University of South Australia

iMOVE

Australian Government
Department of Industry, Science and Resources

AusIndustry
Cooperative Research Centres Program

Government of South Australia
Department for Infrastructure and Transport

A non-parametric framework to transfer household travel survey data across jurisdictions

Authors

Hien (Eric) Duc Han
Research Fellow
University of South Australia
hien.han@mymail.unisa.edu.au

Akshay Vij
Associate Professor
University of South Australia
akshay.vij@unisa.edu.au

Ali Ardeshiri
Senior Research Fellow
University of South Australia
ali.ardeshiri@unisa.edu.au

Acknowledgements

This study was funded by the South Australian Minister for Infrastructure and Transport acting through the Department of Infrastructure & Transport (DIT) South Australia; and the iMOVE Cooperative Research Centre (CRC).

Version: 16 Apr 2024

Table of Contents

EXECUTIVE SUMMARY

Smaller jurisdictions frequently do not have the resources to collect household travel survey (HTS) data. This study introduces a novel non-parametric technique for combining and reweighting raw HTS datasets collected in other jurisdictions to create a ‘pseudo-HTS’ dataset that is representative of the target region, in terms of both the demographic characteristics of the local population of the target region, and the level-of-service of the local transport system.

We hypothesise that a significant proportion of variance in travel behaviour decisions within an urban area can be explained by the relative location of people and jobs within that area. For example, imagine an individual observed in the Melbourne HTS, living in St Kilda East who commutes to the Melbourne CBD on observation day, and stops in Prahran-Windsor on the way back, as shown in Figure 1. Our non-parametric framework tries to find the closest match to these three locations in the target region, Greater Adelaide in this case, such that the relative location of people and jobs between the locations in Greater Adelaide is comparable to the relative location of people and jobs between Melbourne CBD, St Kilda East and Prahran-Windsor. The level of similarity between the source and target regions is used as a weight to control for differences.

We apply the framework to generate a pseudo-HTS for Greater Perth (GP), using HTS data collected in other jurisdictions. We compare inferences from the inferred pseudo-HTS with those from the actual HTS conducted in GP to validate the methodology. We use GP as our target region instead of GA because recent HTS data is available for the former (which is used to validate the methodology).

For TDM sub-models that are estimated at a trip-level, our approach appears to perform reasonably well. In particular, estimation results for trip distribution and mode choice sub-models estimated using the original WA HTS and our inferred pseudo-HTS yield highly comparable estimates. However, sub-models estimated at a more aggregated level may potentially be biased due to limitations inherent in our approach. Our framework tries to match behaviours at a trip or tour-level. Trips or tours that do not have a good match in the target region are discarded. When the transferred trips or tours are aggregated at a higher level, the discarded trips/tours can create biases. For example, our framework underestimates the average number of daily trips undertaken by an individual. Similarly, at a zone-level, for zones in the target region that have good matches in the source region, our framework overestimates the trips attracted by these zones. Conversely, for zones in the target region that do not have a good match in the source region, our framework underestimates the trips attracted by these zones.

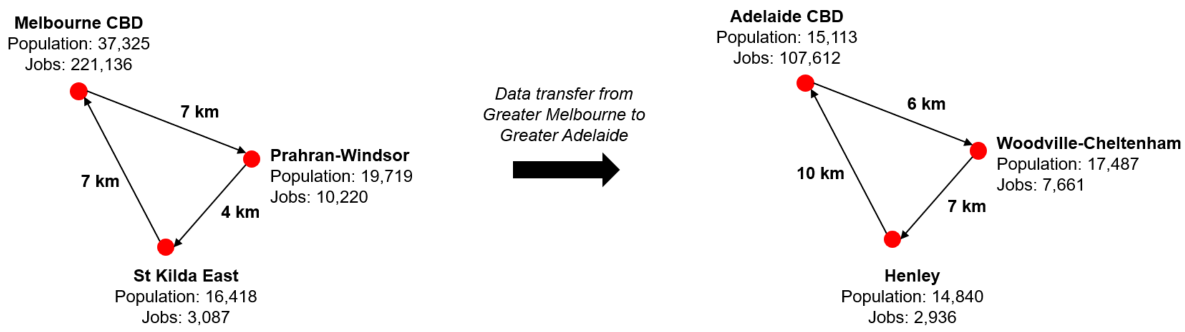


Figure 1: Illustrative example showing how travel behaviours observed in Greater Melbourne would be transferred to Greater Adelaide using our non-parametric framework

