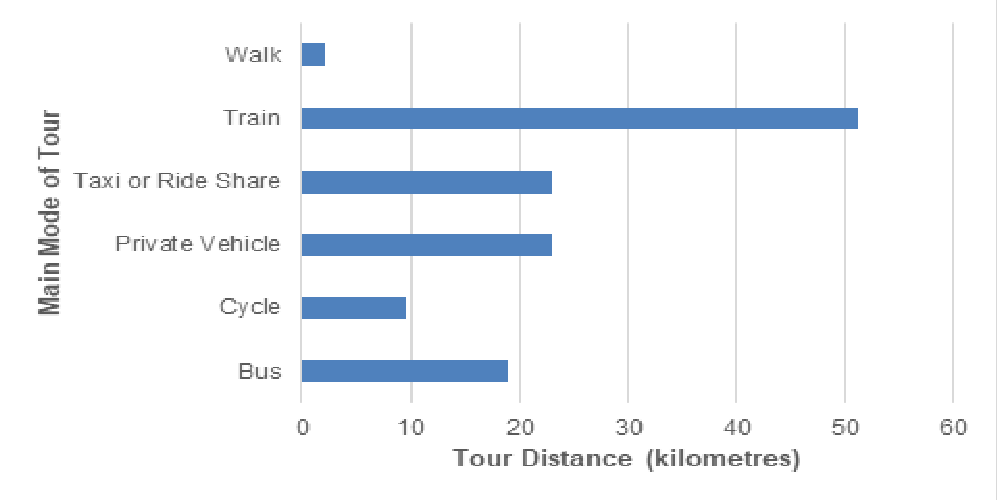


Figure 8: Tour Distance by Main Mode of the Tour

5.2.2 Tour complexity

As shown in Figure 9, the observed complexity of Home-Based Shopping (HBS) tours reflects the tendency for shopping behaviour to involve multiple destinations. While some activities may be spatially clustered (e.g., within a single shopping centre), the data were cleaned to exclude short walk segments that did not reflect meaningful tour complexity. For example, sequences such as *Home to Shopping by Car (7 km), Shopping to Coffee by Walk (120 m)* and *Coffee to Home by Car (7 km)* were recoded as simple HBS tours and the intermediate activity removed from the modelling file. Although such patterns may offer insights into how people organise activities in space, they are not relevant for a strategic mode choice model. In this context, the tour is better classified as a single-purpose car-based shopping trip. Yet, despite our decision to 'simplify the tours', HBS maintained the presence of multiple activity locations, particularly those more than 1 km apart. This suggests that individuals are chaining together spatially dispersed tasks. This could include, for example, visiting a grocery store, a specialty shop, and a service provider (e.g., post office or pharmacy), each located at separate nodes in the urban network. Such behaviour is consistent with efforts to minimise total travel costs or make efficient use of a vehicle trip.

Home-Based Work (HBW) and Home-Based University (HBU) tours also exhibit notable complexity, though to a lesser extent. School (SCH) and Passenger Servicing (SPS) tours tend to be simpler, reflecting their routine and singular nature, often with a direct return to home. Similarly, Home-Based Other (HBO) shows lower complexity, possibly due to its association with short, singular activities.

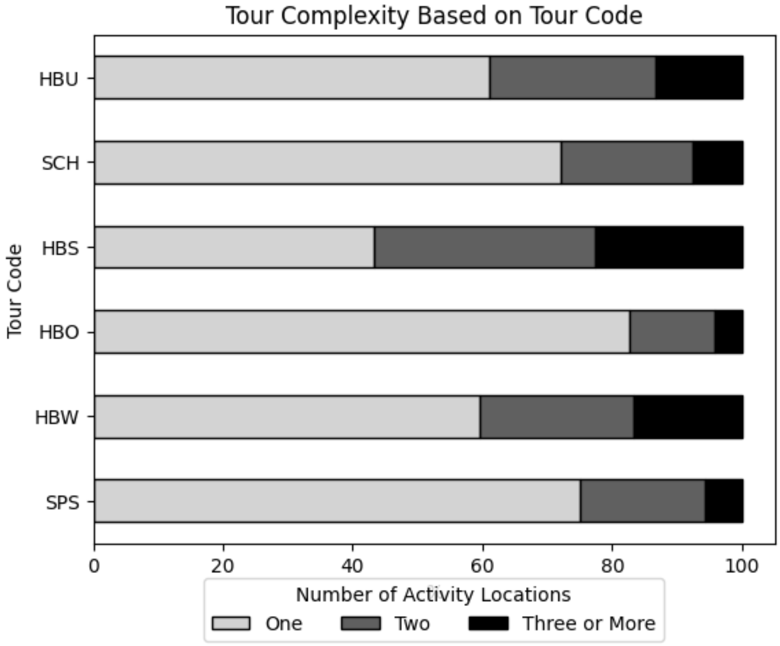


Figure 9: Frequency of Tours and Tour Codes. Number of tours: HBO (7,484 tours), HBW (6,033), HBS (4,182), SPS (2,849), SCH (2,597) and HBU (188)

Figure 10 illustrates how tour complexity varies by time-of-day. Tours that commence during the AM Peak and Off-Peak periods show a higher incidence of multi-activity travel, suggesting that individuals are more likely to chain trips, such as work, shopping, or personal business. In contrast, tours beginning in the PM Peak and Evening are more likely to be simple, single-purpose tours, likely to serve passengers or recreational/social evening engagements.

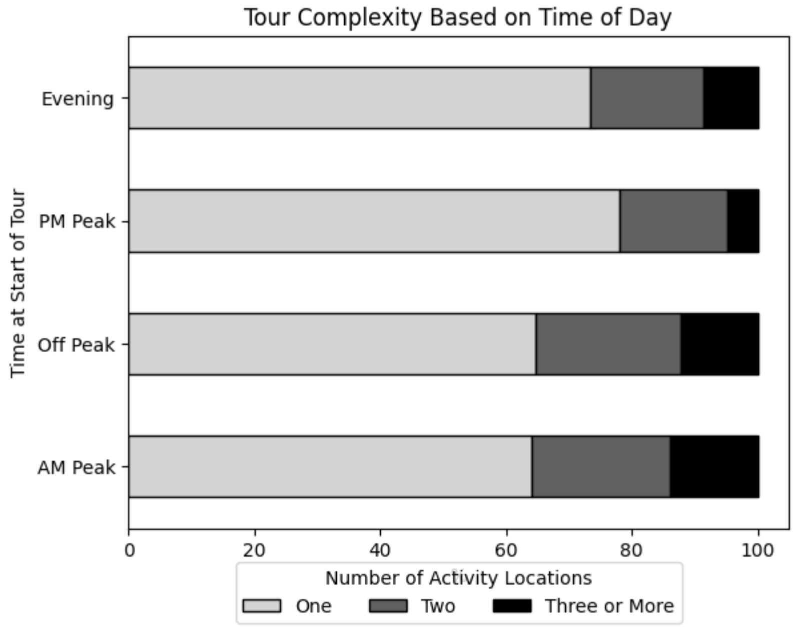


Figure 10: Frequency of Tours and Time of Day. AM peak (8,247) tours; Off (between) Peak (8,494); PM Peak (2,629) and Evening (4,137)

Figure 11 displays the complexity and frequency of tours by the assigned main mode (through the procedures described above). In general, tours with a single activity location dominate, however this is most pronounced in the case of AT (walk-based and cycle-based) tours and least pronounced on PT (rail or bus) and car modes.

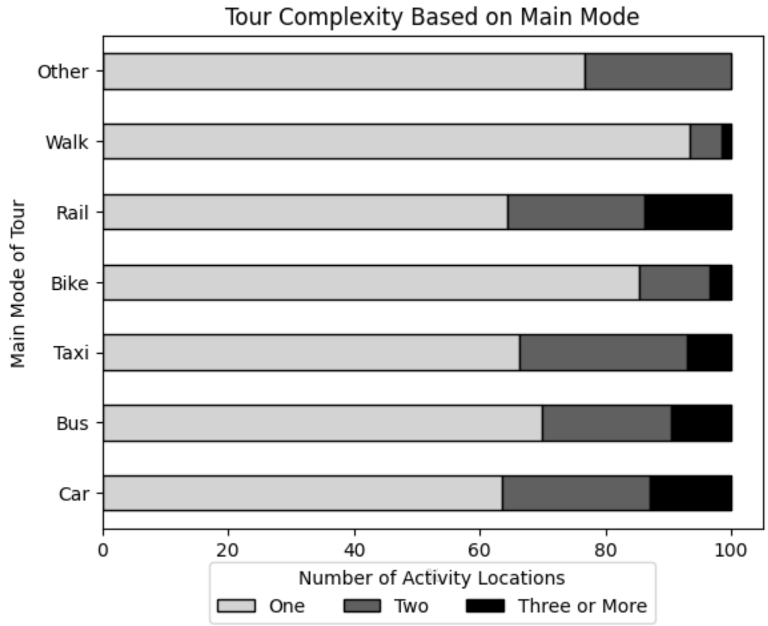


Figure 11: Frequency of Tours and Main Mode - Other (43), Walk (2,704), Rail (511), Bike (404), Taxi (101), Bus (705) and Car (19,045)

5.2.3 Multi-modal tours

Most tours (94%) were unimodal (Table 6) when considering the main mode (by distance) used for each linked trip included in the tours and ignoring the access and egress modes. Of the small proportion of bimodal tours, many of these consisted of walking and private vehicles (Car-Walk), with smaller but notable proportions consisting of combinations of (two of) walking, PT and private vehicle. Only a small number of tours consisted of three or four modes, and no tours in the sample contained more than four modes within a single tour. The relatively small share of multimodal tours and the fewer that involve modes other than walking and private vehicle impose difficulties in estimating more advanced models such as Conditional Trip-Level Mode Choice and Dynamic Choice Models. Table 6 below provides the distribution of tours by the number of modes included in a tour.

Table 6: Modality of RP tours

Number of Modes	**Count**
One	22,304
Two	1,107
Three	76
Four	3

Table 7 highlights clear differences in the extent to which various transport modes support multi-activity tours. Private vehicle travel, particularly by car, is most associated with complex travel patterns, with 35.7% of car-based tours involving multiple activities. Commercial vehicles also show a high proportion (45.0%), likely reflecting the nature of work-related trip chaining. Active modes such as walking (6.6%) and cycling (10.8%) demonstrate some capacity to support multi-activity travel, particularly for local trips. In contrast, PT plays a limited role in facilitating complex tours. Only 2.6% of bus tours and 0.4% of train tours involve more than one activity, suggesting that PT is primarily used for single-purpose travel rather than for chaining multiple activities.

Table 7: Most frequent modes for single- and multiple-activity trips

Mode Sequence	**Single-activity**	**Multi-activity**	**% Multi-activity**
Car	11,825	6,567	35.7%
Commercial vehicle	66	54	45.0%
Taxi (Rideshare/Taxi)	50	8	13.8%
Bus	337	9	2.6%
Train	270	1	0.4%
Bicycle	331	40	10.8%
Walk	2,524	178	6.6%

Despite the main tour mode allocations shown in Table 7, a more detailed examination of mode sequences reveals that PT often forms part of a multimodal travel pattern, as shown in Table 8. While accessing multiple activities is most achieved through car use in combination with walking—for example, the sequence *Car, Walk, Car* accounts for 178 multi-activity tours, there are also complex tours involving both public and private transport. Sequences such as *Car, Bus, Car* (45 multi-activity tours) and *Bus, Walk, Bus* (41) suggest that bus travel can support chaining of activities when integrated with other modes. Moreover, even simple tours such as *Home to Work to Home* may be multimodal. For instance, the sequence *Train, Car* records 50 multi-activity tours and *Walk, Car* 52, indicating that some travellers may share rides with household members in one direction and return using a different mode, such as walking or PT.

Table 8: Top 10 most frequent multi-mode sequences for multi-activity trips, and the appearance of these activities in single activity tours

Mode Sequence	Single-activity	Multi-activity
Car, Walk, Car	–	178
Car, Walk	94	41
Walk, Car	80	52
Car, Bus, Car	–	45
Bus, Car	38	34
Car, Bus	88	11
Bus, Walk, Bus	–	41
Train, Walk, Train	–	34
Car, Train, Car	–	9
Train, Car	20	50

The data shows that multi-activity tours account for just under one-third (32.4%) of all tours, while multi-mode combinations represent just 5.3%. As shown in Table 9, the most common multi-mode combination is Car and Walk, with 1.9% of all tours involving multiple activities across these two modes. Combinations involving Public Transport and Walk (0.63%) and Car and Public Transport (1.46%) are also present, though far less frequent. More complex combinations involving Car, Walk, and Public Transport are rare, comprising just 0.2% of tours. Most travel occurs as single-mode tours (94.7%), reinforcing the dominant role of single-mode travel within the dataset. Given the low incidence of multi-mode tours, caution is warranted in drawing strong conclusions about the behavioural patterns they represent, as the sample size for these combinations is relatively small.

Table 9: Summary of multi-mode and complex (multi-activity) tours

Multi-mode Combinations	Single Activity (i.e. two trips)	Multiple Activities	Percentage
Car and Walk	154	136	1.91%
Public Transport and Walk	25	72	0.63%
Car and Public transport	165	195	1.46%
Car, Walk and Public Transport	-	28	0.19%
Single Mode Tours	15,437	6,813	94.72%

To examine the variety of attributes associated with multi-modal sequences while ignoring directionality, *Bus, Car* tours were combined with *Car, Bus* tours. Given the overwhelming number of tours that were for tour code “SCH”, started at either the AM Peak or Off-Peak times and were for U17s, a large majority of simple tours appear to be students going to and from school. The percentage of non-school based tours using multi-modes was just over 1% with the largest mode share being Car, Walk at 0.6%.

5.2.4 Comparing tours

When comparing the single-mode tours (grouping tours using public transport into one category and active travel into a separate category) with Car-Walk and Bus-Car, clear distinctions in their characteristics emerge. The single-mode PT tours are significantly longer (both in distance and duration) than other tour types, while the AT tours are the shortest. The multi-mode tours (Car-

Walk and Bus-Car), have significantly more activities, they are undertaken by individuals from larger households, who also report the highest number of daily trips (all ANOVA tests significant at 0.01 level). These comparative findings support the selection of predictors used in the choice of tour transport modes, discussed further in Section 7.

Figure 12 provides a few selected examples illustrating these tour differences.

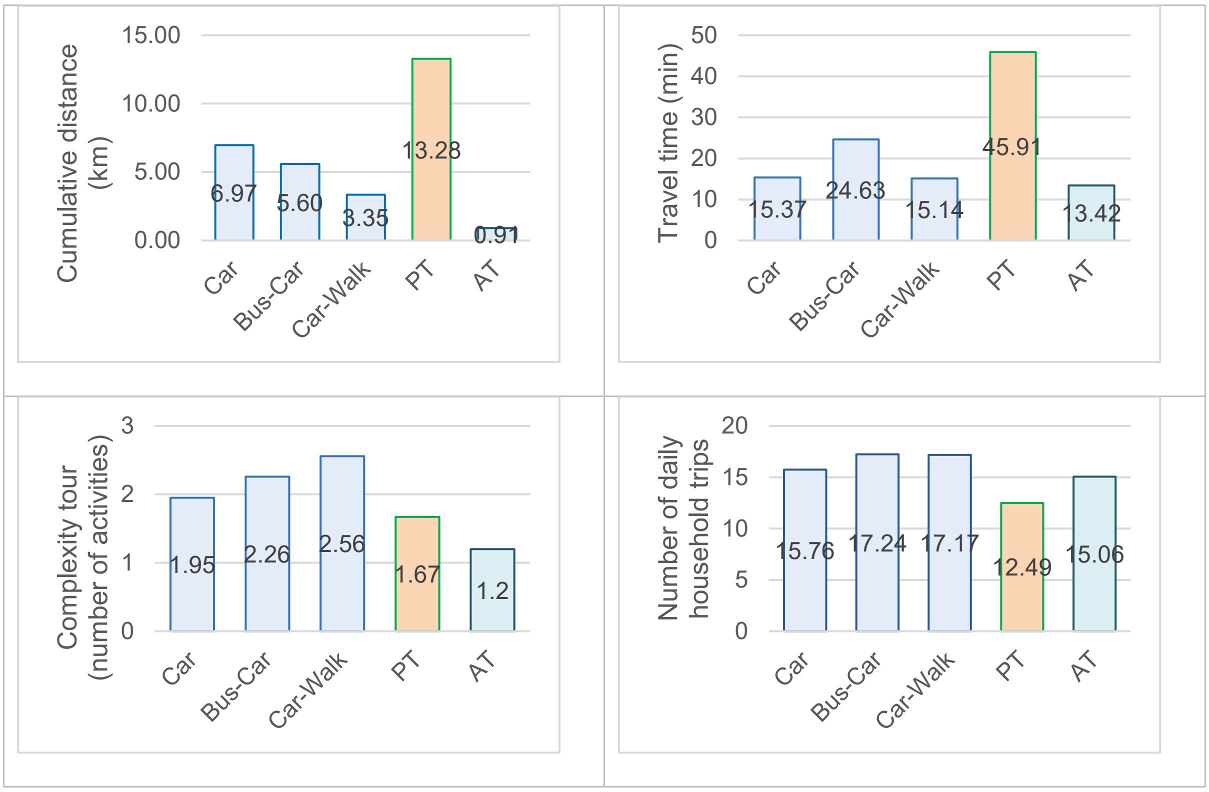


Figure 12: Examples comparing single- and multi-mode tours

6 Network skims to generate revealed preference data

Point-to-point travel skims were developed using a combination of OpenStreetMap (OSM) network distances, GTFS transit schedules, and vehicle travel time matrices from STEM. Skims were extracted for several mode categories: car, transit (all modes), bus-only transit, rail-only transit, park-and-ride (PnR), walk, and bicycle. Transperth GTFS data from the 17th of August 2021 (weekday) and 21st of August 2021 (Saturday) was used to reflect representative service conditions. All access and egress routing for non-car modes was computed using the OSM pedestrian network.

Public transport cost components—including in-vehicle time, access and egress time, and waiting time—were derived using OpenTripPlanner (OTP). However, due to limitations in OTP's native handling of park-and-ride itineraries, a separate methodology was developed. This involved a simplified probabilistic station choice model based on each household's location, OSM-based network distance to candidate rail stations, and station-level parking capacities. Based on the selected station for each origin, PnR skims were constructed by concatenating STEM-derived car travel times for the access leg with OTP-derived transit skims for the egress leg.

Car skims and the car access components of park-and-ride trips were taken directly from STEM's road assignment outputs. As only AM peak and inter-peak matrices were available, PM peak travel times were estimated by transposing the AM peak OD matrix, assuming symmetric conditions. Final skim values were validated against PATHS survey data by comparing stated and derived travel times, after excluding trips with short distances (<100 m) or implausibly low reported durations, or public transport trips with excessive access times. As shown in Figure 13, there is alignment between stated and derived values for most modes and metrics.

To enable comparison and validation between the extracted point-to-point skims and PATHS data, and transformation of the observed PATHS trip records was undertaken as follows. Trip records were processed using modes for each trip leg to identify the overall chosen travel mode and compute relevant travel metrics. Individual trip legs were first mapped from PATHS' internal mode codes to a simplified mode taxonomy (e.g., 'bus', 'rail', 'car', 'walk'). Consecutive duplicate leg modes were collapsed, and each trip's overall mode was inferred using a hierarchical rule set that accounted for sequential multimodal patterns. For example, a trip involving a walk and a rail leg was classified as a rail trip, while a trip starting from home involving a car leg before a public transport leg was classified as a PnR trip. Trip-level attributes such as duration, distance, auxiliary (non-transit/car) travel time, transit time, waiting time, and number of transfers were calculated by aggregating leg-level data.

Caution is warranted when comparing wait times derived from travel diaries and those estimated from network skims. Unlike real-world behaviour, the network skim data does not account for traveller responses to reading timetables or optimising their departure time. As a result, PT wait times from network skims are represented as half the scheduled headway, based on GTFS records. In contrast, the Wait attribute in the survey data captures the cumulative wait time across all transfers within a trip. These methodological differences mean the two measures are not directly comparable.

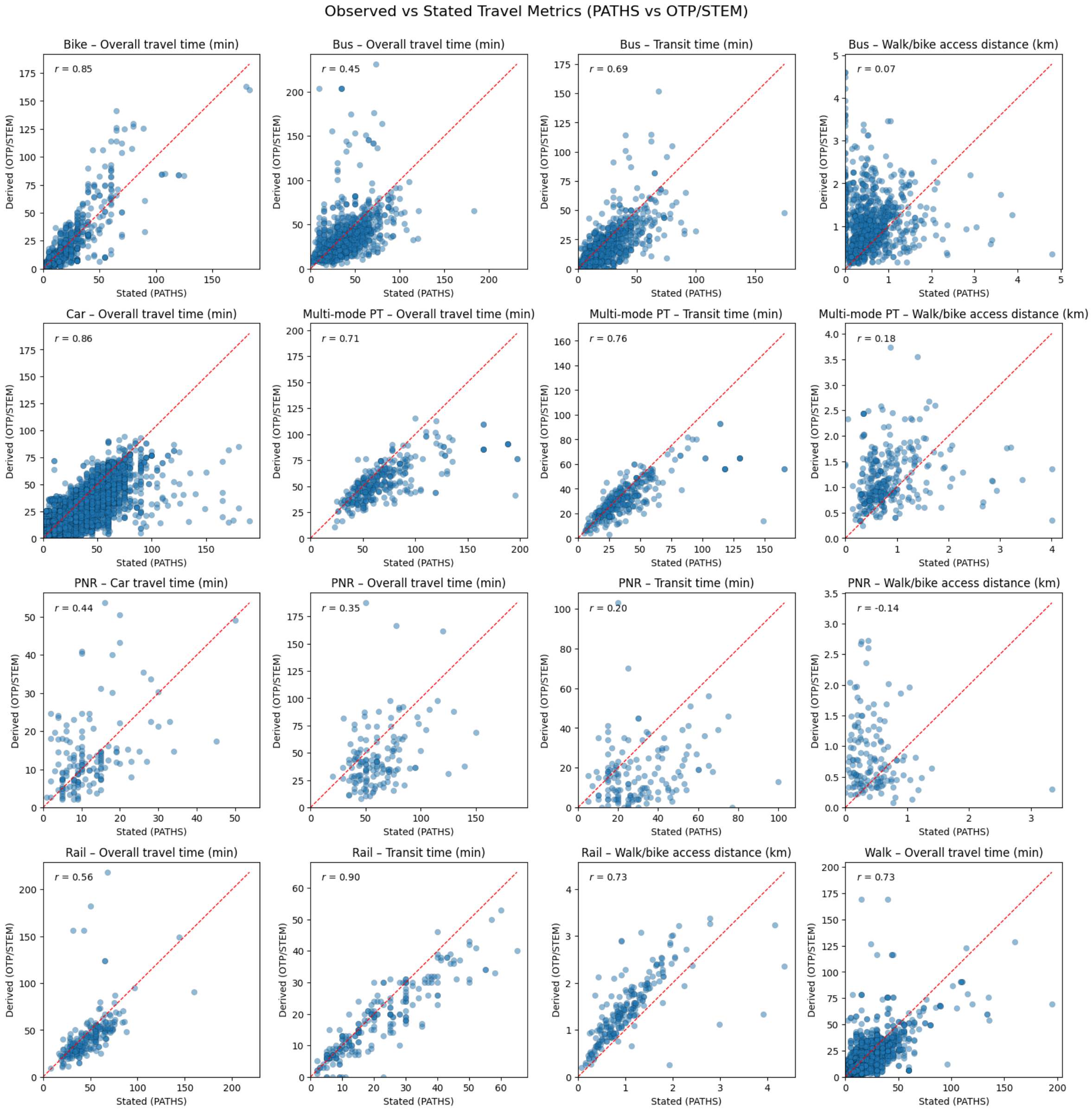


Figure 13: Comparison of observed and stated travel times and distances

6.1 Mode feasibility and model “availability”

Despite having access to OTP and GFTS, for a number of trips not all modes were considered feasible. Itineraries for pure PT trips (i.e. excluding PnR) where the access or egress distance is longer than 2km were not generated. Very short trips, including a subset of trips with the same start and end coordinates could not be generated. This also included a few trips of less than 100 m crow-fly length that could not be assigned to the OpenStreetMap network in a feasible way.

Even when network attributes could be calculated, travel times by PT and AT modes were very high compared to the car travel time, so we need to make a decision on the feasibility of making the trip by a non-car mode or park-and-ride (PnR). Table 10 gives a summary of these rules:

Table 10: Rules used to remove unfeasible trips

Mode	**Rule (to remove)**	**Number of trips removed**	**Remaining affected trips (unavailable)**
Bus	Total travel time over 100 min	29	3,176
Rail	Access time over 55 min	9	959
Taxi	Network distance over 40 km	1	237
Bike	Total travel time over 60 min and exceeds 4 times the car travel time	6	3,176
Walk	Distance over 1.8km	250	34,933

7 Revealed Preference (RP) Mode Choice

The revealed preference (RP) mode choice models presented in this section provide a foundational view of individual travel decisions, based upon observed mode choice and transport supply measurements obtained via network skims (Section 6). These models examine how travellers choose between alternative modes, based on measurable attributes like travel time, cost, and service characteristics. On their own, these provide meaningful insights and policy tools allowing analysts to estimate the relative influence of key factors on mode choice for each trip, supporting important planning tasks such as:

Forecasting for policy evaluation

While mode choice models alone cannot fully capture the broader impacts of new infrastructure or pricing, such as changes in destination choice, trip frequency, or network congestion resulting from others' decisions, they do offer valuable insight into expected shifts in transport mode preferences. By quantifying how travellers respond to attributes such as time, cost, and convenience, these models enable an assessment of the potential mode share impacts of proposed interventions (e.g., road extensions or improvements or public transport fare increases).

Evaluating the economic benefits of new investments

By quantifying how different factors (e.g., travel time, cost, reliability) affect travellers' utility, RP models allow for the estimation of changes in consumer surplus. These surplus calculations are central to cost-benefit analysis (CBA), enabling governments to justify investment in projects public transport, active mode infrastructure or roads. Mode choice models allow for an avenue to monetise the social impact of transport investment and compare this against the capital and operational costs of the intervention.

Transport Operations and Performance

By drawing on detailed network skim variables, RP trip-based models allow for a systematic examination of how different elements of the transport supply influence traveller utility**.** Models presented here decompose the travel experience into measurable components: access and egress times, transfer counts, headways and in-vehicle travel times. This structure enables analysts to identify which aspects of the public transport system most discourage use, allowing planners to identify which service improvements that may lead to higher levels of behavioural response.

We acknowledge that tour-based and activity approaches may better handle trip chaining and full day constraints, but by first establishing the patterns and sensitivities present in trip-level decisions, the current models lay the groundwork for transitioning to more behaviourally consistent tour-based models.

7.1 RP first trip of the tour

The first model is a nested logit (NL) tour-based model, considering as the response variable the mode of the 1st trip in each tour, as representative for the tour (Table 11).

Table 11: Nested Logit model – all first trips of tours

Parameter	Estimate	Rob. t-ratio	*p*-value
Alternate Specific Constants			
Bus	4.250***	9.18	0.000
Rail	3.260***	6.30	0.000
Car	0.290	0.71	0.476
Taxi	0	fixed	-
Bike	1.133**	3.04	0.002
Walk	5.760***	15.50	0.000
Park and Ride (PnR)	1.839***	3.76	0.000
Mode Travel Time Components			
In-vehicle time (Bus)	-0.030***	5.29	0.000
Access and egress time (Bus)	-0.068***	6.79	0.000
Half headway time (Bus)	-0.040***	4.75	0.000
Number of transfers (Bus)	-0.946***	5.25	0.000
In-vehicle time (Rail)	-0.026***	5.54	0.000
Access and egress time (Rail)	-0.028***	5.18	0.000
Half headway time (Rail)	-0.075***	4.32	0.000
Number of transfers (Rail)	-0.574***	4.89	0.000
Travel time (Bike)	-0.047***	12.30	0.000
Travel time (Walk)	-0.138***	20.65	0.000
In-vehicle time (Car)	-0.060***	13.52	0.000
Car travel time (PnR)	-0.001***	3.57	0.000
Taxi travel time (Taxi)	-0.040	1.94	0.053
Mode Travel Cost Components			
Travel cost (Car, Bus Rail, PnR)	-0.157***	25.47	0.000
Travel cost (Taxi)	-0.044*	2.51	0.012
Socio-Demographics			
Employment (Bus)	-0.074*	2.49	0.013
Age between 17-30 (Bus)	1.279***	7.92	0.000
Age between 30-55 (Bus)	-0.417**	2.71	0.007
Single Person HH (Bus)	-0.733***	4.55	0.000
Detached Dwelling (Bus)	-0.602***	3.62	0.000
Age between 17-30 (Rail)	1.274***	7.08	0.000
Age between 30-55 (Rail)	-0.400*	2.36	0.018
Single person household (Rail)	-0.963***	5.82	0.000
Is female (Bike)	-1.398***	7.85	0.000
Age over 75 (Walk)	-0.612***	7.27	0.000
Is employed (PnR)	0.799**	2.80	0.005

Parameter	Estimate	Rob. t-ratio	*p*-value
Vehicle ownership per household (Car)	0.531***	7.01	0.000
Has a driver licence (Car)	1.987***	14.67	0.000
Self-Reported Frequency for Taxi and Ride Share			
Rideshare frequency (Taxi)	1.836***	7.63	0.000
Rideshare frequency (Bus)	0.451**	2.86	0.004
Rideshare frequency (Car)	-0.585***	4.32	0.000
Inclusive Values (Null Hypothesis = 1)			
λ (PT modes)	0.418***	5.56	0.000
λ (Active modes)	1	fixed	-
λ (Car)	0.875***	3.00	0.003

Model Fit Statistics:
Log Likelihood: -5902.28; N = 17,131; N_{ind} = 9620; Number of parameters (K): 40
McFadden Pseudo-R^2 = 0.239, AIC: 11,884.56 (AIC/N = 0.65)
Note: Significance. Codes '***' 0.001 '**' 0.01 '*' 0.05

This NL model includes seven elemental alternatives: Bus, Rail, Car, Taxi, Bike, Walk, and PnR, estimated using data from 17,131 first trip of the tour across 9,620 individuals. McFadden's pseudo R^2 of 0.239 indicates solid explanatory power for a mode choice model.

Travel time components across all modes show the expected negative influence on utility. In-vehicle time is significant for all major modes: car (−0.060), bus (−0.030), and rail (−0.026), with larger penalties observed for active modes and, in particular, walk time (−0.138). Public transport users face additional burdens: access/egress time, service frequency and number of transfers are all strongly negative and statistically significant. Notably, transfer penalties are substantial (−0.946 for bus and −0.574 for rail).

The cost parameters in the RP mode choice model are negative and statistically significant, consistent with theoretical expectations that higher travel costs reduce the likelihood of choosing a given mode. This provides a basis for calculating willingness-to-pay (WTP) estimates in Table 12. The WTP values indicate that in-vehicle time for bus travel is valued at approximately $11.53/hour, while access and egress time is valued more highly at $26.10, and headway at $15.41. For rail, in-vehicle time is valued at $9.77, with access–egress and headway valued at $10.73 and $28.65, respectively. The value of a transfer is estimated at $6.04 for bus and $3.67 for rail. Car travel time is valued at $23.05, and taxi at $54.44, indicating variation in time sensitivity across modes.

Table 12: Willingness to pay for all first trips of the tour

Mode	In-Vehicle Time ($/hour)	Access/Egress Time ($/hour)	Headway / Wait Time ($/hour)	Transfer Penalty ($)
Bus	$11.53	$26.10	$15.41	$6.04
Rail	$9.77	$10.73	$28.65	$3.67
Car	$23.05	–	–	–
Taxi	$54.44	–		

The utility structure used here provided a basis for home-based work tours (HBW) and home-based other (HBO) coupled with non-home-based other (n-HBO). However, HBS required a structural change as Taxi (one observation) and PnR (6 observations) were rarely chosen. The output for RP first-trip of the tour-by-tour segments are presented in Appendix A (Section 12 of the report).

7.2 Multinomial logistic regression for tour mode choice

A multinomial logistic regression model was estimated to examine the predictors of the tour transport modes (Table 13). The model fits the data well (Chi-sq = 21262.692, 50 df, $p < 0.001$; with very significant Pearson and Deviance statistics and moderately high pseudo-R^2 values (over 0.3) indicating that the model captures a substantial amount of the variability in tour mode choice. Classification results (hit ratio of 89.9%) are also better than the classification by chance (65%). Compared to tours completed by other modes, Car tours are longer in distance, more complex, and associated with older travellers, who possess a driving license and are part of larger households, with greater mobility needs.

Table 13: Multinomial Logistic Regression results

Mode	**Predictor**	**B**	**Std. Error**	**Wald**	**df**	**Sig.**	**Exp(B)**
Car	Intercept	0.439	0.402	1.192	1	0.275	
	Cumulative Distance (km)	0.064***	0.009	47.009	1	<0.001	1.066
	Trip Time (min)	-0.056***	0.005	112.892	1	<0.001	0.945
	Age (years)	0.023***	0.004	36.497	1	<0.001	1.023
	HH daily trips (#/count)	0.027**	0.011	5.943	1	0.015	1.028
	Party size min 2 people	0.106	0.155	0.465	1	0.495	1.112
	Couple	0.26	0.192	1.845	1	0.174	1.297
	Driving licence	2.269***	0.175	168.965	1	<0.001	9.67
	Activities in tour	0.342***	0.073	22.165	1	<0.001	1.408
	Household (hh) size (#/count)	0.073	0.067	1.178	1	0.278	1.076
	[SEX=Male]	-0.185	0.125	2.169	1	0.141	0.831
Bus-Car	Intercept	-1.109*	0.496	4.992	1	0.025	
	Cumulative Distance (km)	-0.122***	0.015	66.014	1	<0.001	0.885
	Trip Time (min)	0.045***	0.006	56.125	1	<0.001	1.046
	Age (years)	-0.036***	0.006	40.154	1	<0.001	0.965
	Hh daily trips (#/count)	0.015	0.015	1.015	1	0.314	1.015
	Party size min 2 people	0.069	0.206	0.114	1	0.736	1.072
	Couple	-0.18	0.267	0.455	1	0.500	0.835
	Driving licence	0.299	0.228	1.717	1	0.190	1.349
	Activities in tour	0.593***	0.079	55.95	1	<0.001	1.809
	Household (hh) size (#/count)	0.063	0.089	0.503	1	0.478	1.065
	[SEX=Male]	-0.884***	0.176	25.235	1	<0.001	0.413
Car-Walk	Intercept	-2.109***	0.48	19.33	1	<0.001	
	Cumulative Distance (km)	-0.135***	0.016	67.681	1	<0.001	0.874
	Trip Time (min)	0.024***	0.007	13.142	1	<0.001	1.024
	Age (years)	0.002	0.005	0.157	1	0.692	1.002
	Hh daily trips (#/count)	0.065***	0.014	22.456	1	<0.001	1.067

Mode	Predictor	B	Std. Error	Wald	df	Sig.	Exp(B)
	Party size min 2 people	-0.332	0.195	2.906	1	0.088	0.717
	Couple	1.03***	0.255	16.368	1	<0.001	2.8
	Driving licence	-0.091	0.218	0.176	1	0.675	0.913
	Activities in tour	0.662***	0.075	77.911	1	<0.001	1.939
	Household (hh) size (#/count)	-0.413***	0.09	21.116	1	<0.001	0.662
	[SEX=Male]	-0.301	0.158	3.644	1	0.056	0.74
PT	Intercept	0.04	0.424	0.009	1	0.924	
	Cumulative Distance (km)	-0.042***	0.01	19.626	1	<0.001	0.959
	Trip Time (min)	0.043***	0.005	63.134	1	<0.001	1.044
	Age (years)	-0.005	0.004	1.354	1	0.245	0.995
	Hh daily trips (#/count)	-0.007	0.012	0.31	1	0.578	0.993
	Party size min 2 people	-0.894***	0.174	26.32	1	<0.001	0.409
	Couple	0.141	0.206	0.468	1	0.494	1.151
	Driving licence	0.058	0.186	0.097	1	0.756	1.059
	Activities in tour	0.299***	0.076	15.625	1	<0.001	1.349
	Household (hh) size (#/count)	0.08	0.072	1.223	1	0.269	1.083
	[SEX=Male]	-0.097	0.134	0.525	1	0.469	0.907
AT	Intercept	4.491***	0.424	112.365	1	<0.001	
	Cumulative Distance (km)	-1.224***	0.024	2690.538	1	<0.001	0.294
	Trip Time (min)	0.083***	0.006	211.211	1	<0.001	1.087
	Age (years)	0.016***	0.004	17.014	1	<0.001	1.016
	Hh daily trips (#/count)	0.061***	0.012	27.625	1	<0.001	1.063
	Party size min 2 people	-0.305*	0.16	3.623	1	0.057	0.737
	Couple	0.303	0.2	2.303	1	0.129	1.354
	Driving licence	1.27***	0.184	47.526	1	<0.001	3.561
	Activities in tour	-0.608***	0.078	61.055	1	<0.001	0.544
	Household (hh) size (#/count)	-0.233***	0.07	11.005	1	<0.001	0.792
	[SEX=Male]	-0.01	0.13	0.006	1	0.939	0.99

a. The reference category is: Other. The reference category for SEX = Female.
Note: Significance. Codes '***' 0.001 '**' 0.01 '*' 0.05.

The Bus-Car tours are associated with younger travellers and females, they are longer in duration and include more activities. Conversely, the Car-Walk tours are more likely undertaken by travellers living in couples, but without children at home.

What distinguishes the PT tours is that they are mostly undertaken by solo travellers. Tours by walking and cycling are associated with younger travellers, in smaller households, they are substantially shorter in distance and less complex.

These results show that tour mode choice reflects both mobility needs and distinct lifestyle patterns. While the dominance of car-based tours is expected, the distinctions among multi-modal tours (e.g., Bus-Car and Car-Walk) offer additional insights. They suggest that the integration and coordination of transport services, including schedules, transfers, and fare systems, can act as either enablers or barriers to daily mobility, especially for those engaging in more complex tours involving multiple stops and purposes. For example, individuals in households with a similar number of daily trips may exhibit different mobility strategies: one may complete fewer but more complex tours (stringing activities together efficiently), while another may make more fragmented tours with a simpler structure. Multi-modal travellers tend to have longer durations and more activities, often reflecting the travel patterns of younger and female users who may rely on PT as a key mode (Bus-Car) or being alone travellers (Car-Walk). Given the limited number of non-car-based complex and multimodal tours, enhancing multi-modal connectivity, reducing transfer costs, and designing services that accommodate this type of tours would significantly broaden mobility options and promote sustainable travel behaviours for Perth residents.

7.3 RP Tour-Based Choice Model

The tour-based model incorporated key findings from the multinomial logistic regression by including relevant tour attributes and user characteristics. Specifically, variables such as tour complexity, party size, and household structure were added to capture the diversity of real-world travel behaviour. The model confirmed that PT tours are predominantly undertaken by individuals travelling on their own, while more complex and multimodal tours tend to be associated with younger and female individuals. By accounting for these factors, the model enhances behavioural realism and aligns more closely with observed travel patterns, allowing for more accurate forecasting of mode choice across different traveller segments.

The tour-based mode choice model (Table 14) demonstrated a strong overall fit, with a McFadden pseudo-R^2 exceeding 0.35. This indicates that the model successfully captures meaningful variation in tour-level mode choice decisions. Travel time coefficients were statistically significant and in expected directions, though notably lower in magnitude than those observed in the corresponding trip-based models. A key implication is that value-of-time (VOT) estimates, particularly for PT, are substantially lower at the tour level. For instance, the estimated WTP for in-vehicle time on bus tours was $4.80/hour—less than half the $11/hour observed in the trip model—while the VOT for car travel was approximately $17/hour (Table 15). This suggests that when individuals consider an entire sequence of activities, sensitivity to travel time diminishes, possibly due to greater emphasis on flexibility, convenience, and the coordination of multiple activities.

Tour complexity was statistically significant but offered only modest explanatory power. Similar patterns were observed for group size: while PT riders were more likely to travel alone and car users were more likely to be accompanied by others, these effects were relatively minor compared to the dominant role of travel time and modal flexibility. The private vehicle's time and access advantage appears to be amplified in tour-based decisions, reinforcing its entrenched role in Perth's car-oriented travel landscape.

Table 14: Nested Logit Model for Tour Mode Choice

Parameter	**Estimate**	**Rob. t-ratio**	***p*-value**
Alternate Specific Constants			
Bus	2.845***	4.89	0.000
Rail	1.997***	3.39	0.001
Car	0.046	0.09	0.928
Taxi	-		-
Bike	-0.774	1.47	0.142

Parameter	Estimate	Rob. t-ratio	*p*-value
Walk	4.417***	8.29	0.000
Mode Travel Time Components			
In-vehicle time (Bus)	-0.009***	4.66	0.000
Access and egress time (Bus)	-0.023***	7.14	0.000
Half headway time (Bus)	-0.012***	3.35	0.001
Number of transfers (Bus)	-0.432***	6.87	0.000
In-vehicle time (Rail)	-0.005***	2.77	0.006
Access and egress time (Rail)	-0.011***	6.81	0.000
Half headway time (Rail)	-0.029***	6.31	0.000
Number of transfers (Rail)	-0.240***	5.23	0.000
Travel time (Bike)	-0.025***	16.1	0.000
Travel time (Walk)	-0.057***	32.55	0.000
In-vehicle time (Car)	-0.029***	10.83	0.000
In-vehicle time (Taxi)	-0.036***	3.3	0.001
Taxi travel time (Taxi)	-0.029**	2.54	0.011
Mode Travel Cost Components			
Travel cost (Car, Bus Rail, PnR)	-0.126***	28.27	0.000
Travel cost (Taxi)	-0.024**	2.49	0.013
Socio-Demographics			
Familiar with taxi (Bus)	0.395***	2.74	0.006
Single person household (Bus)	-0.311**	2.12	0.034
Number of working hours (Bus)	-0.074***	2.84	0.005
Age 1 (Bus)	1.372***	9.66	0.000
Age 2 (Bus)	-0.731***	5.56	0.000
House (Bus)	-0.471***	4.48	0.000
Single person household (Rail)	-0.337**	2.26	0.024
Age 1 (Rail)	1.309***	9.38	0.000
Age 2 (Rail)	-0.785***	6.26	0.000
Familiar with taxi (Car)	-0.358***	2.85	0.004
Vehicles per household (Car)	0.627***	10.03	0.000
No driver licence (Car)	1.914***	15.99	0.000
Post-COVID (Taxi)	-0.251	1.44	0.150
Familiar with taxi (Taxi)	1.968***	7.07	0.000
Female (Bike)	1.329***	8.22	0.000
Age 1 (Bike)	-0.123	0.66	0.509
Age 3 (Walk)	-0.444***	5.83	0.000
Single OC (Bus)	0.519**	2.23	0.026
Single OC (Rail)	0.558**	2.4	0.016
Single OC (Bike)	0.589**	2.03	0.042
Partial OC (Car)	0.411**	2	0.046

Parameter	**Estimate**	**Rob. t-ratio**	***p*-value**
Partial OC (Bike)	-0.324	0.88	0.379
Partial OC (Walk)	0.134	0.64	0.522
Vehicle occupancy (Car)	0.287	1.43	0.153
Tour complexity (Car)	-0.022	0.39	0.697
Tour complexity (Walk)	0.369***	4.14	0.000
Inclusive Values (Null Hypothesis = 1)			
λ (PT modes)	0.414***	8.54	0.000
λ (Active modes)	1	fixed	-
λ (Car)	0.897	1.43	0.153

Model Fit Statistics:
Log Likelihood: -6296.28; N = 17,635; N_{ind} = 9,826; Number of parameters (K): 47
McFadden Pseudo-R^2 = 0.349, AIC: 12,686.56 (AIC/N = 0.72)
Note: Significance. Codes '***' 0.001 '**' 0.01 '*' 0.05.