1 INTRODUCTION

Travel demand models (TDMs) are quantitative tools that are used by local, regional and national planning organizations for the development of evidence-based transport policy. TDMs can offer insights on current patterns of travel behaviour and provide a framework for predicting changes in behaviour in response to changes in the transport and land use system. Forecasts from TDMs are used to determine the capacity that new infrastructure must provide; and to facilitate the economic, environmental and social impact assessments of competing initiatives.

Traditionally, TDMs have been calibrated and validated using data collected through household travel surveys (HTSs) that ask participating individuals about their travel patterns over a 1 or 2-day observation period. These data collection methods are expensive. Hartgen and San Jose (2009) report average costs of $487,000 per HTS, and roughly $150 per response, though they note that "many surveys cost considerably more than the average, and the spread of the data is substantial". The Sydney Household Travel Survey currently samples roughly 5,000 households each year from a population of roughly 5 million, and observes their travel patterns over a 24-hour period, at an estimated cost of approximately $1 million per year.

The last HTS in metropolitan Adelaide was conducted in 1999. That data is now more than twenty years old, and not reflective of current or future travel patterns within the region. However, limited resources have precluded the collection of more recent HTS data in the region. This study aims to develop a new methodology for recalibrating TDM model parameters that does not require collection of HTS data for the jurisdiction of interest, but uses HTS data collected by other comparable jurisdictions.

Transport planners in smaller jurisdictions that do not have the resources to invest in their own data collection exercises frequently use datasets from other comparable jurisdictions to calibrate their TDMs. In many cases, the parameters of the TDM may be borrowed directly from established values reported in the literature. For example, the US *National Cooperative Highway Research Program* (NCHRP) has drafted a best-practices report on the calibration and validation of TDMs that includes procedures for transferring datasets across jurisdictions, and recommended values for a subset of standard model parameters (see Cambridge Systematics, 2012).

As part of our previous work, DIT and UniSA developed and tested a similar methodology for the recalibration of SAM that does not rely on primary data collection methods. In particular, we used household travel diary datasets collected in six other jurisdictions within Australia and New Zealand between Jan 2018 and Feb 2020 to update model parameters. These jurisdictions were Greater Melbourne, Southeast Queensland (SEQ), Greater Perth, Greater Hobart and Australian Capital Territory in Australia, and Auckland from New Zealand. Wherever possible, we augmented these datasets with additional information available through the Census and other datasets collected by the Australian Bureau of Statistics (ABS) and the Commonwealth Government. In some cases, we updated model parameters based on established values reported in the literature. Finally, we collected travel diary data from a small sample of 493 residents within Greater Adelaide (GA) to validate our final model parameters as part of the 2021 Greater Adelaide Travel Survey (GATS).

In our previous work, we employed a two-step model averaging approach. First, for each sub-model within SAM, we used household travel diary and transport cost skims from each of the other jurisdictions to estimate model parameters for these jurisdictions. Second, we averaged parameters from the six different jurisdictions, whilst controlling for outliers, to identify the appropriate parameters for GA. By including multiple jurisdictions within our analysis from which to pick parameter estimates for the GA region, we were able to identify outliers more easily. By collecting primary travel diary data from a small sample of GA residents, we were able to identify which jurisdictions most closely correspond to GA.

The present study builds on this previous work through the development of an alternative data averaging approach. We propose combining and reweighting raw HTS datasets collected in other jurisdictions to create a ‘pseudo-HTS’ dataset that is representative of the target region, in terms of both the demographic characteristics of the local population of the target region, and the level-of-service of the local transport system. We apply the framework to generate a pseudo-HTS for Greater Perth (GP), using HTS data collected in other jurisdictions. We compare inferences from the inferred pseudo-HTS with those from the actual HTS conducted in GP to validate the methodology. We use GP as our target region instead of GA because recent HTS data is available for the former (which is used to validate the methodology).