Table 15 presents the VOT for the models estimated considering the main four tour modes: Bus, Rail, Car, and Taxi.

Table 15: Willingness to pay for all main tour mode models

Mode	**In-Vehicle Time ($/hour)**	**Access/Egress Time ($/hour)**	**Headway / Wait Time ($/hour)**	**Transfer Penalty ($)**
Bus	$4.38	$10.97	$5.77	$10.97
Rail	$2.41	$2.41	$1.90	$14.03
Car	$13.70	–	–	–
Taxi	>$60.00	–		

7.4 Tour Complexity (Ordinal and Poisson Regression Models)

Two non-parametric models were estimated to evaluate the number of activities in a tour: ordinal regression and Poisson regression. Their findings converge, providing complementary findings.

They are presented next. Both models have good statistical measures. The ordinal logistic model has a Chi-square 266.358 (16df, p<0.001), with highly significant Pearson and Deviance statistics (<0.001). The scale effects, allowing for non-constant variances across subgroups, are an advanced and flexible approach that improved model fit over basic proportional odds and helped account for heterogeneity (e.g., males have more dispersion in their scores for tour complexity, see Table 16).

The results indicate significant thresholds estimates for complexity of the tour and strictly increasing, with no violations of the ordinal structure. There is no sign of threshold reversal, which would undermine the proportional odds assumption, thus the spacing between thresholds can be interpreted. Here, spacing increases from early thresholds (small differences) to later thresholds, which suggests an unbalanced distribution in the tour complexity, with very few tours including over nine activities.

Table 16: Ordinal Regression Model Results for Tour Complexity

Parameter Estimates	Variables	Estimate	Std. Error	Wald	df	Sig.
Threshold	[complexity_tour = 0]	-5.87***	0.263	496.449	1	<0.001
	[complexity_tour = 1]	0.218**	0.071	9.258	1	0.002
	[complexity_tour = 2]	1.537***	0.08	369.399	1	<0.001
	[complexity_tour = 3]	2.722***	0.109	627.335	1	<0.001
	[complexity_tour = 4]	3.668***	0.139	699.568	1	<0.001
	[complexity_tour = 5]	4.366***	0.163	714.854	1	<0.001
	[complexity_tour = 6]	5.092***	0.191	711.321	1	<0.001
	[complexity_tour = 7]	5.781***	0.219	694.341	1	<0.001
	[complexity_tour = 8]	6.539***	0.254	661.458	1	<0.001
	[complexity_tour = 9]	7.032***	0.28	631.68	1	<0.001
	[complexity_tour = 10]	7.367***	0.299	607.481	1	<0.001
	[complexity_tour = 11]	7.835***	0.329	568.038	1	<0.001
	[complexity_tour = 12]	8.703***	0.398	478.789	1	<0.001
Location	[Couple=0]	0.061	0.043	2.019	1	0.155
	[Driving license=0]	-0.422***	0.07	36.179	1	<0.001
	[SEX=Male]	-0.150***	0.026	34.312	1	<0.001
	Age (years)	0.004***	0.001	24.639	1	<0.001
	Household (HH) size (#/count)	-0.032*	0.014	5.07	1	0.024
	Total number bikes (#/count)	0.065***	0.011	36.558	1	<0.001
	Total number vehicles (#/count)	0	0.015	0	1	0.988
	HH median weekly income (thousands AUD)	0.019	0.013	2.364	1	0.124
Scale	[Couple=0]	0.042	0.022	3.76	1	0.052
	[Driving license=0]	-0.137***	0.041	11.204	1	<0.001
	Age (years)	0.001	0	3.713	1	0.054
	Household (hh) size (#/count)	-0.002	0.007	0.105	1	0.746
	[SEX=Male]	0.066***	0.012	27.574	1	<0.001
	Total number bikes (#/count)	0.006	0.005	1.286	1	0.257
	Total number vehicles (#/count)	0.016*	0.007	5.275	1	0.022
	HH median weekly income (thousands AUD)	0.018**	0.006	7.641	1	0.006

a. Parameters set to zero for reference categories (SEX=Female, Driving license =1, Couple=1).
Note: Significance. Codes ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05.

The location predictors shift the response distribution left or right on the latent (logit) scale. The most significant influences are from lack of a driving license (-0.422), decreasing the log-odds. As shown by the comparative analysis and Poisson regression (next), those who have a driving license are more likely to be in higher complexity categories. Males are less likely to engage in highly complex tours. Age, smaller households, and more bikes slightly increase the log-odds of higher complexity (albeit small effects). Household income and the number of vehicles have no significant effect on the complexity of tours.

The model in Table 16 estimated non-constant residual variances across groups, a generalised ordinal model allowing non-proportional dispersion (scale parameter). These scale parameters help the model relax the proportional odds assumption, effectively allowing different variances in the latent outcome across groups. Variance is smaller among those without a licence, i.e., responses are more tightly clustered, but it is larger among males. There is a slight increase in variance with more vehicles owned by households and higher income. Age, household size and number of bikes do not have a significant impact on variability.

Table 16 offers insights on the tour complexity modelled using Poisson Regression. The average and variance of tour complexity are comparable (1.86 and 1.74, respectively), and the Kolmogorov-Smirnov test marginally supports the Poisson distribution assumption (0.068).

The significant p-value of the omnibus test of the Poisson regression (Chi-sq=188.249, 8df, <0.001) suggests that the predictors, taken together, significantly improve model fit over the null model. Deviance and Pearson statistics indicate acceptable fit (p<0.001) and there is no major evidence of overdispersion (0.763 and 0.969).

Table 17 shows that the expected baseline log count of the tour complexity is 1.413**.** Household size, number of bikes, and gender are not significant and have negligible effects. There is a slight, non-significant decrease in complexity for males vs females (reference category). However, holding a driving license significantly increases the complexity of tours by 21.6%, also each vehicle in the household increases expected complexity by 1.9%. Finally, any additional $1,000 weekly income increases complexity by 1.1%. Being in a couple is associated with a 2.7% decrease in expected complexity and every additional year of age slightly increases expected complexity by 0.2% (Table 16).

Table 17: Results of Poisson Regression for Tour Complexity

Parameter	**B**	**Std. Error**	**95% Wald Confidence Interval**		**Hypothesis Test**			**Exp(B)**
			Lower	**Upper**	**Wald Chi-Sq.**	**df**	**Sig.**	
(Intercept)	0.346***	0.0311	0.285	0.407	123.1	1	<0.001	1.413
[SEX=Male]	-0.007	0.0079	-0.022	0.009	0.712	1	0.399	0.993
Age (years)	0.002***	0.0003	0.001	0.002	49.21	1	<0.001	1.002
Couple=1	-0.028	0.0138	-0.055	-0.001	4.045	1	0.044	0.973
Driving licence=1	0.196***	0.0247	0.147	0.244	62.63	1	<0.001	1.216
Household (HH) size (#/count)	-0.004	0.0045	-0.013	0.005	0.814	1	0.367	0.996
Total number bikes (#/count)	0.007	0.0045	-0.002	0.016	2.454	1	0.117	1.007
Total number vehicles (#/count)	0.019***	0.0033	0.013	0.025	33.18	1	<0.001	1.019
HH median weekly income (thousands AUD)	0.011**	0.004	0.003	0.019	7.246	1	0.007	1.011

Note: Signif. Codes '***' 0.001 '**' 0.01 '*' 0.05.

The results further confirm that the most influential factors (statistically and practically) are holding a driving licence, vehicle ownership and income (smaller but significant increases), with age having smaller but still meaningful effects.

7.5 Conclusions and recommendations from the tour-based model choice modelling

Both the tour-based model and the first-trip-of-tour approach led to similar results. The values of travel time savings are largely consistent across the two data representations, which is not unexpected, given that two-thirds of the tours are simple. While it is evident that tour complexity and shared travel may influence tour mode choice, these factors appear to be outweighed by the cost and time considerations made by households when selecting a travel mode, reflected in the car mode share of just over 80%.

The models presented here were estimated on all tours, but the utility specification is most appropriate for HBW and HBO tours, which comprise 60% of the data. Small adjustments were made for HBS, but other tour classifications, particularly SCH and SPS, would require further investigation, as shared travel is a more prominent factor. For SCH there is a case for introducing multimodal alternatives, such as Car–Walk and Car–Bus, into the decision set. Overall, we believe the models presented reflect the received data and will enable the modelling team to assess the estimation results against existing mode choice models used to support strategic transport decisions.

Despite the solid results, we note that there are areas of further investigation that could enhance our understanding of mode choice, and, more broadly, the mobility and activity decisions of Perth households. These avenues may offer a richer explanation of how and why travel patterns emerge in the local context.

- Shared travel requires further attention: Approximately 35% of tours involved multiple individuals, often travelling together for part of or for all the journey. This complicates assumptions of individual decision-making and suggests that future models should more explicitly account for joint travel behaviour, beyond simple “service passenger” tours.

- Spatial and socioeconomic context: Due to limited data sources, the current models do not fully incorporate urban form, spatial proximity, parking supply, or activity types at destinations. Similarly, income interactions with cost were not implemented due to time constraints, but represent a clear opportunity for future enhancement.

- Secondary data: modelled travel time was used from STEM matrices at a zone-to-zone level to estimate car travel times, whereas modelled journeys using road network and timetable data was used for public transport and active travel modes. While acceptably correlated when compared to observed travel times, further refinement of estimation processes and data sources could be used to further increase the accuracy of these times, such as network-level car travel time data incorporating delay and congestion measures.

- Coding of trips into tours: in some cases, it was observed that trips consisted of 50m walks between adjacent destinations, complicating estimation when in practice it only makes sense to walk between different elements of a single activity, A refinement of the coding used into generating trips could assist in better understanding mode choice behaviour.

8 Stated Preference (SP) Data

This section includes results from the PATHS stated preference (SP) survey, which gathered information on respondents' general travel behaviour based on a selected 'pivoted' trip from their daily activity patterns. Responses to four stated choice experiments with eight attributes and basic socio-demographics represented the main variables in the dataset. It is important to note that the unit of analysis in the SP survey was the individual trip, not the entire chain or tour of trips (Ipsos 2019: 4). Consequently, integration with revealed preference (RP) data could only be achieved at the trip level.

8.1 Sample description

The SP dataset includes records on 3,147 respondents, who were asked four choice scenarios with four alternatives (except for 131 respondents, with 2 or 3 alternatives per scenario) and three respondents completed only three scenarios[2].

As anticipated, 79% of the trips used a car as the main mode, followed by active travel at 9.1%, Taxi/Uber/Rideshare at 6.2%, with the remaining trips using public transport accessed by various modes (Table 18). It is important to note that the mode share for walking is significantly lower than the share reported for all trips (as discussed in Section 3) because the reference trip for the SP experiment was typically a middle- or long-distance urban trip.

Table 18: Mode choice for the chosen RP trip

Main mode	**Count**
Car driver	2,447
Car passenger	42
PnR	111
WnR	22
KnR	44
Taxi/Share	194
Cycle	279
Walk	8

The focus on SP was on commuting trips (22.5%), followed by shopping (21%), social (12.6%), recreational (8.9%), personal business (7.4%), pick-up/delivery (2.6%) and school/education (2.3%) (Figure 14).

---

[2] Person ID: Y19H1340103P01, Y19H0930804P01 and Y19H1761312P01

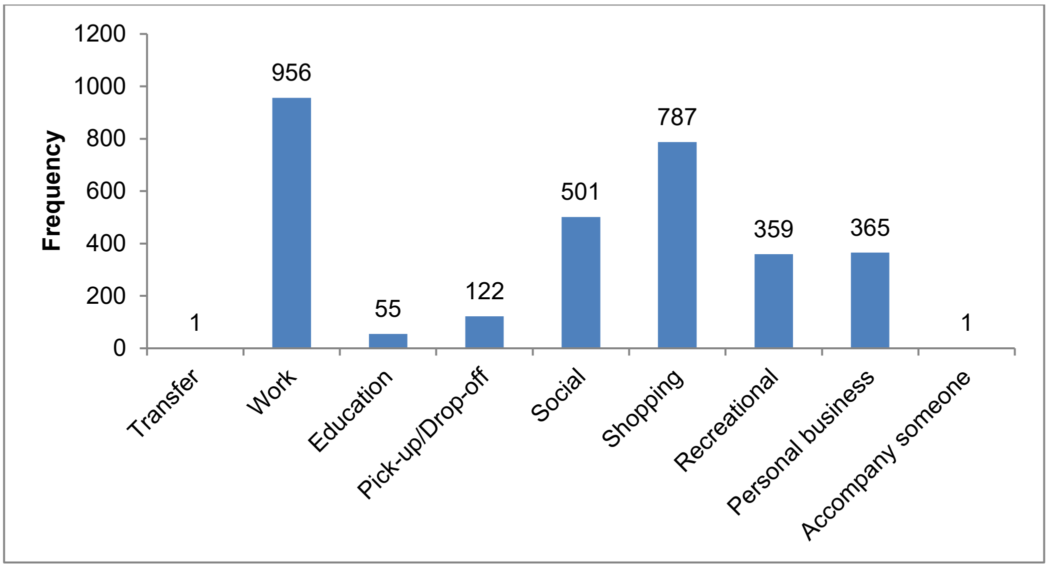


Figure 8: Distribution of trip purposes in the SP

As indicated, for each scenario, respondents were asked to choose between four alternatives (mostly) out of six modes (Table 19). In some cases, only two or three options were considered based on their availability. For example, 'Car drive' or 'PnR' assumed that the individual had access to a car.

Table 19: Aggregated choice statistics for SP

	1 Car drive	**2 PnR**	**3 Walk to PT (WnR)**	**4 Taxi/Share**	**5 AT - Cycle**	**6 AT - Walk**
Not Chosen	2,654	**6,264**	**6,353**	**7,014**	**7,411**	**7,362**
Chosen	**5,401**	1,732	2,073	1,330	1,035	1,014
Total	8,055	7,996	8,426	8,344	8,446	8,376

Within the main transport mode, the SP differentiates between Bus, Light rail, and Train as options for PT.

While the trip purpose seemed to have been accounted for in the SP, similar to the RP distribution of trip purposes, giving more prominence to work, shopping, social, and recreational trips, the mode for the chosen trip was evenly considered in the experiments (Figure 15).

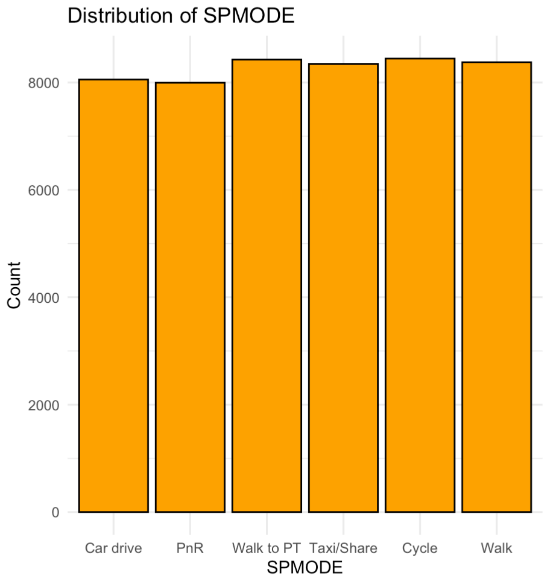


Figure 15: Distribution of the Main Transport Mode (SPMODE)

Section 8.2 describes the attribute values for the SP design.

8.2 Attribute levels

The attributes used in the design included: in-vehicle travel time or walking/cycling time for active travel and access/egress, combined waiting and transfer time, parking availability, operating costs/fares and parking costs, main mode of PT and number of transfers, and for taxi/Uber/Rideshare, whether the ride is shared with strangers. As illustrated in Figure 16, the attribute levels were pivoted based on the reported travel time and cost for the chosen RP mode, specifically in relation to car driving. Notably, walking and cycling are included as access/egress modes for car driving, but are not accounted for in Taxi, Uber, or Rideshare modes, despite their relevance to these modes. The SP design served as the basis for informing our choice models, as detailed in Section 7.

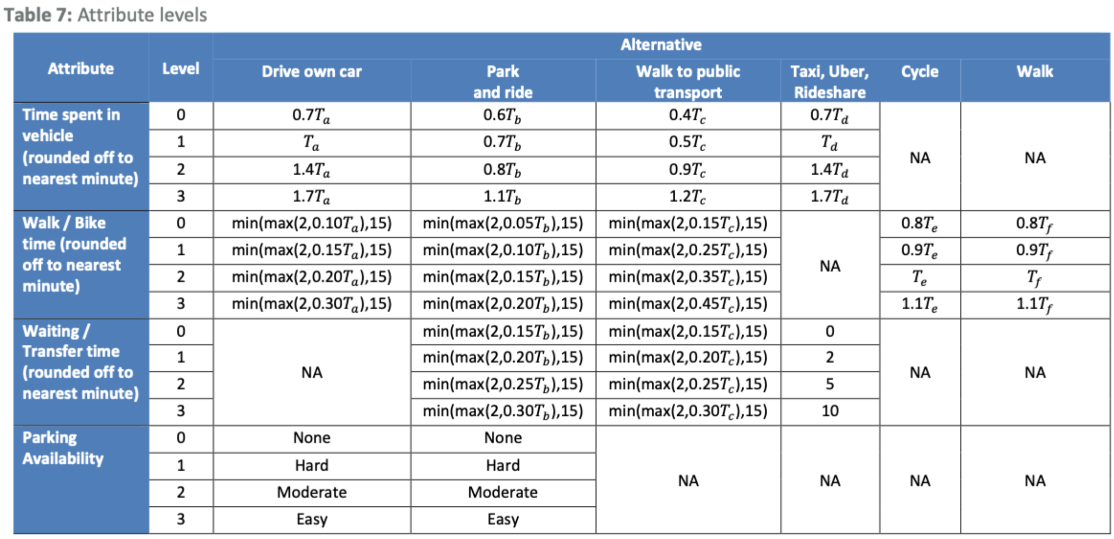


Table 7: Attribute levels

Attribute	Level	Alternative					
		Drive own car	Park and ride	Walk to public transport	Taxi, Uber, Rideshare	Cycle	Walk
Time spent in vehicle (rounded off to nearest minute)	0	$0.7T_a$	$0.6T_b$	$0.4T_c$	$0.7T_d$	NA	NA
	1	T_a	$0.7T_b$	$0.5T_c$	T_d		
	2	$1.4T_a$	$0.8T_b$	$0.9T_c$	$1.4T_d$		
	3	$1.7T_a$	$1.1T_b$	$1.2T_c$	$1.7T_d$		
Walk / Bike time (rounded off to nearest minute)	0	min(max(2,0.10T_a),15)	min(max(2,0.05T_b),15)	min(max(2,0.15T_c),15)	NA	$0.8T_e$	$0.8T_f$
	1	min(max(2,0.15T_a),15)	min(max(2,0.10T_b),15)	min(max(2,0.25T_c),15)		$0.9T_e$	$0.9T_f$
	2	min(max(2,0.20T_a),15)	min(max(2,0.15T_b),15)	min(max(2,0.35T_c),15)		T_e	T_f
	3	min(max(2,0.30T_a),15)	min(max(2,0.20T_b),15)	min(max(2,0.45T_c),15)		$1.1T_e$	$1.1T_f$
Waiting / Transfer time (rounded off to nearest minute)	0	NA	min(max(2,0.15T_b),15)	min(max(2,0.15T_c),15)	0	NA	NA
	1		min(max(2,0.20T_b),15)	min(max(2,0.20T_c),15)	2		
	2		min(max(2,0.25T_b),15)	min(max(2,0.25T_c),15)	5		
	3		min(max(2,0.30T_b),15)	min(max(2,0.30T_c),15)	10		
Parking Availability	0	None	None	NA	NA	NA	NA
	1	Hard	Hard				
	2	Moderate	Moderate				
	3	Easy	Easy				

Figure 9. Screenshot SP Attribute levels (Ipsos, 2019: 13, Table 7)

The pivoting on the car trips has created situations where the attribute levels were unrealistic even for a hypothetical scenario, particularly for travel times. The histograms presented in Figures 17-19 show pronounced positively skewed distributions, which would affect any model estimation.

These outlying cases were eliminated from the data used in the choice models, as explained below.

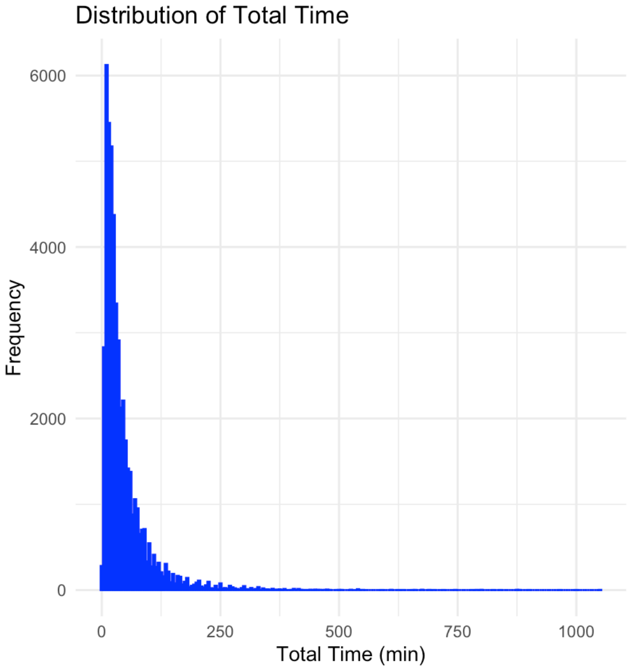


Figure 17: Distribution of Total Time (min)

The skewed distributions of travel times stem from long car trips (over 60 min), which would significantly increase door-to-door travel time if undertaken by cycling or walking. As indicated, we recognise that the realism of the choice tasks is likely to have been influenced by these durations, leading us to exclude options with total trip durations exceeding 150 min from the choice models (Figures 18-19).

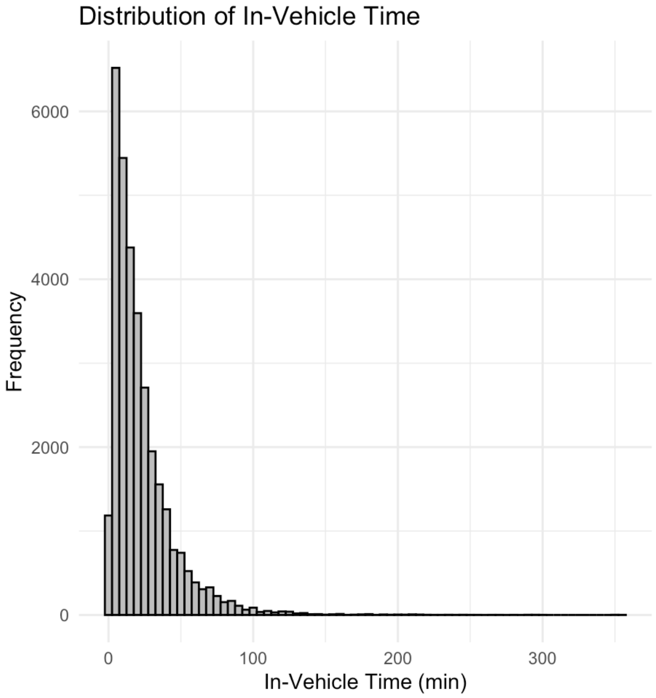


Figure 18: Distribution of In-Vehicle Time (min)

We reiterate that the primary challenge with attribute values in the SP design consisted in the inclusion of AT options with durations that exceed reasonable walking and cycling times for typical daily trip purposes such as commuting, shopping, social activities, recreation, personal errands, pick-up/delivery, and school/education.

To address this, we decided to exclude choice situations where the total travel time exceeded 150 min, particularly if active travel options were over 120 min. While ‘reengineering’ the choice set formation, we have understood the correspondence between travel times across different modes, accounting for their respective speeds (e.g., a 5 km trip might take 5 min by car, 10 minutes by PT, 25 min by bike, and 60 min walking), the inclusion of options with these large travel times is unfitting.

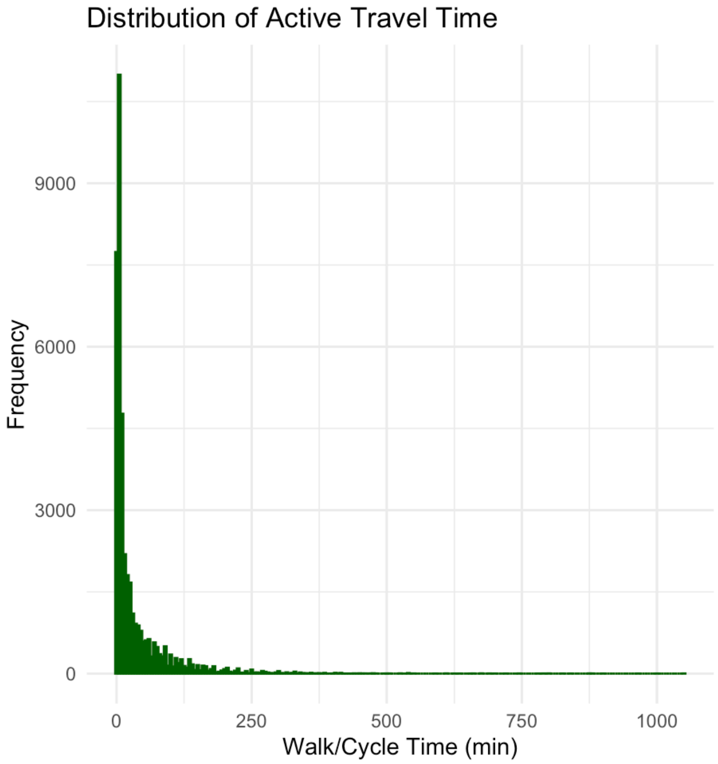


Figure 19: Distribution of Active Travel AT (min)

Figures 20 and 21 provide pertinent transfer times between various transport modes and frequencies of transfers, thus they were left unchanged. Additionally, the unusually high parking cost values were included as per the SP design, as they are considered unlikely to bias the model's choice results (Figure 22).

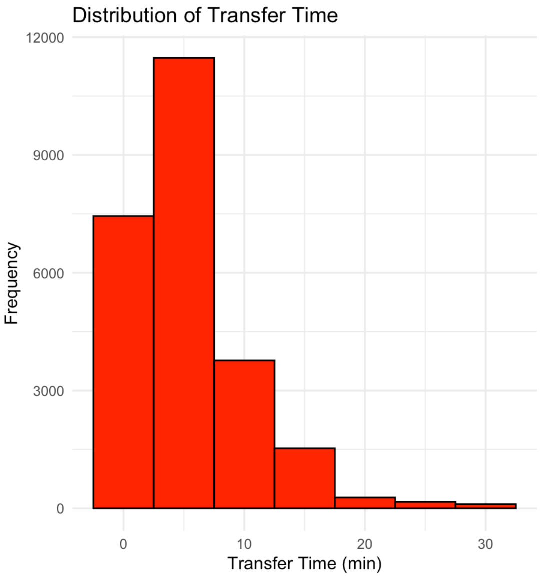


Figure 20: Distribution of Transfer Time (min)

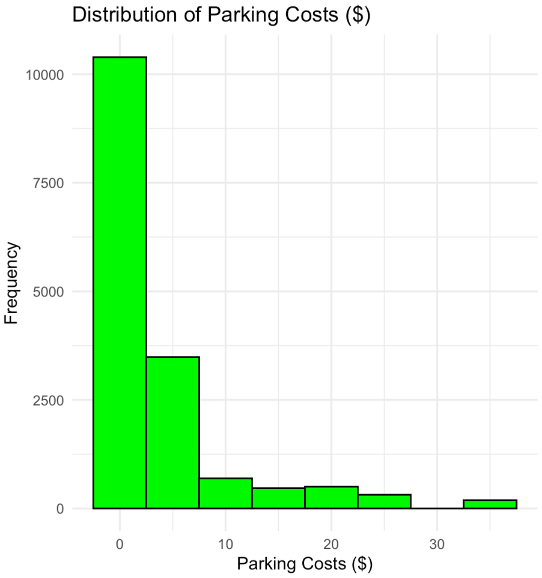


Figure 21: Distribution of parking costs ($)

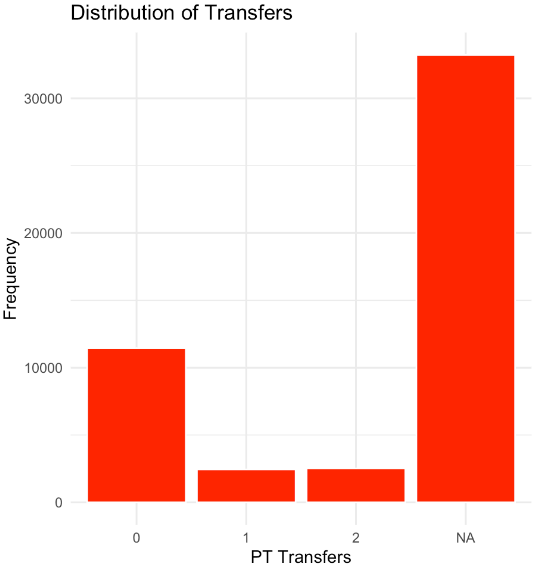


Figure 22: Distribution of transfers

While the distribution of operation costs and fares is positively skewed (Figure 23), there was only one record eliminated from the analysis (e.g., Y19H1440906P01, work trip with a cost of $250, and scenario costs between $19.90 and $32.40).

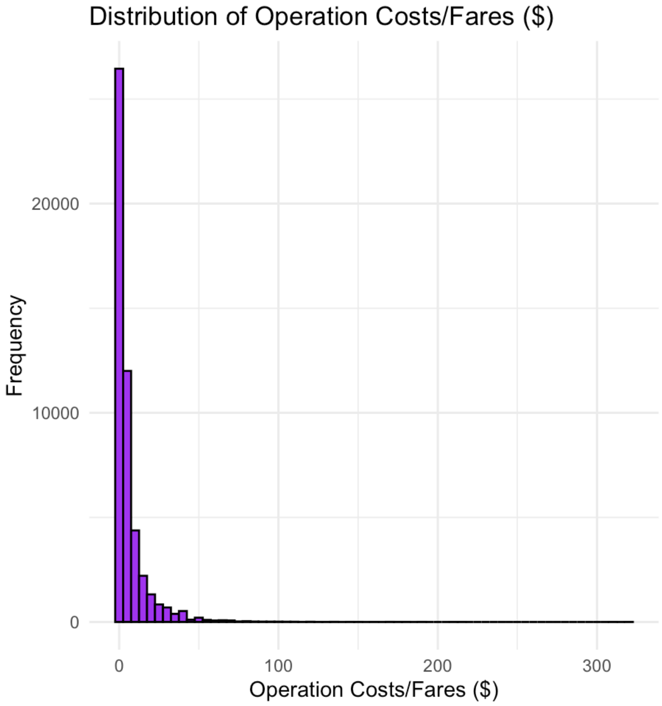


Figure 23: Distribution Operation Costs/Fares ($)

Figures 24 to 27 provide boxplots that visualise more clearly the presence of outliers. The charts further confirm the issue of unrealistic travel times resulted from simple conversion of driving times into active travel times, with disregard for the feasibility of using active travel for these trips. There are two options for dealing with such occurrences: a) eliminating choice sets that include extreme outliers (over 25% of the data); and b) eliminating only options with these extreme outliers (above 8% of the data) and reorganising the choice sets.

The team has completed the analysis with both options and compared the results of the two processes in Appendix B of this report.

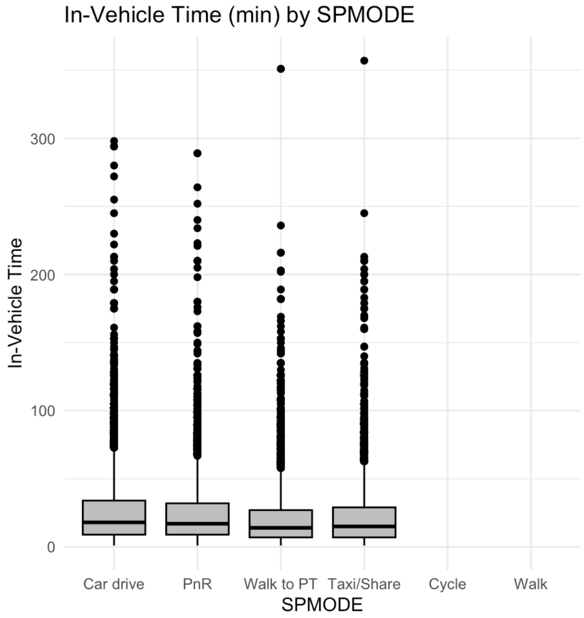


Figure 10 In-Vehicle time by mode

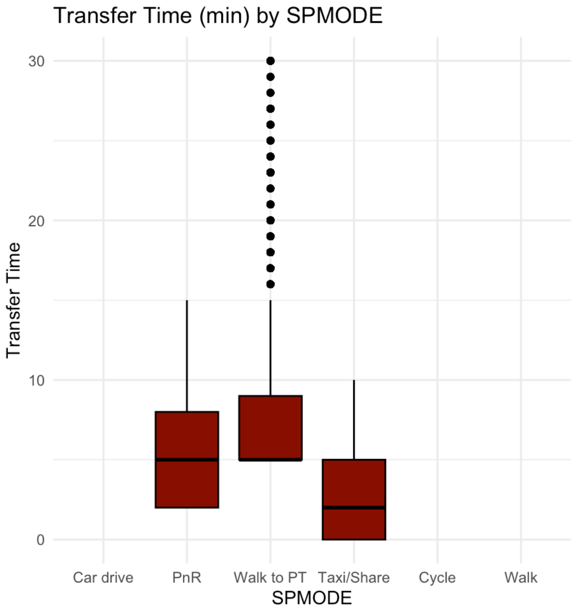


Figure 25: Transfer time by mode

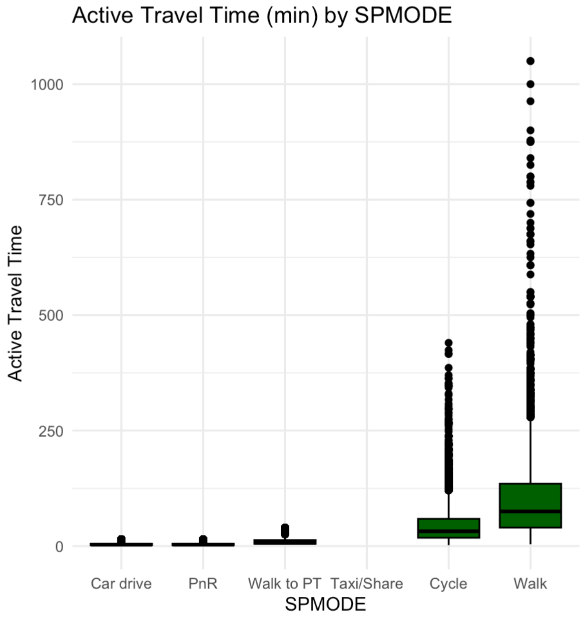


Figure 26: Walking and cycling time by mode

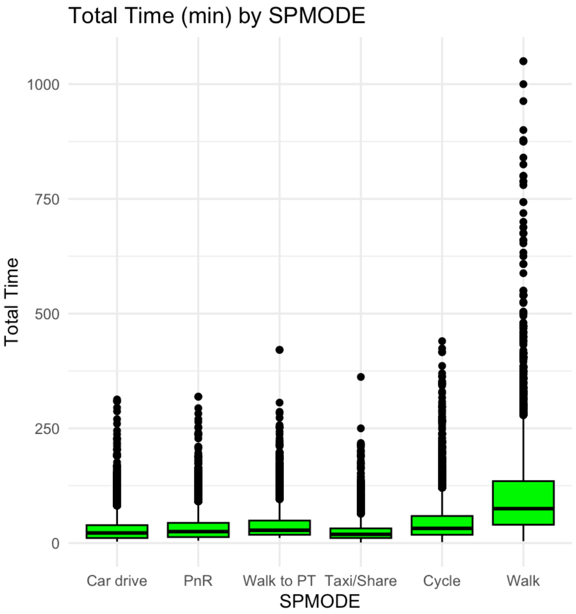


Figure 27: Total time by mode

8.3 Comparison of RP and SP values

For users of the SP data, a key concern is the apparent disconnect between the trip analysed in the SP survey and the corresponding reference trip from the RP data. This discrepancy is critical, as the analysis aims to integrate these two data sources effectively. To illustrate the severity of the mismatch, consider the following examples:

- **Y19H0690528P01:** The reference car trip was 80 minutes with a reported cost of $99, but the SP scenario reflects a short trip of only 5 km.

- **Y19H0790213P01:** The reference car trip was 50 minutes with a cost of $9.63, yet the SP scenario includes durations ranging from 49 to 350 minutes (the latter for active travel) with costs up to $112.

- **Y19H1500219P01:** The reference car trip of 145 minutes (TRAVTIME_REFTRIP) is inconsistent with the reported DURATION_REFTRIP of 5 minutes for a 5.9 km trip. The SP scenario's attribute levels appear unchecked, with travel times varying between 138 and 875 minutes and costs as high as $175.

While additional examples could be provided, these cases underscore the challenges in integrating data from both stated preference (SP) and revealed preference (RP) sources, even at the trip level.

Furthermore, descriptive statistics reveal several outliers in the RP trips, which may represent special travel circumstances. For instance, some reference trips reported durations of less than three minutes, which were used to construct SP scenarios that are unrealistic and should therefore excluded from the analysis. Conversely, six records with travel times of 1,200 minutes were also removed as outliers.

In addition, the random trip presentation in the SP experiment from each linked tour of the participant, remained unjustified. It was found that using the first trip of each tour had the most explanatory power (Section 7, Appendix A). As such, when joined to RP data, SP has limited value to the model when the second or further trip was used within the SP experiment.

8.4 Decision on the choice tasks for modelling

As outlined in Appendix A, a thorough preliminary analysis was undertaken to assess the quality and usefulness of the data as received. This review revealed instances of unrealistic travel times, characterised by extreme values that were either excessively large or implausibly short. Such data anomalies risk skewing parameter estimates and distorting behavioural interpretations in discrete choice models. To address these concerns, we excluded choice tasks containing unrealistic travel times, resulting in a 22% reduction in the stated preference (SP) dataset. The final dataset used for analysis included 38,605 valid choice tasks, down from the original 49,643. This reduction aimed to enhance the validity of the data and ensure that the models reflect credible/plausible travel decision-making processes.

Appendix A provides results from the RP mode choice for three tour purposes: 1) home-based work; 2) home-based and non-home-based other; and 3) home-based shopping. The mode is for the first trip of the tour, more relevant for the choice of the tour especially if the availability of the mode remains for the following trips in the chain (Hasnine & Habib, 2018). The results confirm the role of time and cost with comparable VOT obtained by RP mode choice at the trip level.

Appendix B presents the results of multinomial logit (MNL) and latent 2-class choice models (LCM) estimated using both the original and reduced datasets. The analysis demonstrates improved model fit and behavioural realism when using the reduced dataset. For example, coefficients associated with travel times and costs became more consistent with expectations, and anomalies such as unrealistic preferences for active travel modes for long trips were corrected. Based on these findings, we will proceed with modelling using the reduced dataset to ensure that the results are robust, reliable, and reflective of genuine travel behaviour. This step is essential for drawing meaningful insights that can support effective policy evaluation and decision-making.

Conversely, choice modelling for RP data enhanced/integrated with SP information poses its own challenges. The imputation of the attributes of the non-chosen alternatives, which are still relevant to the traveller requires addressing two decisions: 1) determining the suitable choice set; 2) estimating the attribute values for the non-chosen alternatives.

Unfortunately, the first decision is particularly problematic and cannot be informed by the SP data. Many options suggested in the SP are unrealistic and may not align with real-world context and individual circumstances. For the second decision, the imputation of the travel times and fares has been completed considering network data and traffic conditions and the available public information on timetables. Yet, many other attributes influence the choice of modes (or the main RP mode) and the imputation presented difficulties. For example, key factors such as operation costs, parking, combination with active travel, required further assumptions, which introduce uncertainty and highlight the need for sensitivity analyses to assess how variations in imputed data influence the results.

9 Choice modelling results for SP

This section presents the stated preference (SP) data analysis incorporating RP features, with an emphasis on uncovering the substitution patterns and unobserved heterogeneity that influence decision-making. The modelling approach incrementally progressed from a multinomial logit (MNL) model to more sophisticated frameworks, including availability and nested logit (NL) models. These advancements aim to better capture the complex relationships between mode attributes and contextual factors, such as state dependence and tour characteristics.

An availability model was used as an intermediary step to identify substitution patterns among modes by varying their presence across choice scenarios. This provides critical insights into the error structure, guiding the specification of a nested logit (NL) model. The nested structure allows for correlated error terms within groups of similar modes, addressing the independence assumptions of the MNL model and offering a more nuanced view of mode substitution.

By including state dependence—reflecting respondents' revealed mode choices—and attributes such as tour complexity and duration, the model provides a detailed understanding of how individual and trip-specific factors shape transportation decisions.

9.1 The multinomial logit (MNL)

The first step in analysing the travel mode choice involved estimating a Multinomial Logit (MNL) model, which serves as a baseline framework to evaluate the effects of mode-specific attributes on choice behaviour. This model was applied to the 'cleaned' dataset, after removing extreme outliers, to ensure robust parameter estimation. Socio-demographic variables and revealed preference (RP) information were initially excluded, allowing the analysis to focus solely on the stated preference (SP) data and the impact of travel-specific attributes. Then, their additional contribution to the explanatory power of the models was assessed.

The utility functions for each mode were specified based on their defining characteristics. For example, the MNL with socio-demographics included the following utility functions:

The Utility for Car: accounts for in-vehicle travel time, access time (walking), operational costs, parking costs and parking availability. In terms of social-demographic variables we have included the respondent's age and whether they hold a class C driver's licence.

U(CAR) = ACSCAR + BTTCAR*IN_VEH_TIME + BACC*ACCESS_EGRESS + BCOST*OPERATIONAL_COST + BPARKCOST*PARK_COST + BPARKAVCAR*PARK_AVAILABILITY+ BAGE*AGE +BLIC*LICENCE

(1)

The utilities for public transport (PT) are presented by access mode PnR and WnR; Attributes include whether the main mode was rail (base) or bus (entered in the utility), in-vehicle travel time, wait time, the fare, number of transfers, walking access time (WnR) and parking availability (PnR). Social demographics include whether the respondent participates in the workforce. Note that parameters for access time and operational cost (fares) are constrained across car and public transport (after testing for differences):

U(PNR)= APNR + BBUS*BUS + BTTPT*IN_VEH_TIME + BTTWAIT*WAIT_TIME+BCOST*FARE+ BTRF*NUMBER_TRANSFERS + BPARKPNR*PARK_AVAILABILITY+BWORK *WORK

(2)

U(WNR)= APNR + BBUS*BUS + BTTPT*IN_VEH_TIME + BTTWAIT*WAIT_TIME + +BCOST*FARE + BTRF*NUMBER_TRANSFERS +BACC*ACCESS_EGRESS + BWORK *WORK (3)

Taxis and shared mobility: includes in-vehicle travel time, cost, and whether the service is shared. The age of the respondents is also included. Travel costs are constrained across all modes, but travel time and age are unique parameters.

U(TAXI)= BTTTAXI*N_VEH_TIME+BCOST*FARE +BSHARE*SHARE_WITH_OTHERS +BAGE_TAXI*AGE (4)

Active travel options (cycling, walking): include travel time for walking or cycling and whether the respondent is female.