The remainder of the report is organized as follows: Chapter 2 presents the data-averaging methodology developed by this study. Chapter 3 compares data from the pseudo-HTS and actual HTS for GP in terms of summary statistics. Chapters 4-7 undertake a more detailed comparison between the two datasets in terms of trip production, trip attraction, trip distribution and mode choice. Chapter 8 concludes with a summary of key findings and directions for future research.

2 METHODOLOGY, APPROACH AND TECHNICAL FRAMEWORK

As mentioned previously, transport planners in smaller urban areas that do not have the resources to invest in their own data collection exercises frequently use datasets from other comparable jurisdictions to calibrate their TDMs. In many cases, the parameters of the TDM may be borrowed directly from established values reported in the literature. For example, the US *National Cooperative Highway Research Program* (NCHRP) has drafted a best-practices report on the calibration and validation of TDMs that includes procedures for transferring datasets across jurisdictions, and recommended values for a subset of standard model parameters (see Cambridge Systematics, 2012). The transferred model parameters can subsequently be used to generate a pseudo-HTS for the target region using microsimulation methods.

For example, Greaves (2000), Stopher et al. (2003) and Greaves (2006) have previously used variations of this general method to generate pseudo-HTS data for Adelaide. The approach involves utilizing a sample of actual households with known socio-demographic information and then simulating travel pattern characteristics based on distributions derived from data and models from other HTSs. Broadly, their methodology comprised three distinct steps. First, they generated a synthetic population comprising a subset of households, with accompanying socio-demographic characteristics, using the 1996 Census Household Sample File (HSF) sourced from the 1996 ABS Census data specific to the Adelaide region. Second, they estimated model parameters denoting the impact of key transport, land use and demographic variables on key travel behaviours of interest, using HTS data from Dallas and Salt Lake City in the US, primarily due to the absence of sufficiently recent national HTS data in Australia. Third, they simulated travel behaviours for the synthetic population, feeding their socio-demographic characteristics from the first step as input variables into the models estimated in the second step. The simulated data were subsequently compared against the 1999 Metropolitan Adelaide Household Travel Survey (AHTS) to assess the simulated data's ability to replicate the observed travel patterns in the AHTS survey.

In general, findings from these and other related studies in the broader literature on the transferability of travel demand models indicate that the method is sufficiently accurate in most cases at generating travel behaviours for the target region. However, the method can only generate behaviours specified explicitly as dependent variables in the travel demand models. For example, if the mode choice sub-model excludes modes such as taxis, ridesharing and e-scooters, then the model cannot predict these mode choice behaviours, and consequently, use of these modes is not captured by a pseudo-HTS generated using this method. If the analyst were interested in understanding or predicting use of these excluded modes in the target region, they would have to redo the second and third steps of the methodology, as described above, making sure to include these modes explicitly as dependent variables in the second step.

In this study, we develop an alternative non-parametric method that seeks to mitigate these limitations of the traditional parameter-transfer approach. We hypothesise that a significant proportion of variance in travel behaviour decisions within an urban area can be explained by the relative location of people and jobs within that area. For example, imagine an individual living in a low-density and sparsely populated suburb in Tasmania, 10 km away from the Hobart CBD, who commutes to the CBD every day by car. If this individual were transferred to the Perth metropolitan area, such that they lived roughly 10 km away from the Perth CBD, our non-parametric framework would predict that the individual commutes to the Perth CBD every day by car. Travel behaviours observed for other individuals belonging to other jurisdictions can similarly be transferred to the target region, and the collection of trips thus transferred can be treated as a ‘pseudo-HTS’ for the target region.