U(CYCLE)= ACS_CYCLE + BTTCYC**TRAVEL_TIME* + BFEMCYC*FEMALE (5)

U(WALK)= ACS_WALK + BTTWALK**TRAVEL_TIME* + BFEMWLK*FEMALE (6)

The MNL model results, presented in Table 20, provide a foundational understanding of mode choice preferences under the assumption of independent and identically distributed (IID) error terms. While the MNL model is a straightforward approach, it sets the stage for more sophisticated analyses to address limitations such as substitution patterns and correlated error structures. The coefficients have the expected signs, negative for time and costs, reflecting a preference for minimising travel time and expenses, and positive for parking availability, albeit not statistically significant. Wait time for public transport is non-significant and this may be a balance between risk (insufficient wait time to be certain of making the service) and the dislike for longer waiting times. It is perhaps that 2 minutes presents a high amount of risk to respondents.

Age and possessing a driver's license are positively associated with car travel, being in the workforce is positively associated with using public transport, and age also significantly and positively influences the choice of taxi. The interaction terms indicate that females are less likely to cycle and walk.

Table 20: Results of the MNL RP model

Variable Description	**Parameter Estimate**	**Asymptotic z**	***p*-value**
Alternate Specific Constants			
ASC for car	1.887***	9.55	0.000
ASC for park and ride (PnR)	-0.262	1.14	0.255
ASC for walk and ride (WnR)	0.269	1.42	0.155
ASC for cycling	1.514***	7.92	0.000
ASC for walking	2.772***	13.75	0.000
Mode Attributes			
Operation cost / fares (Car, Taxi, PnR,WnR)	-0.030***	8.93	0.000
In-vehicle travel time (car)	-0.013***	4.04	0.000
In-vehicle travel time (PnR, WnR)	-0.011***	3.46	0.001
In-vehicle travel time (taxi)	-0.018***	4.08	0.000
Travel time (cycling)	-0.052***	16.24	0.000
Travel time (walk)	-0.061***	25.03	0.000
Walking access time (car, WnR)	-0.031***	3.87	0.000
Parking availability (car)	0.036	1.67	0.095
Parking availability (PnR)	0.054	1.50	0.133

Variable Description	**Parameter Estimate**	**Asymptotic z**	***p*-value**
Parking cost (car)	-0.414*	11.57	0.000
Waiting time (WnR, PnR)	-0.001	0.06	0.952
Number of transfers (WnR, PnR)	-0.046	1.31	0.191
Main mode is ‘bus’ (WnR, PnR)	-0.019	0.38	0.705
Sharing the ride (taxi)	0.080	1.11	0.269
Socio-Demographic Variables			
Age (car)	0.012***	6.54	0.000
Holds a driver licence (car)	1.661***	7.59	0.000
Participates in the workforce (WnR, PnR)	0.395***	6.28	0.000
Age (taxi)	0.005*	2.35	0.019
Is female (cycling)	-0.627***	8.11	0.000
Is female (walk)	-0.247**	2.83	0.005
Model Fit Statistics: Log Likelihood: -9,792.72; N = 9,709; Number of parameters (K): 26 McFadden Pseudo-R^2 = 0.074 AIC: 19,825.4 (AIC/N = 2.043); Chi-square: 10,758 (df = 21, p < 0.001)			

Note: Signif. Levels: ***p < 0.001, **$p < 0.01$, *$p < 0.05$.

The willingness to pay for travel times savings derived from this model are $24.50/hr for car, $22.75/hr for PT and $36.30/hr for shared mobility and taxi, consistent with reported times in Australia and elsewhere, and slightly higher than the evaluations presented in Section 7.

9.2 Availability multinomial logit to investigate substitution patterns (A-MNL)

The availability models were used to investigate substitution patterns among transport modes, providing insights into how the presence or absence of specific options influences choice behaviour. Initially, we explored a comprehensive availability model where all potential cross-substitutions were considered; however, this model did not converge due to its complexity. A simplified model revealed significant substitution patterns, particularly between PnR, WnR, and Taxi modes relative to the presence of Car. These findings highlight the dominant preference for Car and the perception of other motorised modes as substitutes. Additionally, there was a notable substitution effect between PnR and WnR, suggesting that respondents perceive these modes as sharing characteristics beyond the observable utility variables. Similar patterns were observed between the two active modes, Cycle and Walk, highlighting their shared behavioural attributes.

A refined model structure was developed to capture the key substitution effects between private modes (Car and Taxi), public modes (PnR and WnR), and active modes (Cycle and Walk). This structure aligns more intuitively with the data and is supported by the substitution patterns observed. The final availability model, presented in Table 21, incorporates these insights and serves as the basis for informing the specification of our NL model. By explicitly considering the availability of alternatives and their impact on substitution patterns, this model improves our understanding of mode hierarchy and the interdependencies among choices, laying the groundwork for further, more robust analysis based on the PATHS datasets.

Table 21: Results of the A-MNL model

Variable Description	Parameter Estimate	Asymptotic z	P-value
Alternative-Specific Constants (ASC)			
ASC for car	1.528***	7.19	0.000
ASC for PnR	-0.641**	2.72	0.010
ASC for WnR	0.186	0.93	0.350
ASC Cycling	1.247***	6.12	0.000
ASC Walking	2.455***	11.58	0.000
Attributes			
Operation cost/Fares (Car, Taxi, PnR, WnR)	-0.033***	9.54	0.000
In-vehicle travel time (car)	-0.012***	3.86	0.000
In-vehicle travel time (PnR, WnR)	-0.012***	3.78	0.000
In-vehicle travel time (Taxi)	-0.013***	2.88	0.000
Travel time (Cycling)	-0.051***	16.36	0.000
Travel time (Walk)	-0.060***	24.79	0.000
Walking Access time (car, WnR)	-0.036***	4.89	0.000
Parking availability (car)	0.029	1.34	0.180
Parking availability (PnR)	0.081*	2.15	0.030
Parking cost (car)	-0.041***	11.31	0.000
Waiting time (WnR, PnR)	0.000	0.01	0.990
Number of transfers (WnR, PnR)	-0.100***	2.87	0.000
Main mode Bus (WnR, PnR)	-0.004	0.09	0.930
Sharing the ride (Taxi)	0.083	1.11	0.270
Demographics			
Age (car)	0.011***	6.33	0.000
Holds a driver’s licence (car)	-1.672***	7.64	0.000
Participates in the workforce (WnR, PnR)	0.348***	5.73	0.000
Age (Taxi)	0.005*	2.03	0.040
Female (Cycling)	-0.607***	7.92	0.000
Female (Walk)	-0.222***	2.56	0.010
Presence/Absence Parameters			
U(Car | presence of Taxi)	-0.002	0.04	0.970
U(Taxi | presence of Car)	-0.825***	9.37	0.000
U(PnR | presence of WnR)	-0.29***	3.84	0.000
U(WnR | presence of PnR)	-0.295***	4.46	0.000
U(Cycle | presence of Walk)	-0.277***	3.54	0.000
U(Walk | presence of Cycle)	-0.242**	2.70	0.010
Model Fit Statistics: Log Likelihood: -9,708.72; N = 9,709; Number of parameters (K): 32 McFadden Pseudo-R^2 = 0.082, AIC: 19,480.1 (AIC/N = 2.006)			

Note: Signif. Levels: ***$p < 0.001$, **$p < 0.01$, *$p < 0.05$.

9.3 Nested logit model (NL)

The Nested Logit (NL) model extends the analysis of mode choice by explicitly addressing the correlation in unobserved components of utility within groups of similar transport modes. While the MNL model assumes independence across alternatives and the availability model (A-MNL) introduces a way to identify substitution effects, the NL model, built on the findings from the availability model, organised the choice alternatives into three intuitive and behaviourally meaningful nests, mentioned above:

- Private Vehicles (Car, Taxi)
- Public Transport (PnR, WnR)
- Active Travel (Cycle, Walk)

By introducing correlation structures within these nests, the NL model better captures substitution patterns among similar modes while preserving interpretability and analytical elegance. This approach aligns with the findings from the availability model A-MNL, where substitution parameters revealed strong relationships between closely related modes, such as PnR and WnR, and Cycle and Walk. However, rather than relying on availability attributes alone, the NL model leverages an explicit nested structure to account for these patterns systematically.

The NL model (Table 22) advances the analysis of mode choice by addressing correlations in the unobserved components of utility within groups of similar transport modes. Compared to the MNL the NL model offers a significant improvement in model fit, as evidenced by a lower AIC and a higher McFadden's pseudo-R^2. This enhancement reflects the NL model's ability to capture the shared utility components among alternatives within well-defined nests, as in this case. The utility coefficients for travel time and cost remain negative and significant, consistent with earlier models, while their moderated magnitudes indicate the NL model's success in accounting for correlations within nests.

The Inclusive Value (IV) parameters for the private vehicle and PT nests fall within the range of 0 and 1, confirming the presence of substitution patterns and correlation in the unobserved components of utility within these groups. The PT nest exhibits particularly strong substitution, with a low IV parameter highlighting the close relationship between PnR and WnR as complementary multimodal options. Similarly, the private vehicle nest shows strong substitution between Car and Taxi, reinforcing a hierarchy where Car is the dominant choice and Taxi serves as a contingency solution. In contrast, the IV parameter for the active travel nest is not significantly different from 1, suggesting weaker or inconclusive substitution between Cycle and Walk. This may reflect greater heterogeneity in how these modes are perceived or used, influenced by individual preferences or contextual factors such as trip distance and infrastructure availability.

The NL model provides statistical support for organising alternatives into intuitive and behaviourally meaningful nests and offers an elegant balance between analytical sophistication and interpretability. Compared to a potential cross-nested logit model, the NL framework is simpler to communicate and aligns naturally with practical transport categories. By capturing robust substitution patterns in the private vehicle and public transport nests and highlighting the nuances within active travel, the NL model provides valuable insights for transport policy and planning. These results demonstrate the NL model's superiority over earlier frameworks, delivering a comprehensive and behaviourally realistic understanding of mode choice.

Table 22: Results NL Model

Variable Description	Parameter Estimate	Asymptotic z	P-value
Alternative-Specific Constants (ASC)			
ASC for car	2.337***	8.91	0.000
ASC for PnR	-0.753*	2.13	0.033
ASC for WnR	0.249	0.81	0.418
ASC Cycling	1.242***	5.20	0.000
ASC Walking	2.524***	9.62	0.000
Attributes			
Operation cost/Fares (Car, Taxi, PnR, WnR)	-0.046***	9.17	0.000
In-vehicle travel time (car)	-0.016***	3.95	0.000
In-vehicle travel time (PnR, WnR)	-0.019***	4.22	0.000
In-vehicle travel time (Taxi)	-0.019***	3.70	0.000
Travel time (Cycling)	-0.055***	8.76	0.000
Travel time (Walk)	-0.063***	15.53	0.000
Walking Access time (car, WnR)	-0.049***	5.15	0.000
Parking availability (car)	0.045	1.57	0.117
Parking availability (PnR)	0.088	1.79	0.074
Parking cost (car)	-0.056***	10.84	0.000
Waiting time (WnR, PnR)	-0.003	0.19	0.848
Number of transfers (WnR, PnR)	-0.138**	2.80	0.005
Main mode Bus (WnR, PnR)	-0.046	0.72	0.471
Sharing the ride (Taxi)	0.102	1.06	0.288
Demographics			
Age (car)	0.016***	6.25	0.000
Holds a driver's licence (car)	-2.306***	8.11	0.000
Participates in the workforce (WnR, PnR)	0.599***	4.79	0.000
Age (Taxi)	0.008**	2.70	0.007
Female (Cycling)	-0.651***	7.13	0.000
Female (Walk)	-0.256***	2.67	0.008
Inclusive Parameters (statistical test against null = 1)			
Inclusive Parameter (Car, Taxi)	0.662***	12.41	0.000
Inclusive Parameter (PnR, WnR)	0.566***	6.90	0.000
Inclusive Parameter (Cycling, Walk)	0.923	1.07	0.242
Model Fit Statistics: Log Likelihood: -9,713.56; N = 9,709; Number of parameters (K): 29 McFadden Pseudo-R^2 = 0.081, AIC: 19,484.1 (AIC/N = 2.007)			

Note: Signif. Levels: ***p < 0.001, ***p < 0.01, *p < 0.05.*

9.4 Estimating SP with RP trip and tour information

Table 23 shows the MNL model with set dependence and tour information. Travel time negatively impacts utility across all modes, with Cycling and Walking being the most sensitive. Parking availability positively affects car use, while parking costs are significantly negative. In contrast, variables like waiting time, number of transfers, and bus usage for PnR and WnR show no significant influence, suggesting limited impact when other factors are considered.

State dependence parameters highlight strong correlations in the repeated mode choices, with significant positive coefficients for all modes, particularly for active travel (Cycling and Walking), indicating substantial inertia in individuals' mode selection. Demographic variables also play a critical role; age and workforce participation significantly influence WnR and PnR choices, while holding a driver's license strongly reduces the likelihood of choosing non-car modes. Revealed preference information from travel diaries further enriches the model, with the number of tours negatively associated with utility and total travel time positively associated with selecting car as the mode of travel.

The NL model given in Table 23 improves upon the MNL framework by accounting for hierarchical structures in mode choices. As before, the inclusive value parameters indicate significant substitution patterns within mode nests, particularly for Car and Taxi (0.658, $p < 0.001$) and PnR and WnR (0.803, $p < 0.001$), reflecting strong interdependencies among these options. Similar to the NL model presented in Table 19, Cycling and Walking exhibit weaker substitution, as their inclusive parameter is not statistically significant. The model underscores the importance of hierarchical nesting in capturing the interrelationships between similar modes, offering a more behaviourally realistic framework for mode choice analysis.

Key parameters reveal nuanced insights into travel behaviour. In-vehicle travel time remains a critical factor across modes, with particularly steep penalties for active travel modes (e.g., Cycling: -0.052, Walking: -0.047). Parking availability positively influences car use, although marginally (0.042, $p = 0.057$), while parking cost has a substantial negative impact (-0.04, $p < 0.001$).

Demographic factors further shape preferences. Consistent with previous models, Age positively influences Car use and Taxi adoption, while being female reduces the likelihood of choosing Cycling or Walking. State dependence remains strong across modes, indicating persistent habitual travel patterns. As indicated by A-MNL and by the NL models, the availability of the mode is a critical factor for mode choice. This is particularly relevant for car choice, where possessing a driver's license loses its effect in choosing the car as the travel mode when analysed together with the mode availability (-1.498, $p < 0.001$).

Table 23: Results MNL with RP trip and tour information

Variable Description	**Parameter Estimate**	**Asymptotic z**	**P-value**
ASC for car	2.051***	9.70	0.000
ASC for PnR	-0.269	1.15	0.248
ASC for WnR	0.234	1.22	0.224
ASC Cycling	1.285***	6.64	0.000
ASC Walking	2.016***	9.84	0.000
Attributes			
Operation cost/Fares (Car, Taxi, PnR, WnR)	-0.030***	8.85	0.000
In-vehicle travel time (car)	-0.015***	4.65	0.000
In-vehicle travel time (PnR, WnR)	-0.012***	3.73	0.000
In-vehicle travel time (Taxi)	-0.017***	3.97	0.000
Travel time (Cycling)	-0.052***	15.40	0.000
Travel time (Walk)	-0.047***	19.62	0.000
Walking Access time (car, WnR)			
Parking availability (car)	0.042	1.91	0.057
Parking availability (PnR)	0.052	1.43	0.153
Parking cost (car)	-0.04***	10.82	0.000
Waiting time (WnR, PnR)	-0.002	0.13	0.893
Number of transfers (WnR, PnR)	-0.040	1.13	0.259
Main mode Bus (WnR, PnR)	-0.009	0.19	0.852
Sharing the ride (Taxi)	0.097	1.32	0.186
Demographics			
Age (car)	0.011***	6.22	0.000
Holds a driver's licence (car)	-1.498***	6.68	0.000
Participates in the workforce (WnR, PnR)	0.385***	6.14	0.000
Age (Taxi)	0.007***	2.86	0.004
Female (Cycling)	-0.577***	6.77	0.000
Female (Walk)	-0.245***	2.61	0.001
Revealed preference information			
Number of daily tours	-0.079**	2.79	0.005
Total daily time travelling	0.003**	2.78	0.005
State Dependence, PnR	1.819***	7.07	0.000
State Dependence, WnR	1.911***	10.98	0.000
State Dependence, Taxi	1.85***	6.17	0.000
State Dependence, Cycle	4.524***	11.80	0.000
State Dependence, Walk	2.847***	15.73	0.000
Model Fit Statistics: Log Likelihood: -9,369.71; N = 9,709; Number of parameters (K): 32 McFadden Pseudo-R^2 = 0.113, AIC: 18,803.4 (AIC/N = 1.937)			

Note: Signif. Levels: ***p < 0.001, ***p < 0.01, *p < 0.05.*

The goodness-of-fit measures show once again a significant improvement of the NL over MNL model with RP trip and tour information.

Table 24: Results nested logit (NL) with RP trip and tour information

Variable Description	**Parameter Estimate**	**Asymptotic z**	**P-value**
ASC for car	2.576***	9.15	0.000
ASC for PnR	-0.533*	2.01	0.045
ASC for WnR	0.080	0.37	0.715
ASC Cycling	1.202***	5.16	0.000
ASC Walking	1.763***	7.12	0.000
Attributes			
Operation cost/Fares (Car, Taxi, PnR,WnR)	-0.048***	9.31	0.000
In-vehicle travel time (car)	-0.022***	4.88	0.000
In-vehicle travel time (PnR, WnR)	-0.014***	3.85	0.000
In-vehicle travel time (Taxi)	-0.02***	3.67	0.000
Travel time (Cycling)	-0.057***	9.51	0.000
Travel time (Walk)	-0.049***	13.71	0.000
Walking Access time (car, WnR)	-0.037***	3.78	0.000
Parking availability (car)	0.061*	2.03	0.042
Parking availability (PnR)	0.068	1.62	0.105
Parking cost (car)	-0.055***	10.30	0.000
Waiting time (WnR, PnR)	-0.001	0.10	0.919
Number of transfers (WnR, PnR)	-0.092*	2.25	0.025
Main mode Bus (WnR, PnR)	-0.030	0.53	0.595
Sharing the ride (Taxi)	0.121	1.23	0.217
Demographics			
Age (car)	0.016***	6.04	0.000
Holds a driver’s licence (car)	-2.111***	7.24	0.000
Participates in the workforce (WnR, PnR)	0.459***	5.59	0.000
Age (Taxi)	0.009***	2.95	0.003
Female (Cycling)	-0.611***	6.51	0.000
Female (Walk)	-0.258**	2.52	0.012
Revealed preference information			
Number of daily tours	-0.102***	2.60	0.009
Total daily time travelling	0.003*	2.19	0.028
State Dependence, PnR	1.908***	6.53	0.000
State Dependence, WnR	2.101***	9.93	0.000
State Dependence, Taxi	2.611***	6.96	0.000
State Dependence, Cycle	4.846***	9.28	0.000
State Dependence, Walk	3.02***	11.96	0.000
Inclusive Parameters			
Inclusive Parameter (Car, Taxi)	0.658***	12.53	0.000
Inclusive Parameter (PnR, WnR)	0.803***	3.36	0.000
Inclusive Parameter (Cycling, Walk)	0.907	1.42	0.272

Model Fit Statistics:
Log Likelihood: -9,306.14; N = 9,709; Number of parameters (K): 35
McFadden Pseudo-R^2 = 0.119, AIC: 18,682.4 (AIC/N = 1.925),

Note: Signif. Levels: ***p < 0.001, ***p < 0.01, *p < 0.05.*

While the findings have not changed qualitatively, the nested logit structure sharpened the effect of the attributes, socio-demographics, and state dependence/habit. RP information on the number of daily tours negatively impacting the utility and mode choice and the total travel time positively associated with selecting Car as the mode of travel, confirm the current empirical evidence supporting convenience and travel time savings.

The integration of RP and SP data revealed notable challenges. Substantial data cleaning was necessary to include in the analysis only realistic choice sets and options. This resulted in excluding over 20% of the SP data. Moreover, not all SP trips are 'connected' to an RP tour and pivoted on the main mode. Despite these limitations, the cleaned SP data still offers insights into mode choice behaviour and supported the joint RP-SP modelling framework.

State dependence, or inertia in mode choice, is strong, particularly for private vehicles, PnR, and WnR modes. However, the inclusion of tour complexity attributes, such as the number of tours undertaken in a day, showed only limited and negative associations with stated preferences for car use. This may suggest that busy individuals and households with multiple tours tend to prefer distinct modes of transport to satisfy their mobility needs. As shown by the RP statistics, the length and duration of a tour decrease as the number of tours increases. This relationship is also reflected in the opposite signs of the number of tours and total daily travel time. Additional investigations with the total number of stops per tour, indicated weak, non-significant links to mode choice for the trip.

10 Joint RP-SP Choice Models

While the joint estimation of RP and SP data aimed to enhance behavioural insight for Perth's mode choice modelling, several limitations emerged. A key issue was the mismatch between how transport agencies define travel modes and how they were presented in the SP design. This was compounded by criticisms of the SP instrument itself—namely, whether the presented scenarios were realistic and feasible. These concerns are reflected in the results: the RP data offered substantially stronger explanatory power than the SP data.

Initial models, which relied on reference trips from respondents' tours, often failed to reflect representative journeys, leading to poor model fit. A more effective approach used only the first trip of each tour in the RP data, improving stability and allowing the SP component to be reintegrated into a stronger joint model. In the final model (Table 22), the RP segment achieved a significant improvement in log-likelihood (from -31,272.39 to -5,999.33), while the SP segment remained weaker (-9,952.42 from -13,278.77). Although the SP component helped validate the direction of some key parameters (e.g., transfer penalties, walk and access time), the RP model emerged as the more robust basis for strategic modelling in Perth.

The parameter estimates from the jointly estimated RP and SP mode choice model (Table 25) align well with behavioural theory. Time-related coefficients are consistently negative, reflecting the disutility of longer travel durations. Walking (-0.169) and cycling (-0.068) attract the highest penalties, consistent with the physical effort required. Car travel time is also valued negatively (-0.054), even among car users, while bus (-0.028) and rail (-0.016) show smaller time sensitivities. The relatively low penalty on rail time may reflect its higher service quality and longer average tour distances (over 50 km), supporting its perceived efficiency.

Transfer penalties are substantial for both bus (-0.59) and rail (-0.12), underscoring the deterrent effect of needing to change services. Wait times are also negatively valued—especially for bus (-0.027), but not significant for rail (-0.015)—highlighting the importance of service frequency. Access time to rail is less penalised than for bus, suggesting that travellers are more tolerant of longer access trips when using faster, higher-capacity rail modes.

Socio-demographic variables provide further behavioural nuance. Older respondents are more likely to choose PT, while single-person households are less inclined to use bus or rail. Notably, car access remains a structural determinant: individuals without a licence are strongly deterred from car use (+2.007), while vehicle-rich households are more likely to drive (+0.516). Familiarity with taxis increases their likelihood of being chosen in stated preference settings.

Table 25: Results Joint RP-SP Model

Parameter	**Estimate**	**Rob. t-ratio**	***p*-value**
Alternate Specific Constants			
Bus	-2.106***	5.77	0.000
Rail	-3.225***	6.73	0.000
Car	-5.941***	20.44	0.000
Taxi	-6.863***	33.99	0.000
Bike	-4.584***	27.54	0.000
Walk	0	fixed	-
Park and Ride (PnR)	-0.028***	25.41	0.000
Mode Travel Time Components RP SP joint estimates			
In-vehicle time (Bus)	-0.028***	6.89	0.000
Access and egress time (Bus)	-0.065***	6.33	0.000
Half headway time (Bus)	-0.027**	2.60	0.009
Number of transfers (Bus)	-0.59***	5.19	0.000
In-vehicle time (Rail)	-0.016***	5.10	0.000
Access and egress time (Rail)	-0.483***	4.22	0.000
Half headway time (Rail)	-0.015	1.42	0.155
Number of transfers (Rail)	-0.12***	4.84	0.000
Travel time (Bike)	-0.068***	11.01	0.000
Travel time (Walk)	-0.169***	31.64	0.000
In-vehicle time (Car)	-0.054***	12.12	0.000
Car travel time (PnR)	-0.0008**	3.21	0.012
Taxi travel time (Taxi)	-0.029**	2.54	0.011
Mode Travel Cost Components RP SP joint estimates			
Travel cost (Car, Bus Rail, PnR)	-0.159***	29.42	0.000
Travel cost (Taxi)	-0.057***	6.00	0.000
Socio-Demographics			
Employment (Bus)	-0.108***	2.70	0.007
Age between 17-30 (Bus)	1.244***	6.43	0.000
Age between 30-55 (Bus)	-0.395**	2.13	0.033
Single Person HH (Bus)	-0.557***	3.23	0.001
Detached Dwelling (Bus)	-0.965***	7.18	0.000
Age between 17-30 (Rail)	1.263***	5.31	0.000
Age between 30-55 (Rail)	-0.435*	1.94	0.053
Single person household (Rail)	-0.958***	4.36	0.000
Is female (Bike)	-1.427***	8.04	0.000
Age over 75 (Walk)	-0.633***	7.28	0.000
Any Employment (PnR)			
Vehicle ownership per household (Car)	0.516***	0.08	0.000
Has a driver licence (Car)	2.007***	0.14	0.000
Self-Reported Frequency for Taxi and Ride Share			

Parameter	Estimate	Rob. t-ratio	*p*-value
Rideshare frequency (Taxi)	2.192***	0.23	0.000
Rideshare frequency (Bus)	0.652**	0.21	0.002
Rideshare frequency (Car)	-0.528***	0.15	0.000
Alternate Specific Constants			
SP car	8.126***	13.73	0.000
SP park and ride (PnR) rail	0.369	0.66	0.511
SP walk and ride (WnR): rail	1.616***	4.97	0.000
SP walk and ride (PnR): bus	0.593	1.12	0.264
SP walk and ride (WnR): bus	1.74***	4.82	0.000
Mode Travel Cost Components SP only estimates			
Parking Availability (Car, PnR)	0.114	0.99	0.321
Wait time (WnR, PnR)	0.007	0.20	0.842
Reference Trip Purpose			
Reference Trip is Work (Car)	-0.688**	2.21	0.027
Reference trip is recreation (Walk)	1.365**	2.54	0.011
Reference trip is social (Taxi)	1.443***	3.80	0.000

Model Fit Statistics:
Log Likelihood: -15,927.78; N = 26,815; N_{ind} = 9,766; Number of parameters (K): 58 McFadden Pseudo-R^2 = 0.199, AIC: 31,971.56 (AIC/N = 1.19)
Note: Significance. Codes ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05.

11 Operationalising mode choice models

To test the ‘Donor approach’, we focused on Home-Based Work (HBW) tours. These, along with Home-Based Other (HBO), are the largest proportion of the dataset and produced the most stable results in initial estimation. When approximating the Donor model, we retained the provided parameter constraints and nesting structure, but did not constrain the alternative specific constants (ASCs) or their interactions. The Donor model is expected to have calibrated ASCs to match observed mode shares. By freely estimating these parameters, we effectively calibrated our models to the observed shares in PATHS.

A key issue was that the Donor model classified Rail and Bus as a single mode. We maintained that Rail and Bus are distinct and competing alternatives, since Rail offers a higher-speed option for longer trips. This difference complicated direct comparisons with the donor model. We also found strong evidence not to include Park-and-Ride (PnR) in the Transit nest. PnR shares more unobserved characteristics with private vehicles than with Transit. This is consistent with observed behaviour in Perth, where many travellers drive significant distances to access the Mandurah and Yanchep rail lines. Figures 27a and 27b illustrate the differences in nesting, with Figure 27a showing the donor-constrained structure and Figure 27b the revised structure estimated from PATHS.

Figure 27a. Model error structure from the Donor Model

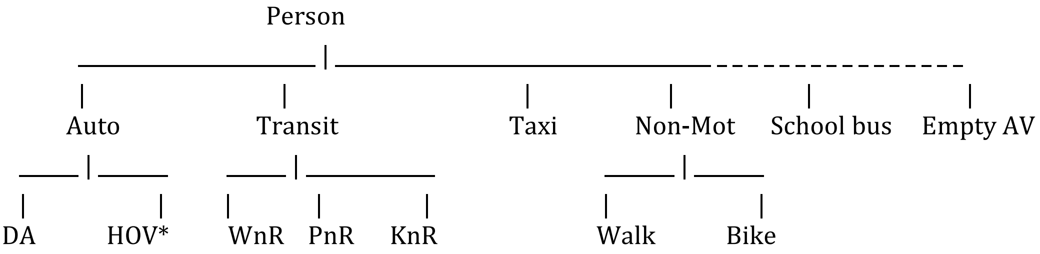
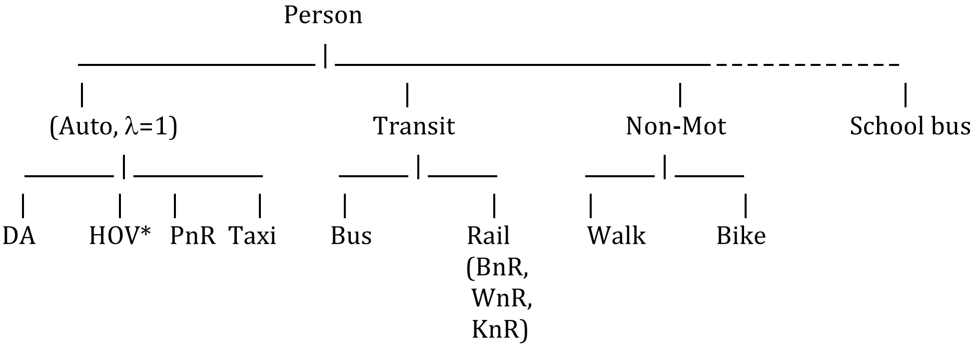


Figure 27b. Model error structure for PATHS estimates



To examine the differences between models, we focused on the structure (Figures 27a vs 27b) and parameter constraints. Only cost and time parameters were constrained, with all other parameters being freely estimated. The results, therefore, show only a partial contrast. While constraining cost and time parameters produced a substantial loss of explanatory power, these results do not capture the full mismatch between the donor model and the PATHS data.

We tested four model specifications (Table 26).

- **Model 1:** Estimation on PATHS data using a simplified utility space. This represents a balance between accuracy and parsimony and is intended for operationalisation.
- **Model 2:** Imposing the Donor nesting structure but freely estimating all parameters.
- **Model 3:** Imposing both the nesting structure and the cost and time parameters from the Donor model.
- **Model 4:** Keeping the Donor parameter constraints but applying the nesting structure revealed from PATHS data.

Table 26: Model summaries

Model	Log Likelihood	McFadden's R^2	λ (PT)	λ (Active)	λ (Car)
Model 1: Simplified estimation model with new PATHS structure	–2,823.88	0.302	0.35	0.87	0.97
Model 2: Donor structure, free estimation	–2,836.35	0.299	1.06	0.88	1.31
Model 3: Donor structure with parameter constraints	–3,129.82	0.226	0.75	0.36	20.49
Model 4: New PATHS structure with Donor parameter constraints	–3,132.14	0.226	0.57	0.32	1.2

The results given in Table 26 suggest that Model 1 provides the strongest fit and the most behaviourally plausible substitution patterns. It balances parsimony and accuracy and is therefore the recommended model for operationalisation in STEM. By contrast, Model 2 shows that when parameters are freely estimated under the Donor nesting structure, the model produces unstable λ values; this highlights the incorrect error structure including for PnR in the nest, with the other public transport alternatives. Models 3 and 4 impose donor parameter constraints, which result in even weaker fit and lower McFadden's R^2. They also generate unrealistic λ values, particularly in the Car nest where values exceed acceptable limits. This indicates that the Donor cost and time sensitivities are unlikely to be transferable to Perth conditions.

Hence, the simplified estimation model (Model 1) best represents observed behaviour. It separates Rail and Bus, treats PnR as car-like, and produces consistent sensitivities to time and cost. We feel that this model structure is more suitable for forecasting and for economic evaluation of projects in Perth.

To illustrate the estimation process used to evaluate models, in the next section we elaborate on Model 1 (suggested operationalising model) and Model 3 (as near as we could get to reproducing the Model 3, Donor-constrained).

11.1 Estimation results for operational model and donor-constrained model

Model 1 (Table 27) is the preferred operationalisation model. It retains a manageable number of parameters, while allowing for mode-specific sensitivities to travel time and cost. Rail and Bus options are separated, reflecting their distinct roles in Perth's transport system, with Rail operating as a higher-speed option for longer-distance travel. Transfers are estimated explicitly and carry strong penalties, consistent with observed behaviour. Socio-demographic variables such as age, household type, driver licence holding, vehicle availability, and gender add explanatory power and capture heterogeneity in traveller responses. The inclusive values for the nests fall within acceptable ranges, supporting behavioural realism. Overall, this specification balances parsimony and accuracy and is practical for embedding in the strategic transport model.

Table 27: Model 1 simplified estimation model with new structure

Parameter	**Estimate**	**Rob. t-ratio**	***p*-value**
Alternative Specific Constants			
Bus	6.583***	8.032	0.000
Rail	5.929***	7.168	0.000
Car	–	–	–
HOV	3.630***	5.647	0.000
Taxi	2.069**	2.751	0.006
Bike	1.455	1.818	0.069
Walk	7.172***	8.962	0.000
PnR	3.209***	4.697	0.000
Mode Travel Time Components			
In-vehicle time (PT)	-0.052***	-9.577	0.000
Access/egress time (PT)	-0.054***	-9.068	0.000
Wait time (PT)	-0.049***	-3.264	0.001
Transfers (PT)	-0.606***	-3.815	0.000
In-vehicle time (Car)	-0.061***	-8.553	0.000
In-vehicle time (HOV)	-0.089***	-7.983	0.000
Car travel time (PnR)	0.009	1.032	0.302
Transfers (PnR)	-1.461***	-4.637	0.000
Travel time (Bike)	-0.051***	-10.220	0.000
Travel time (Walk)	-0.191***	-12.926	0.000
Mode Travel Cost Components			
Parking cost (Car)	-0.135***	-17.010	0.000
Travel cost (Car, PT, PnR)	-0.331***	-9.575	0.000
Travel cost (Taxi)	-0.091***	-4.234	0.000
Socio-Demographics			
Vehicles per household (PT)	-0.611***	-5.534	0.000
Single person household (PT)	-0.634***	-3.153	0.002
Aged 17–30 (PT)	0.578***	3.802	0.000

Parameter	**Estimate**	**Rob. t-ratio**	***p*-value**
Detached dwelling (Bus)	-0.412**	-3.087	0.002
Vehicles per household (Car)	0.745***	5.966	0.000
No licence (Car)	2.774***	7.729	0.000
Female (Bike)	-1.382***	4.885	0.000
Inclusive Values (Null Hypothesis = 1)			
λ (PT modes)	0.352***	8.731	0.000
λ (Active modes)	0.871	0.866	0.375
λ (Car)	0.966	0.411	0.684

Model Fit Statistics: Log Likelihood: -2,823.88; N = 5,564; Nind = 4,726; K = 29; McFadden Pseudo-R^2 = 0.302; AIC = 5,707.75; BIC = 5,906.48

Note: Significance levels: ***p < 0.001, **p < 0.01, *p < 0.05.

In contrast, Model 3 (Donor-constrained) is highly constrained, with a limited set of socio-demographic effects, as shown in Table 28. While this specification offers parsimony, it comes at the expense of explanatory power and theoretical validity. In particular, the implausibly large inclusive value for the car nest (λ=20.492) highlights a structural weakness that undermines the behavioural realism of the model. The restricted treatment of travel times and costs also underestimates disutility, leading to utilities that are less sensitive to policy-relevant variables such as in-vehicle time, access time, and cost. Additional testing found that the cost parameter is not sensitive to income. This seems reasonable for Perth, with a high level of private vehicle ownership across all income classes. In addition, the Model 3 (Donor-constrained) appears to overlook the fact that younger travellers are more inclined to use public transport, single occupant households are more likely to use their car, and males are more likely to cycle. Together, these limitations cast doubt on whether Model 3 (Donor-constrained) would be suitable for reliable prediction or policy analysis.

Table 28: Model 3 Donor structure with parameter constraints

Parameter	**Estimate**	**Rob. t-ratio**	***p*-value**
Alternative Specific Constants			
Bus	0.853*	2.225	0.026
Rail	0.514	1.314	0.189
Car	–	–	–
HOV	-56.461**	-2.296	0.022
Taxi	-2.613***	-4.150	0.000
Bike	-0.876*	-2.077	0.038
Walk	0.288	0.756	0.450
Mode Travel Time Components			
In-vehicle time (PT)	-0.025 (fixed)	–	–
Wait/Walk time (PT)	-0.075 (fixed)	–	–
Transfers (PT)	-0.250 (fixed)	–	–
Car travel time (PnR)	-0.001*	-2.426	0.015
In-vehicle time (Car)	-0.025 (fixed)	–	–
In-vehicle time (HOV)	-0.583***	-4.149	0.000
Travel time (Bike)	-0.063 (fixed)	–	–
Travel time (Walk)	-0.063 (fixed)	–	–
Mode Travel Cost Components			
Parking cost (Car)	-0.127***	-22.488	0.000
Travel cost (Q1 income)	-0.320 (fixed)	–	–
Travel cost (Q2 income)	-0.129 (fixed)	–	–
Travel cost (Q3 income)	-0.075 (fixed)	–	–
Travel cost (Q4 income)	-0.037 (fixed)	–	–
Socio-Demographics			
Vehicles per household (PT)	-0.611***	-5.534	0.000
Vehicles per household (Taxi)	-0.261	-1.148	0.251
Vehicles per household (Bike)	-0.419**	-2.471	0.013
Inclusive Values (Null Hypothesis = 1)			
λ (PT modes)	0.748***	8.726	0.000
λ (Active modes)	0.356***	6.543	0.000
λ (Car)	20.492*	2.511	0.012
λ (Taxi)	1.000	fixed	–

Model Fit Statistics:
Log Likelihood: -3,129.82; N = 5,564; N_{ind} = 4,726; Number of parameters (K): 16 McFadden Pseudo-R^2 = 0.226, AIC: 6,397 (AIC/N = 1.33)

Note: Significance levels: ***p < 0.001, **p < 0.01, *p < 0.05.

11.1.1 Comparison of Model 1 and Model 3

A short comparison between the properties of Model 1 (as presented in Table 27) and Model 3 (as presented in Table 28) is presented in Table 29 below, to show the differences ‘at a glance’.

Table 29: Comparison of Model 1 and Model 3

Category	**Model 1 (Table 27)**	**Model 3 (Table 28)**
ASCs	Separate constants for Bus, Rail, Car, Bike, and Walk, with Car as the base.	Smaller constants, some modes constrained or fixed. Disproportionately low ASC for HOV.
Travel Time Components	Mode-specific treatment of in-vehicle, access/egress, wait, and transfers for Bus and Rail. Separate times for Bike, Walk, Car, Taxi, and PnR.	Shared or fixed time parameters across PT and Active.
Cost Components	Single cost parameter across Car, Bus, Rail, and PnR. Separate estimates for Taxi and parking costs.	Costs segmented by income groups (Q1–Q4). We kept parking as it may approximate land use densities.
Socio-Demographics	Includes household structure, age groups, gender, licence, vehicle ownership,.	Limited to vehicle ownership for PT, Taxi, and Bike.
Inclusive Values (λ)	λ for PT, Active, and Car all estimated below 1.	λ values vary; PT and Active below 1, Car nest above 1.
Model Fit	Higher pseudo R^2, lower AIC.	Lower pseudo R^2, higher AIC.

In the next section, we present the summary specifications for other tours. We have worked through similar investigations to arrive at these models, however, we have chosen to show only the parameter estimates, which would represent the first step to implementing these models into the DTMI’s STEM or into the Perth Transport Model (PTM).

11.2 Parameter matrix for operationalisation

The parameter estimates show substantive and significant differences from the Donor models. The strongest contrasts are the higher sensitivities to travel time, the use of tour-specific cost parameters, and the richer socio-demographic attributes. Each model represents a stronger fit to the PATHS data. Together, they are expected to provide more robust economic evaluations of transport projects in Perth.

Table 30 highlights several important features of the estimating models. The alternative specific constants (ASCs) differ across purposes, showing that the base attractiveness of modes varies by trip type. Bus and Rail are strongly positive for HBW and HBO tours, but weaker for HBS, reflecting lower preference for public transport when shopping. Whilst walking and cycling have positive ASC, the actual market share of cycling remains very low (ASC are location parameters and not market shares). Taxi shows a large variation, becoming unattractive for some tours, while HOV is strong for HBW and HBO, but much weaker in HBU.

The time components confirm that transfers are the most significant deterrent to using public transport. Penalties for transfers in PT range from –0.6 to –0.9, which is much higher than the in-vehicle time coefficients. This is consistent with international evidence that travellers weigh reliability and convenience heavily when making mode choices. Bike and Walk travel times also show high disutility, particularly for longer tours, which reflects practical limits on active modes in Perth. We feel there is some degree of latitude to fix time parameters across tour modes for HBW and HBO in one group and HBU, HBS and SPS in the other group. Whilst not necessary, having these consistent parameter values may simplify the operationalisation of STEM.

The cost components show significant and consistent effects across modes. Parking costs for cars are important, and the strong effect of the travel cost parameter across cars, PT, and PnR reflects the impact of monetary outlays on mode choice. Importantly, there is no evidence that cost sensitivity declines with income, which is different from the Donor model assumptions.

Socio-demographic effects provide further behavioural insight. Younger adults are more likely to use PT, while car licence holding and vehicle ownership predict car use. Females are less likely to cycle, and single-person households are more likely to rely on the car. Detached dwellings are associated with lower bus use. These effects vary across tour purposes and highlight the heterogeneity captured in the STEM models.

Finally, the inclusive values (λ) reveal how substitution between alternatives differs by purpose. Bus and rail share substitution patterns in HBW and HBO, but the correlation is weaker in HBS. Walk and Bike are strongly related in HBW and HBO, but much less so in HBU. We found no compelling evidence to create a nest for private vehicle (Car, PnR, HOV and Taxi), thus we fixed the inclusive value to 1, essentially meaning that these alternatives are in their own single alternative branch.

Car availability and driver licence holding strongly reinforce car use, while simultaneously discouraging public transport, with the effect particularly pronounced in non-work (HBU, SPS) segments. Household structure and dwelling type matter for bus uptake, with detached dwellings significantly less likely to use buses, reflecting lower service access in suburban areas. Younger travellers (17–30 years old) show a stronger inclination toward public transport, especially for shopping and social tours, while gender differences emerge in active modes: women are less likely to cycle but more likely to walk, particularly in discretionary travel. Interestingly, taxi use is suppressed by car ownership in commuting but becomes more attractive for discretionary purposes in car-rich households.