This is admittedly a naïve assumption that does not control for differences in other mitigating factors between the source and target regions, such as public transport network, road congestion and climate. In principle, one could make the non-parametric framework more sophisticated to account for differences in each of these variables. However, that might necessitate using a parametric framework of sorts similar to the original approach used by, among others, Stopher et al. (2003), as a parsimonious solution to the problem of how to incorporate a large number of explanatory variables. The benefit of the non-parametric framework is that the transferred behaviours are not constrained by model assumptions. If the individual in our example above took a taxi to work instead in Hobart, then they would be assumed to take a taxi to work if they were living in Perth at a comparable distance away from the CBD as they were in Hobart.

Our non-parametric framework may be described in terms of three critical steps:

1. **Preparing characteristics of origin and destination zones**: This involves compiling the relevant characteristics of origin and destination zones for each trip recorded in the given HTS.

2. **Transferring trips from the source HTS to the target HTS**: Trips from the source HTS are transferred to the target HTS based on the characteristics of the origin and destination zones.

3. **Combining transferred trips from different source jurisdictions and validation**: Transferred trips from all given HTS are combined into a single dataset meant to represent the target HTS, which allows for the validation of results.

In the subsequent sections, we delve into each of these steps in detail.

2.1 Preparing the characteristics of the origin and destination zones

After analysing the travel patterns observed in comparable Household Travel Surveys (HTS), we endeavour to align the origin and destination zones[1] of the given HTS with those of the target HTS based on the following transportation-related variables:

A. **Population and Employment**: This entails matching the number of residents and jobs located at each trip origin and destination.
B. **Distance**: We consider the geographical distance between each trip origin and destination.

To acquire this information, we could rely on data from land use and transportation databases. However, due to constraints in time and resources, we opt to substitute information from alternative sources. Specifically, we derive the number of residents and jobs for each zone from the ABS 2016 Census of Population and Housing, which provides comprehensive demographic and employment data. Additionally, to simplify the calculation, we determine the distance between each trip origin and destination zones as the geodetic distance between the coordinates attached to the centroids of these zones.

[1] We adopt SA2 as the zonal level throughout Stage III of the report, unless otherwise specified.

Table 1. An actual trip extracted from Greater Melbourne HTS

Trip ID	Origin	Destination	Mode	Purpose
1	Pakenham – North	Narre Warren – South West	Bus	Work
2	Narre Warren – South West	Pakenham -South	Bus	Shopping
3	Pakenham -South	Pakenham – North	Bus	Home

In Table 1, we present the actual travel movements of an individual as recorded in the Greater Melbourne Household Travel Survey (HTS). Specifically, the individual undertakes three distinct trips, commencing from what is presumed to be their residence in "Pakenham – North." Subsequently, the person travels to "Narre Warren – South West," with a stopover in "Pakenham - South," before returning to "Pakenham – North." The characteristics of each zone are depicted in Figure 2. For instance, "Pakenham – North" is inhabited by 18,486 individuals and accommodates 3,427 workers. In contrast, "Narre Warren – South West" is home to 13,989 residents and attracts 10,339 workers. Furthermore, the geodetic distance between "Pakenham – North" and "Narre Warren – South West" is measured at 15.4 km.

The next logical step is to identify zones within the target jurisdiction that exhibit similar characteristics to those of "Pakenham – North" and "Narre Warren – South West." However, a significant challenge arises when comparing the actual count of dwellings or employed persons in each zone, as well as the distance between the origin and destination zones. This challenge stems from the vast differences in scale between zones in different jurisdictions. For example, the Melbourne CBD SA2 zone records a population of 57,735 and 221,136 employed persons. Consequently, when a trip recorded in the Greater Melbourne HTS includes "Melbourne CBD," it becomes less feasible to transfer the trip to any zones in the study jurisdiction, such as Greater Adelaide, due to the lack of a zone with a comparable scale to that of "Melbourne CBD." To mitigate the impact of these scale differences, it is essential to normalize the characteristics of all zones across both the given and target HTS.