Table 30: Estimated models for five trip purposes (HBW, HBO, HBS, HBU, and SPS)

Parameter Group	**HBW**	**HBO**	**HBS**	**HBU**	**SPS (only tour segments that are servicing a passenger)**
Alternative Specific Constants (ASCs)					
Bus	6.583	5.321	3.386	5.666	2.197
Rail	5.929	3.876	2.676	4.702	1.899
Car	– (base)	– (base)	– (base)	– (base)	NA
HOV	3.630	2.114	3.401	-0.100	– (base)
Taxi	2.069	1.754	-1.357	-0.447	-0.265
Bike	1.455	0.987	-0.249	4.068	4.311
Walk	7.172	6.215	5.092	6.764	5.812
PnR	3.209	2.854	-2.644	4.095	0.852
Time Components					
In-vehicle time (PT)	–0.052	–0.049	-0.015	-0.016	-0.015
Access/egress time (PT)	–0.054	–0.061	-0.031	-0.022	-0.020
Wait time (PT)	–0.049	–0.052	-0.015	-0.016	-0.020
Transfers (PT)	–0.606	–0.711	-0.919	-0.698	-0.766
In-vehicle time (Car)	–0.061	–0.058	-0.017	-0.025	NA
In-vehicle time (HOV, Taxi)	–0.089	–0.077	-0.017	-0.101	-0.134
Car travel time (PnR)	0.009	0.004	0.000	0.000	0.000
Transfers (PnR)	–1.461	–1.238	-0.346	0.000	0.000
Travel time (Bike)	–0.051	–0.048	-0.063	-0.154	-0.224
Travel time (Walk)	–0.191	–0.183	-0.186	-0.110	-0.130
Cost Components					
Parking cost (Car)	–0.135	–0.128	-0.352	-0.065	-0.120
Travel cost (Car, PT, PnR)	–0.331	–0.305	-0.152	-0.614	-0.397
Travel cost (Taxi)	–0.091	–0.086	-0.071	-0.204	-0.200
Socio-Demographics					
Vehicles per HH (PT)	–0.611	–0.588			
Vehicles per HH (Taxi)	–0.261	–0.248			
Single person HH (Bus)	–0.634	–0.612			
Aged 17–30 (PT)	0.578	0.551	1.340		
Detached dwelling (Bus)	–0.412	–0.386		-0.980	
Vehicles per HH (Car)	0.745	0.718	0.588	1.387	1.780
licence (Car)	2.774	2.639	3.488	1.821	
Female (Bike)	-1.382	-1.297	-1.164	-1.394	
Female (Walk)			0.579		
Inclusive Values (λ)					
λ (Bus, Rail)	0.352	0.365	1.000	0.524	0.462
λ (Walk, Bike)	0.871	0.856	0.842	0.371	0.398
Car, HOV, Taxi, PnR	1.000	1.000	1.000	1.000	1.000

12 Conclusion and Limitations

Using the Perth Area Travel and Household Survey (PATHS) and associated stated preference (SP) component, this report examined the usefulness of these data to estimate a tour-based mode choice model by integrating Revealed Preference (RP) and Stated Preference (SP) into choice modelling of tour data. The analysis highlighted several key insights and challenges in understanding and modelling mode choice behaviour for tours, particularly within the context of Perth's transport network.

The report underscores that most tours in the dataset are simple in structure, often involving a single activity at the destination, such as home-work-home or home-recreation-home. They are also predominantly unimodal. Even in cases where tours became slightly more complex, such as adding a passenger drop-off before continuing to a primary activity, a single mode generally dominated the tours. This prompted the team to recommend the Simplified Main Tour Mode approach. Whilst this strategy aligns with the data's characteristics, it sacrifices some behavioural detail. Despite this finding, the team examined the data for multi-modal combinations (Section 5.2.3) and found some for combinations of Car and Walk and Car and Bus. These tours typically belonged to younger household members (<17 years) and SCH tours.

Tour complexity is relatively rare. While tour-based modelling was an initial goal, only 33% of tours were complex (consisting of more than one activity), only 5.5% involved more than one transport mode and fewer than 2% contained both. The analysis reveals that most tours in the dataset were simple, typically involving a single activity at the destination (e.g., home-work-home, home-recreation-home) and relying on a single mode of transport. Even when complexity increases, such as including a passenger drop-off before proceeding to the primary activity, a dominant mode generally persists, which led to the use of the Simplified Main Tour Mode approach, which assigns a primary mode for the entire tour.

Overall, the WTP results highlight the importance of PT service accessibility (access/egress times and service frequency over in-vehicle travel time). This explains the dominance of car use for trips with complex or indirect public transport access, and the strong PT mode share to centres well-connected by PT such as the CBD. Also, parking costs are a major lever for behavioural change. When parking costs were separated from vehicle running costs in model estimation, the magnitude of the coefficient was much larger. This suggests that increases to parking prices, especially in constrained activity centres, can meaningfully shift mode choice.

As the sequence of models has shown, joint RP/SP estimation offer insights, but in this case RP models were stronger. The SP component of PATHS contributed useful directional checks, but less robust, partly due to design limitations in the survey instrument. Ultimately, the standalone RP models, particularly those using the first trip of the tour, offered greater reliability and we recommend them for application in strategic transport modelling.

Reliable mode choice models must be grounded in the observed behaviour of Perth’s residents. The unique nature of the city’s car dependence, dispersed urban form, and evolving public transport network means that generic models or those made for other cities often fail to capture local differences and considerations, which the models in this report aimed to address.

The revealed preference PATHS datasets, providing observed data from a variety of Perth residents, enriched with secondary data (such as network attributes, land-use indicators and service availability measures), provided a robust foundation for model development. While the integration of secondary sources such as OpenTripPlanner and STEM travel time and distance matrices strengthened the credibility of key travel time and cost estimates, more could be done to further improve these measures when imputed for non-chosen alternatives for RP modelling.

Consequently, the research team recommends for any future research developing specialised routing tools to enhance the accuracy of alternate travel time metrics for the RP non-chosen alternatives.

The models presented in this report meet the primary objective of developing robust, tour-based mode choice models using the PATHS datasets. These models are designed to support the transport agency in benchmarking our results against existing models currently used in Perth's strategic transport planning. A secondary aim has been to evaluate the value of the stated choice data in improving the estimation of time and cost parameters, which are critical inputs for the economic appraisal of transport policy and infrastructure investments. While the results offer a strong foundation, we acknowledge several limitations and outline recommendations for extending future research and enhancing the utility of the PATHS data.

We recognise there were a number of important variables that were not included in the model and may have limited its ability to fully capture the complexity of travel behaviour. Household income, for example, was omitted, despite its potential to explain heterogeneity in willingness-to-pay for time and cost savings. The time of day a tour commenced was also excluded, even though morning departures are often associated with more complex tours, and with declining availability of PnR bays, both of which are likely to influence mode choice. The model did not account for household interactions and joint travel, which were evident in the dataset, particularly in school and serving passenger tours, where multiple household members travelled together. Finally, the tour complexity analysis did not segment by the tour type, which may have led to some further insights into the way household structure and person demographics shape daily activity patterns.

Similarly, there is a need for more detailed spatial and network measures to improve model accuracy, and to better reflect the decision-making context faced by travellers. The current models do not incorporate comprehensive spatial attributes, such as land-use mix, walkability, or real-time network performance, all of which are likely to influence travel behaviour. Including these variables would provide a more complete representation of the urban environment and offer greater explanatory power.

A further comparison between PATHS-based RP models and donor-constrained models suggests caution in transferring parameter estimates from other contexts. The donor models imposed structural assumptions such as grouping bus and rail together and placing Park-and-Ride within the public transport nest, which do not align with observed Perth behaviour. More critically, the donor time coefficients were small, implying weak responsiveness to travel time savings and underestimating the benefits of infrastructure or service improvements. In contrast, the PATHS RP models provided stronger and more differentiated sensitivities, with higher values of time and realistic transfer penalties. For Perth, this indicates that donor models may not provide a sound foundation for policy evaluation, whereas PATHS-informed models offer greater behavioural realism and reliability.

Finally, we acknowledge that operationalising the choice models developed in this study will require further collaboration with DTMI, particularly to ensure alignment with both the existing STEM and upcoming PTM frameworks. Not all tour purposes were included in the current model, and a broader set of purposes, including SCH, should be incorporated to better reflect the diversity of travel behaviour. The elemental modes used in this analysis do not fully align with those adopted by the agency, and without adjustment, this misalignment may limit the direct transferability of results. To embed the models into forecasting tools, coordination is needed to confirm mode definitions and tour classifications, and to identify where simplified, but robust, model structures may be necessary to meet operational needs.

References

Arentze, T. A., Hofman, F., Joh, C. H., & Timmermans, H. J. P. (1999). The development of ALBATROSS: Some key issues. In *Traffic and Mobility: Simulation—Economics—Environment* (pp. 57-72). Springer Berlin Heidelberg.

Auld, J., & Mohammadian, A. (2009). Framework for the development of the agent-based dynamic activity planning and travel scheduling (ADAPTS) model. *Transportation Letters*, *1*(3), 245–255. https://doi.org/10.3328/TL.2009.01.03.245-255.

Axhausen, K. W., & Gärling, T. (1992). Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems. *Transport Reviews*, *12*(4), 323-341.

Azevedo, Carlos Lima, Neeraj Milind Deshmukh, Balakumar Marimuthu, Simon Oh, Katarzyna Marczuk, Harold Soh, Kakali Basak, Tomer Toledo, Li-Shiuan Peh, and Moshe E. Ben-Akiva. "Simmobility short-term: An integrated microscopic mobility simulator." *Transportation Research Record* 2622, no. 1 (2017), 13-23.

Bastarianto, F. F., Hancock, T. O., Choudhury, C. F., & Manley, E. (2023). Agent-based models in urban transportation: review, challenges, and opportunities. In *European Transport Research Review* (Vol. 15, Issue 1). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1186/s12544-023-00590-5.

Bowman, J. L., & Ben-Akiva, M. E. (2001). Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research A*, *35*(1), 1-28.

Bradley, M., Bowman, J. L., & Griesenbeck, B. (2010). SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution. *Journal of Choice Modelling*, *3*(1), 5-31.

Buliung, R. N., & Kanaroglou, P. S. (2007). Activity–travel behaviour research: conceptual issues, state of the art, and emerging perspectives on behavioural analysis and simulation modelling. *Transport Reviews*, 27(2), 151-187.

City of Perth (2021). Community Profile, https://profile.id.com.au/perth/car-ownership#:~:text=Analysis%20of%20the%20car%20ownership,19.2%25%20respectively%20for%20Greater%20Perth.

Davidson, W., Vovsha, P., Freedman, J., & Donnelly, R. (2010, September). CT-RAMP family of activity-based models. In *Proceedings of the 33rd Australasian Transport Research Forum (ATRF)* (Vol. 29, p. 29).

Gärling, T., Kwan, M. P., & Golledge, R. G. (1994). Computational-process modelling of household activity scheduling. *Transportation Research B*, 28(5), 355-364.

Hägerstrand, T. (1970). What about people in regional science, *Regional Science Association Papers*, 29.

Hasnine, M. S., & Habib, K. N. (2018). What about the dynamics in daily travel mode choices? A dynamic discrete choice approach for tour-based mode choice modelling. *Transport Policy*, *71*, 70–80. https://doi.org/10.1016/j.tranpol.2018.07.011.

Hasnine, M. S., & Habib, K. (2020). Tour-based mode choice modelling as the core of an activity-based travel demand modelling framework: a review of state-of-the-art. *Transport Reviews*, 41(1), 5–26, https://doi.org/10.1080/01441647.2020.1780648.

Hesam Hafezi, M., Sultana Daisy, N., Millward, H., & Liu, L. (2022). Framework for development of the Scheduler for Activities, Locations, and Travel (SALT) model. *Transportmetrica A: Transport Science*, *18*(2), 248-280.

Ho, C., & Mulley, C. (2013). Tour-based mode choice of joint household travel patterns on weekend and weekday. *Transportation*, *40*(4), 789–811, https://doi.org/10.1007/s11116-013-9479-0.

IPSOS (2019). *PATHS Stated Preference Methodology & Experimental Design*, Contract 228/15 Provision of Household Travel and Stated Preference Surveys, Ipsos Reference #17-009257-01.

Jones, P. M. (1979). ‘HATS’: a technique for investigating household decisions. *Environment and Planning A*, 11(1), 59-70.

Jones, P. M., Dix, M. C., Clarke, M. I., & Heggie, I. G. (1983). *Understanding travel behaviour*, Gower Publishing, Brookfield, VT, United States.

Kitamura, R., & Fujii, S. (1998). Two computational process models of activity-travel behavior. *Theoretical foundations of travel choice modeling*, 251-279.

Kagho, G. O., Balac, M., & Axhausen, K. W. (2020). Agent-Based Models in Transport Planning: Current State, Issues, Expectations. *Procedia Computer Science*, *170*, 726–732. https://doi.org/10.1016/j.procs.2020.03.164.

Louviere, J.J., Hensher, D.A., & Swait, J.D. (2000). *Stated Choice Methods: Analysis and Applications*. Cambridge University Press.

Mai, T., Fosgerau, M., & Frejinger, E. (2015). A nested recursive logit model for route choice analysis. *Transportation Research B*, *75*, 100-112.

Miller, E. J., Roorda, M. J. & Carrasco, J. A. (2005). A tour-based model of travel mode choice. *Transportation*, 32, 399–422, https://doi.org/10.1007/s11116-004-7962-3.

Pendyala, R. M., Kitamura, R., Kikuchi, A., Yamamoto, T., & Fujii, S. (2005). Florida Activity Mobility Simulator: Overview and Preliminary Validation Results. *Transportation Research Record*, *1921*(1), 123–130, https://doi.org/10.1177/0361198105192100114.

Rai, R. K., Balmer, M., Rieser, M., Vaze, V. S., Schönfelder, S., & Axhausen, K. W. (2007). Capturing human activity spaces: New geometries. *Transportation Research Record*, 2021(1), 70-80.

Rasouli, S., & Timmermans, H. (2013). Activity-based models of travel demand: promises, progress and prospects. *International Journal of Urban Sciences*, 18(1), 31–60. https://doi.org/10.1080/12265934.2013.835118.

Recker, W. W., McNally, M. G., & Root, G. S. (1986a). A model of complex travel behavior: Part I—Theoretical development. *Transportation Research Part A: General*, 20(4), 307-318.

Recker, W. W., McNally, M. G., & Root, G. S. (1986b). A model of complex travel behavior: Part II—An operational model. *Transportation Research Part A: General*, 20(4), 319-330.

Tajaddini, A., Rose, G., Kockelman, K. M., & Vu, H. L. (2020). Recent progress in activity-based travel demand modeling: rising data and applicability. In: de Luca, S., Di Pace, R., & Fiori, C. (Eds.) *Models and Technologies for Smart, Sustainable and Safe Transportation Systems*, Chp.2, https://doi.org/10.5772/intechopen.87681, https://www.intechopen.com/chapters/73240.

Vovsha, P., Hicks, J. E., Vyas, G., Livshits, V., Jeon, K., Anderson, R., & Giaimo, G. (2017). Combinatorial tour mode choice. Presented at the *96th Annual Meeting of the Transportation Research Board*, January 8–12.

13 Appendix A: RP Mode Choice (First Trip of the Tour) by Tour Purpose

13.1 Home-based work: first trip

The HBW tour model includes seven elemental alternatives—Bus, Rail, Car, Taxi, Bike, Walk, and Park-and-Ride (PnR)—organised under a three-level nested logit structure: PT (Bus, Rail), Active (Bike, Walk), and Car-based modes (Car, Taxi, PnR). The model is estimated on 5,569 observations from 4,733 individuals. Model fit statistics indicate strong explanatory power with a McFadden's pseudo-R^2 of 0.30 against observed shares, suggesting good model performance for a tour-based RP context.

The travel time coefficients are all negative and statistically significant, reinforcing the expected disutility associated with time spent traveling. The in-vehicle time parameters for bus (–0.044) and rail (–0.035) are slightly more sensitive than in the full-sample model, while the penalty for car travel time (–0.060) is notably higher, indicating greater responsiveness among HBW travellers to car time. Access/egress times are particularly penalising for both bus (–0.080) and rail (–0.040), as are transfer penalties (–0.888 for bus, –0.676 for rail), consistent with the structured nature of commute trips. Walking time shows the largest negative effect (–0.183), suggesting HBW users are especially averse to long walks as part of their commute. Interestingly, the taxi time coefficient (–0.04) is also high, despite taxi being infrequently chosen.

The estimated cost parameters in the HBW model are negative and statistically significant (-0.135 for car and PT) but not for taxi (-0.023 for taxi), aligning with expectations that higher monetary outlays reduce mode utility and supporting the derivation of willingness-to-pay (WTP) metrics for travel time savings across modes.

Table A 1: RP Only Trip Choice Models for Home-Based Work

Parameter	**Estimate**	**Rob. t-ratio**	***p*-value**
Alternate Specific Constants			
Bus	4.896***	7.34	0.000
Rail	3.665***	5.11	0.000
Car	-0.752	1.09	0.278
Taxi	-	-	-
Bike	1.169**	2.29	0.022
Walk	5.927***	10.27	0.000
Mode Travel Time Components			
In-vehicle time (Bus)	-0.043***	6.31	0.000
Access and egress time (Bus)	-0.080***	5.70	0.000
Half headway time (Bus)	-0.051***	3.31	0.000
Number of transfers (Bus)	-0.888***	4.32	0.000
In-vehicle time (Rail)	-0.035***	5.23	0.000
Access and egress time (Rail)	-0.040***	3.17	0.000
Half headway time (Rail)	-0.069***	4.41	0.000
Number of transfers (Rail)	-0.656***	5.50	0.000
Travel time (Bike)	-0.045***	10.38	0.000
Travel time (Walk)	-0.183***	6.31	0.000
In-vehicle time (Car)	-0.072***	5.70	0.000

Parameter	**Estimate**	**Rob. t-ratio**	***p*-value**
Car travel time (PnR)	-0.001***	3.31	0.000
Taxi travel time (Taxi)	-0.107*	9.32	0.053
Mode Travel Cost Components			
Travel cost (Car, Bus Rail, PnR)	-0.135***	17.7	0.000
Travel cost (Taxi)	-0.025	0.61	0.543
Socio-Demographics			
Employment (Bus)	-0.039**	0.51	0.609
Age between 17-30 (Bus)	0.545***	2.38	0.018
Age between 30-55 (Bus)	-0.417***	0.57	0.567
Single Person HH (Bus)	-0.771***	2.88	0.004
Detached Dwelling (Bus)	-0.470***	2.48	0.013
Age between 17-30 (Rail)	0508**	1.99	0.047
Age between 30-55 (Rail)	-0.098	0.43	0.665
Single person household (Rail)	-0.774***	3.33	0.001
Female (Bike)	-1.398***	4.76	0.000
Age over 75 (Walk)	-1.365***	1.47	0.141
Vehicle ownership per household (Car)	0.720***	5.21	0.000
Has a driver licence (Car)	2.410***	8.47	0.000
Self-Reported Frequency for Taxi and Ride Share			
Rideshare frequency (Taxi)	1.836***	3.81	0.000
Rideshare frequency (Bus)	0.451**	2.42	0.015
Rideshare frequency (Car)	-0.585***	2.68	0.007
Inclusive Values (Null Hypothesis = 1)			
λ (PT modes)	0.416***	5.56	0.000
λ (Active modes)	1	fixed	-
λ (Car)	0.907**	3.10	0.002

Model Fit Statistics:
Log Likelihood: -2,774.19; N = 5569; N_{ind} = 4733; Number of parameters (K): 38
McFadden Pseudo-R^2 = 0.300, AIC: 4,060.16 (AIC/N = 0.73)
Note: Significance. Levels: ***$p < 0.001$, ***$p < 0.01$*, *$p < 0.05$*.

Table A2: Willingness to pay for all first trips of the tour HBW

Mode	**In-Vehicle Time ($/hour)**	**Access/Egress Time ($/hour)**	**Headway / Wait Time ($/hour)**	**Transfer Penalty ($)**
Bus	$19.08	$35.60	$22.46	$6.55
Rail	$15.58	$17.78	$30.52	$4.99
Car	$31.79	–	–	–
Taxi	Insufficient information	–		

13.2 Home-Based and Non-Home-Based Other (HBO)

The RP nested logit model for the Home Based and non-Home Based Other (HBO) segment captures mode choice behaviour across a range of discretionary and personal trip purposes, including recreation, visiting friends, entertainment and personal business (Table A2). The seven elemental alternatives are nested within public transport (bus and rail) and private vehicle (car and taxi) with active modes (bike and walk) each belonging to a degenerate branch. With 4,816 observations from 3,723 individuals, compared with observed shares, the model captures a lower McFadden pseudo-R^2 of 0.17, consistent with the high variability and lower predictability of non-commuting travel behaviour.

Travel time parameters across modes are generally negative, as expected, with most statistically significant. Bus and rail in-vehicle times (–0.027 and –0.028, respectively) are modestly valued, while access/egress and wait times show more variation. Notably, the coefficient on rail access time is small and statistically insignificant, suggesting limited sensitivity to first/last-mile distance in this tour type. The transfer penalty for bus users is large and significant (–1.044), indicating a strong disutility from multi-leg PT journeys, though rail transfers are not statistically significant. For private modes, car travel time (–0.047) is significant and larger in magnitude than for PT modes, suggesting discretionary travellers value personal convenience and time efficiency. Active travel modes also show strong time sensitivity, with the walk time coefficient at –0.12 and the bike coefficient at –0.036.

The estimated willingness-to-pay (WTP) values for the HBO (Home-Based Other) segment reveal clear patterns in how travellers value reductions in travel time and service characteristics across modes. In-vehicle time for both bus and rail modes is valued similarly, with WTPs of \$14.20 and \$14.83 per hour respectively, indicating comparable disutility during the main segment of public transport trips. The disutility of bus transfers is notably high, with a WTP of \$9.28, underscoring the potential traveller frustration associated with indirect or complex bus routes. In contrast, rail transfers attract a much lower penalty (\$1.92), suggesting greater tolerance for transfers within the rail system, possibly due to more reliable infrastructure or station environments.

Car users show a higher value of time (\$25.00 per hour), reflecting either income effects or a stronger preference for uninterrupted travel. Taxi users exhibit the highest WTP to reduce travel time at \$51.47 per hour, though this result may be affected by a small sample or short trip durations, warranting caution in interpretation. These differences in marginal values offer a useful basis for prioritising improvements in directness and travel time savings across transport modes, particularly in reducing transfer penalties and in-vehicle time for public transport.

Table A3: Willingness to pay for all first trips of the tour HBW

Mode	**In-Vehicle Time (\$/hour)**	**Access/Egress Time (\$/hour)**	**Headway / Wait Time (\$/hour)**	**Transfer Penalty (\$)**
Bus	\$14.20	\$42.13	\$8.19	\$9.28
Rail	\$14.83	\$2.30	\$45.06	\$1.92
Car	\$31.79	–	–	–
Taxi	\$51.47	–		

Table A 2: RP Only Trip Choice Models for Home-Based and Non-Home-Based Other

Parameter	**Estimate**	**Rob. t-ratio**	***p*-value**
Alternate Specific Constants			
Bus	1.335	1.66	0.098
Rail	0.756	0.72	0.474
Car	0.668	1.77	0.077
Bike	-0.456	1.25	0.211
Walk	4.341***	12.22	0.000
Park and Ride (PnR)	-0.486	1.38	0.167
Mode Travel Time Components			
In-vehicle time (Bus)	-0.027*	1.78	0.075
Access and egress time (Bus)	-0.079***	3.21	0.001
Half headway time (Bus)	-0.015	0.76	0.448
Number of transfers (Bus)	-1.044*	2.17	0.030
In-vehicle time (Rail)	-0.028**	3.08	0.002
Access and egress time (Rail)	-0.004	0.34	0.735
Half headway time (Rail)	-0.084	1.80	0.071
Number of transfers (Rail)	-0.216	0.86	0.392
Travel time (Bike)	-0.036***	5.94	0.000
Travel time (Walk)	-0.12***	13.63	0.000
In-vehicle time (Car)	-0.047***	6.05	0.000
Car travel time (PnR)	-0.0001	0.40	0.686
Taxi travel time (Taxi)	-0.002	0.08	0.933
Mode Travel Cost Components			
Travel cost (Car, Bus Rail, PnR)	-0.112***	9.33	0.000
Travel cost (Taxi)	-0.055***	3.18	0.001
Socio-Demographics			
Employment (Bus)	-0.277**	2.55	0.011
Age between 17-30 (Bus)	1.456***	4.05	0.000
Age between 30-55 (Bus)	-0.088	0.20	0.841
Single Person HH (Bus)	-0.534	1.49	0.137
Detached Dwelling (Bus)	-1.15**	2.81	0.005
Age between 17-30 (Rail)	1.488***	3.46	0.001
Age between 30-55 (Rail)	-0.276	0.59	0.555
Single person household (Rail)	-1.139**	2.73	0.006
Female (Bike)	-1.541***	5.53	0.000
Age over 75 (Walk)	-0.594***	5.19	0.000
Vehicle ownership per household (Car)	0.364***	3.55	0.000
Has a driver licence (Car)	0.786***	4.08	0.000
Self-Reported Frequency for Taxi and Ride Share			
Rideshare frequency (Taxi)	1.313***	3.67	0.000
Rideshare frequency (Bus)	-0.221	0.43	0.665

Parameter	Estimate	Rob. t-ratio	*p*-value
Rideshare frequency (Car)	-0.313	1.27	0.203
Inclusive Values (Null Hypothesis = 1)			
λ (PT modes)	0.436**	2.47	0.014
λ (Active modes)	1	fixed	-
λ (Car)	0.549***	6.79	0.000

Model Fit Statistics:
Log Likelihood: -2376.43; N = 4816; Nind = 3723; Number of parameters (K):38
McFadden Pseudo-R^2 = 0.171, AIC: 4015.48 (AIC/N = 0.83)
Taxi is the reference category.
Note: Signif. Levels: ***p < 0.001, ***p < 0.01, *p < 0.05.*

13.3 Home-Based Shopping

Table A3 below shows the results of the RP MNL mode choice model for home-based shopping (HBS) trips.

The attributes for rail (travel time and transfers) are not significant, for bus, only travel time is negatively significant, but for AT travel time is a deterrent from their use for shopping. For car only the costs are statistically significant, reflecting the preference for car in shopping trips for Perth residents. Notably, while females are less likely to cycle for commuting and other purposes, they are more likely to cycle for shopping trips.

Table A 3: Multinomial Logit Model – HBS Trips

Parameter	Estimate	Rob. t-ratio	*p*-value
Alternate Specific Constants			
Bus	-3.244***	-4.03	0.000
Rail	-3.700***	-4.47	0.000
Car	-5.509***	-8.50	0.000
Bike	-4.074***	-7.03	0.000
Walk	-	-	-
Mode Travel Time Components			
Travel time (Bus)	-0.029*	-2.36	0.018
Wait time (Bus)	0.003	0.19	0.849
Number of transfers (Bus)	-0.540	-1.32	0.187
Travel time (Rail)	-0.009	-0.57	0.569
Wait time (Rail)	-0.171*	-2.27	0.023
Number of transfers (Rail)	-0.295	-0.54	0.589
Long distance travel (Rail)	1.629**	2.62	0.009
Travel time (Car)	-0.004	-0.13	0.897
Travel time (Bike)	-0.072*	-2.14	0.033
Travel time (Walk)	-0.173***	-8.64	0.000
Short distance (Walk)	1.056**	2.59	0.010
Mode Travel Cost Components			
Parking cost (Car)	-0.278*	-2.17	0.030
Travel cost	-0.131	-1.84	0.066
Employment (Bus)	0.417	1.20	0.230

Parameter	**Estimate**	**Rob. t-ratio**	***p*-value**
Detached Dwelling (Car)	1.011***	5.33	0.000
Vehicle ownership per household (Car)	1.286***	4.02	0.000
has a driver licence (Car)	2.254***	7.32	0.000
Female (Bike)	-1.234**	2.64	0.008

Model Fit Statistics:
Log Likelihood: -707.79; N = 3,819; N_{ind} = 3,232; Number of parameters (K): 22
McFadden Pseudo-R^2 = 0.246, AIC: 1,459.58 (AIC/N = 0.38)
Note: Signif. Levels: ***p < 0.001, ***p < 0.01, *p < 0.05.*

14 Appendix B: Preliminary Choice Modelling Investigation on SP Data

The purpose of this appendix is to detail a preliminary analysis on the usefulness of the data as received. We noted that the data contained unrealistic travel times (very large or very short), Here we eliminated choice sets that contained unrealistic travel times leading to a 22% of the SP unusable datapoints. The final number of choice tasks used for the analysis is 38,605 from the original 49,643.

This appendix presents the MNL and the Latent 2-class choice (LCM) models with both the received data and the reduced data set. We present this as supporting evidence to remove the choice tasks with unrealistic travel times.

14.1 Multinomial Logit Choice SP Model

A simple MNL model has been estimated considering all SP data (Table B1), with utility functions according to the travel options after eliminating the very extreme outliers. The models do not include socio-demographics, nor RP information.

For example, the active travel options include only walking/cycling time; the taxis /shared only the in-vehicle time, cost, and if they are shared or not; PT the access mode and time, in-vehicle time, transfers and waiting time, fare, parking availability for PnR; for car, in-vehicle time, costs for running and parking. In relations 1) to 6) below, the names of the variables are as in the SP dataset.

u(CarD)=Acar+bTTcar*TIMEINVE+bparkCar*PARKAVAI+bopcost*OPCOSTSF+bparkc* PARKCOST (1)

u(PnR)=APnR+bTTPnR*TIMEINVE+bparkPnR*PARKAVAI+bTTwait*WAITTRAN+bopcost*OPCOSTSF+bPT*MAINPTSE+btrf*NOPTTRAN (2)

u(WnR)=AWnR+bTTWnR*TIMEINVE+bTTwalk*WALKBIKE+bTTwait*WAITTRAN+bopcost*OPCOSTSF+bPT*MAINPTSE+btrf*NOPTTRAN (3)

u(TaxiS)=bTTTaxi*TIMEINVE+bopcost*OPCOSTSF+bshare*SHARED (4)

u(Cycle)=ABike+bTTBike*WALKBIKE (5)

u(Walk)=AWalk+bTTWalk*WALKBIKE (6)

Table B 1: Multinomial Logit (MNL) Results All Data

MODE	CHOICE	Parameter	Std. error	z	Sig. level p
CAR	*ASC*	_1.96210***_	*0.13187*	*14.88*	*<0.001*
	BTTCAR (In-vehicle travel time)	-0.00704***	0.00183	-3.85	0.001
	BPARKCAR (Parking availability)	0.03720**	0.01823	2.04	0.0413
	BOPCOST (Operation costs)	-0.02288***	0.00227	-10.06	<0.001
	BPARKC (Parking costs)	-0.03673***	0.00298	-12.33	<0.001
PnR	*ASC*	*-0.08827*	*0.16929*	*-0.52*	*0.6021*
	BTTPNR (In-vehicle travel time)	-0.00840***	0.00206	-4.09	<0.001
	BPARKPNR (Parking availability)	0.03965	0.03083	1.29	0.1985

MODE	CHOICE	Parameter	Std. error	z	Sig. level p
	BTTWAIT (Waiting & transfer time)	0.02573***	0.00755	3.41	0.0007
	BPT (Bus)	-0.04279*	0.02399	-1.78	0.0745
	BTRF (Number transfers)	-0.0167	0.03009	-0.55	0.5789
Public Transport (AT access)	*ASC*	*0 .59712****	*0.12885*	*4.63*	*<0.001*
	BTTWNR (In-vehicle travel time)	-0.00892***	0.00204	-4.38	<0.001
	BTTWALK (Active travel time)	-0.04320***	0.00153	-28.17	<0.001
TAXI/ Share	BTTTAXI (In-vehicle travel time)	-0.00762***	0.00243	-3.14	0.0017
	BSHARE (Shared ride)	0.04132	0.06399	0.65	0.5185
AT	*ASC Bike*	*0.57846****	*0.12301*	*4.7*	*<0.001*
	BTTBIKE (Travel time)	-0.03570***	0.00199	-17.92	<0.001
	ASC Walk	*1.88315****	*0.12658*	*14.88*	*<0.001*

Note: Signif. Levels: ***p < 0.001, **p < 0.01, *p < 0.05.

Considering the unrealistic travel times (very large or very short), the team eliminated these options from the SP dataset in two ways: a) eliminating completely the choice sets; and b) reducing the choice set sizes by eliminating only the alternatives. The former has led to 38,605 from 49,643, a reduction of over 22% of the SP usable dataset. The results change after removing choice sets with unrealistically small or large travel times, as shown in Table B2.

Table B 2: MNL Results Reduced Number of Choice Sets

MODE	CHOICE	Parameter	Std. error	z	Sig. level p
CAR	*ASC*	*1.81614****	*0.18298*	*9.93*	*<0.001*
	BTTCAR (In-vehicle travel time)	-0.00547***	0.00320	-1.71	0.087
	BPARKCAR (Parking availability)	0.04882**	0.02475	1.97	0.0486
	BOPCOST (Operation costs)	-0.02879***	0.00399	-7.21	<0.001
	BPARKC (Parking costs)	-0.03975***	0.00425	-9.36	<0.001
PnR	*ASC*	*-0.36326*	*0.24229*	*-1.50*	*0.1338*
	BTTPNR (In-vehicle travel time)	-0.00426***	0.00403	-1.06	0.2908
	BPARKPNR (Parking availability)	0.05508	0.04374	1.26	0.2079
	BTTWAIT (Waiting & transfer time)	0.03780***	0.01379	2.74	0.0061
	BPT (Bus)	-0.04536*	0.03441	-1.32	0.1875
	BTRF (Number transfers)	-0.0261	0.04178	-0.62	0.5323
Public Transport (AT access)	*ASC*	*0.4463****	*0.18488*	*2.41*	*0.016*
	BTTWNR (In-vehicle travel time)	-0.00511***	0.00360	-1.42	0.156
	BTTWALK	*-0.06033****	*0.00287*	*-21.02*	*<0.001*

MODE	CHOICE	Parameter	Std. error	z	Sig. level p
	(Active travel time)				
TAXI/ Share	BTTTAXI (In-vehicle travel time)	-0.0043	0.00481	-0.88	0.3762
	BSHARE (Shared ride)	-0.05231	0.0900	-0.58	0.5612
AT	*ASC Bike*	*0.71663****	*0.1764*	*4.06*	*<0.001*
	BTTBIKE (Travel time)	-0.04703***	0.00363	-12.97	<0.001
	ASC Walk	2.19196***	*0.18558*	*11.81*	*<0.001*

Model fit statistics: LL= -6,620.93894; N = 6,461, K = 19; Inf.Cr.AIC = 13279.9 AIC/N = 2.055;
Root Likelihood: Geom. Mean of P^ =0.3589
Chi-sq. [14](LR test) = 7,056.12475, P< 0.001
Note: Signif. Levels: ***p < 0.001, **p < 0.01, *p < 0.05.

The results suggest a stronger preference for car and active travel, followed by public transport accessed by walking, compared to taxi and shared mobility (reference category), but lower for PnR. Travel times have the expected negative coefficients, however, waiting times are positive. The costs are also with the expected negative sign.

14.2 Latent 2-Class Choice Model

Additionally, an LCM was estimated (Table B3) and the results suggest two similar size classes with distinct parameters. The fit improved substantially, but this is only indicative of heterogeneity in the responses (LL=-12,260.961, compared to -22,549.293, McFadden Pseudo R^2=0.456). The choice to limit the number of classes to two was based on the examination of model fit statistics, class membership size and consistency of parameter signs with expectations.

Class 1 includes individuals that are not sensitive to car travel time and costs, but to parking availability. They have a preference for active travel, but longer travel times are a deterrent from using them (as illustrated by the significant and negative coefficients). They are willing to accept long wait and transfer times, although PnR is not favoured. Conversely, Class 2 is sensitive to cost and time (signs in the expected direction), resembling the trade-offs documented in the wide literature.

The class 1 results seem to include those respondents who received choice sets with unrealistically large walking and cycling times and further cleaning of the data is required.

Table B 3: LCM All SP Data

Variable	Class 1 (49.2%)				Class 2 (50.8%)			
	Parameter	Std. error	z	p	Parameter	Std. error	z	p
ASC Car	*2.29623****	*0.57389*	*4.00*	*<0.001*	*1.91789****	*0.54338*	*3.53*	*0.0004*
BTTCAR (In-vehicle travel time)	0.00421	0.00928	0.45	0.650	**-0.01848*****	**0.00580**	**-3.19**	**0.0014**
BPARKCAR (Parking availability)	**0.19766****	**0.07599**	**2.60**	**0.0093**	-0.02055	0.05270	-0.39	0.6966
BOPCOST (Operation costs)	-0.0098	0.00625	-1.56	0.1187	**-0.05471*****	**0.00839**	**-6.52**	**<0.001**
BPARKC (Parking costs)	-0.0144	0.01374	-1.05	0.2942	**-0.07612*****	**0.01060**	**-7.18**	**<0.001**
ASC PnR	*-1.38788***	*0.58308*	*-2.38*	*0.0173*	*0.52433*	*0.55111*	*0.95*	*0.3414*
BTTPNR (In-vehicle travel time)	-0.0185*	0.01052	-1.76	0.0787	**-0.01105****	**0.00467**	**-2.36**	**0.0181**
BPARKPNR (Parking availability)	**0.28357****	**0.12283**	**2.31**	**0.0210**	**-0.12955***	**0.06932**	**-1.87**	**0.0617**
BTTWAIT (Waiting & transfer time)	**0.1181*****	**0.03878**	**3.04**	**0.0023**	**-0.00950**	**0.01415**	**-0.67**	**0.5021**
BPT (Bus)	**0.32976***	**0.18294**	**1.80**	**0.0715**	-0.10746	0.08879	-1.21	0.2261
BTRF (Number transfers)	0.08521	0.18156	0.47	0.6389	-0.08096	0.05348	-1.51	0.1301
ASC PT WnR	*0.6705*	*0.51014*	*1.31*	*0.1887*	*0.93371***	*0.47661*	*1.96*	*0.0501*
BTTWNR (In-vehicle travel time)	0.0092	0.01329	0.07	0.9447	**-0.01468*****	**0.00386**	**-3.80**	**<0.001**
BTTWALK (Active travel time)	**-0.2840*****	**0.03167**	**-8.97**	**<0.001**	**-0.02525*****	**0.00235**	**-10.74**	**<0.001**
BTTTAXI (In-vehicle travel time)	0.00047	0.00835	0.06	0.9550	**-0.02392*****	**0.00757**	**-3.16**	**0.0016**
BSHARE (Shared ride)	0.31957	0.22436	1.42	0.1543	-0.14707	0.25920	-0.57	0.5704
ASC Bike	*3.2584****	*0.56936*	*5.72*	*<0.001*	0.55138	0.47076	*1.17*	*0.2415*
BTTBIKE (Travel time)	**-0.20723*****	**0.02776**	**-7.47**	**<0.001**	**-0.03466*****	**0.00349**	**-9.93**	**<0.001**
ASC Walk	*8.62755****	*0.79083*	*10.91*	*<0.001*	0.71390	0.51084	*1.40*	*0.1623*

Note: Signif. Levels: ***p < 0.001, ***p < 0.01, *p < 0.05.*

The results change after removing choice sets with unrealistically small or large travel times, as shown in Table B4.

Table B 4: LCM with Reduced Number of SP Choice Sets

Variable	Class 1 (49.4%)				Class 2 (50.6%)			
	Parameter	Std. error	z	p	Parameter	Std. error	z	p
ASC Car	0.79589***	*0.63215*	*1.26*	*0.208*	2.65764***	*0.74272*	*3.58*	*0.0003*
BTTCAR (In-vehicle travel time)	0.05071**	**0.01653**	**3.07**	**0.0022**	-0.03619***	**0.00815**	**-4.44**	**<0.001**
BPARKCAR (Parking availability)	*0.03809*	*0.10755*	*0.35*	*0.7232*	*0.09617*	*0.07382*	*1.30*	*0.1927*
BOPCOST (Operation costs)	-0.10206**	**0.02144**	**-4.76**	**<0.001**	*0.00418*	*0.01292*	*0.32*	*0.7463*
BPARKC (Parking costs)	*0.0161*	*0.02715*	*0.59*	*0.5532*	-0.10241***	**0.01781**	**-5.75**	**<0.001**
ASC PnR	-1.56129**	**0.83601**	**-1.87**	**0.0618**	*0.88068*	*0.77080*	*1.14*	*0.2532*
BTTPNR (In-vehicle travel time)	-0.0782**	**0.02160**	**-3.62**	**<0.001**	*0.00742*	*0.00787*	*0.94*	*0.3457*
BPARKPNR (Parking availability)	*0.06244*	*0.17387*	*0.36*	*0.7195*	*-0.02668*	*0.09987*	*-0.27*	*0.7894*
BTTWAIT (Waiting & transfer time)	0.21494**	**0.0816**	**2.63**	**0.0084**	*-0.00394*	*0.03170*	*-0.12*	*0.9012*
BPT (Bus)	0.08457*	*0.34279*	*0.25*	*0.8051*	*-0.00015*	*0.17522*	*-0.01*	*0.9993*
BTRF (Number transfers)	*-0.11301*	*0.24551*	*-0.46*	*0.6453*	*-0.00179*	*0.08599*	*-0.02*	*0.9834*
ASC PT WnR	*-1.13171*	*0.69073*	*-1.64*	*0.1013*	1.74478***	**0.67510**	**2.58**	**0.0098**
BTTWNR (In-vehicle travel time)	*-0.00743*	*0.02038*	*-0.36*	*0.7154*	*-0.00624*	*0.00751*	*-0.83*	*0.4060*
BTTWALK (Active travel time)	-0.26186***	**0.03107**	**-8.43**	**<0.001**	-0.03745***	**0.00444**	**-8.43**	**<0.001**
BTTTAXI (In-vehicle travel time)	0.02734*	0.01507	1.81	0.0696	-0.02781***	*0.01788*	*-1.56*	*0.1199*
BSHARE (Shared ride)	-0.58481**	**0.28937**	**-2.02**	**0.0433**	*0.33030*	*0.30855*	*1.07*	*0.2844*
ASC Bike	1.3678**	0.66583	2.05	0.04	1.70064**	**0.71792**	**2.37**	**0.0178**
BTTBIKE (Travel time)	-0.19788***	**0.03548**	**-5.58**	**<0.001**	-0.03095***	**0.00641**	**-4.83**	**<0.001**
ASC Walk	5.84387***	**0.82386**	**7.09**	**<0.001**	2.52921***	**0.72247**	**3.50**	**0.005**

Note: Signif. Levels: ***p < 0.001, **p < 0.01, *p < 0.05.

The results reveal significant differences in model outcomes when using the full dataset versus the dataset after removing choice tasks with unrealistic travel times. Here is an overview of the differences and the rationale for using the reduced choice set:

Improved Fit and Behavioural Realism:

- With the full dataset, Class 1 exhibited a lack of sensitivity to car travel times and costs but a heightened sensitivity to parking availability. This unusual behaviour could be linked to the unrealistic travel times in the dataset, which may have skewed preferences.

- After cleaning the dataset, Class 1 showed a more expected sensitivity to travel times and costs (e.g., BTTCAR became positive and significant). However, the in-vehicle travel time coefficient > 0, remains a concern. Class 2 also reflected realistic trade-offs commonly documented in the literature, reinforcing the validity of the reduced dataset.

2. **Enhanced Model Parameters**:

   - Unrealistic travel times likely exaggerated certain coefficients in the full dataset. For example, active travel times (BTTWALK and BTTBIKE) showed disproportionately negative effects in Class 1 under the full dataset, potentially due to extreme values.

   - After cleaning, these coefficients became more aligned with behavioural expectations, emphasising the importance of valid and realistic data for accurate parameter estimation.

3. **Consistency Across Classes**:

   - In the full dataset, Class 1 displayed an unusual preference for active travel despite long travel times, coupled with an unrealistic tolerance for waiting and transfer times (BTTWAIT). This behaviour was largely corrected in the reduced dataset, where the preferences became more consistent with practical travel behaviour.

Rationale for Using Reduced Choice Tasks

The removal of choice tasks with unrealistic travel times ensures that the models reflect genuine travel behaviour rather than artifacts of data inaccuracies. Unrealistic times can distort preferences and inflate or suppress sensitivity to key variables, leading to erroneous conclusions. By using the reduced dataset, the models yield more reliable and interpretable results, improving their applicability for policy evaluation and decision-making.

Thus, the reduced choice tasks not only improve model fit (evidenced by better log-likelihood values and more meaningful parameter estimates) but also ensure that the findings are grounded in behavioural realism.