We then adopt Min-Max scaling method to restrain the zonal characteristics to a fixed range of between 0 and 1. This method provides some advantages such as simple implementation, preservation of the relationship between the original data points, and highly interpretable. The formula as of below:

$$X_{normalize} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Particularly, we collect the population (employment) of SA2 zones of the whole Australia, then apply the scaling method altogether. As provided in Table 2, the population and employment of “Pakenham – North” after scaling are 0.3202 and 0.0155, respectively. Meanwhile, “Narre Warren – South West” experience slightly fewer residents (13,989 versus 18,486), but significantly higher employment (10,339 versus 3,427) in comparison with “Pakenham – North”. As a result, the normalize population and employment of “Narre Warren – South West” are 0.2423 and 0.0468, respectively.

In terms of distances, we begin by gathering information on the distances between each pair of zones within each jurisdiction. Subsequently, we pool these pairs’ distances[2] together and utilize Min-Max scaling to normalize them. As illustrated in Table 2, the normalized distance

[2] Pairs of origin and destination zones exclusively belong to each respective jurisdiction within the dataset. In other words, no pair of zones consists of an origin zone from one jurisdiction and a destination zone from another. For instance, an origin zone may be in Greater Melbourne, while the destination zone would also be within Greater Melbourne, ensuring that all trips remain within the same jurisdictional boundaries.

between "Pakenham – North" and "Narre Warren – South West" is calculated as 0.0962. Moreover, the smallest recorded distances occur between Collingwood and Fitzroy in Melbourne, while the largest distances are observed between Beaudesert and Kilcoy in Queensland.

Figure 2. Visualization of an actual trip extracted from Greater Melbourne HTS

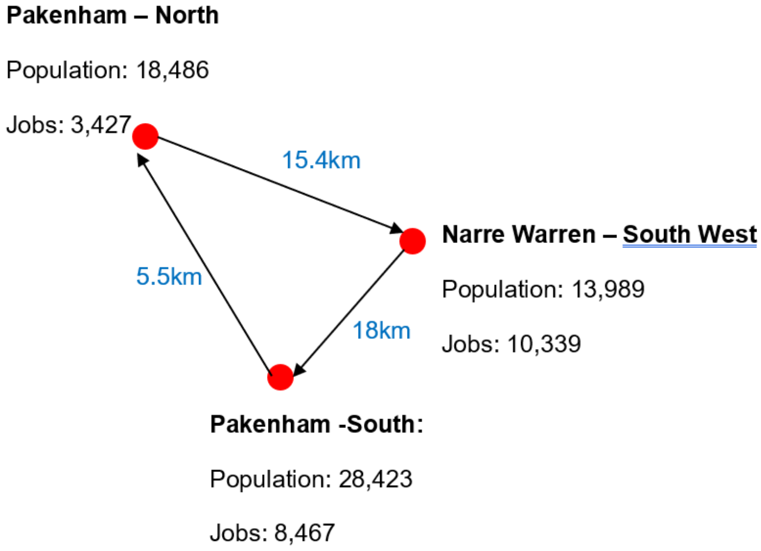


Table 2. Normalize zonal characteristics.

Pakenham – North		**Narre Warren - South West**		
Population	Employment	Population	Employment	Distance
18,486	3,427	13,989	10,339	15.4km
Normalized variable				
Origin-Pop	Origin-Emp	Dest-Pop	Dest-Emp	N_distance
0.3202	0.0155	0.2423	0.0468	0.0962

2.2 Transferring trips from the given jurisdiction to the target jurisdiction

In the preceding sections, we discussed the crucial zonal information required and the initial processing steps necessary to make this data comparable, setting the stage for meaningful comparisons later on. This section outlines the methodology for transferring trips from any given Household Travel Survey (HTS) to the target HTS within the study jurisdiction[3]. The primary objective is to identify zones in the target jurisdiction that closely resemble the characteristics of the zones recorded in the given HTS. To achieve this, various methods can be employed to generate similarity scores and address the challenge effectively.

Similarity scores quantify the resemblance or likeness between two objects, such as data points, based on their characteristics or features. These scores aim to measure how closely related or comparable the objects are. Various methods exist to compute similarity, including Euclidean distance, cosine similarity, Jaccard similarity, and Pearson correlation coefficient. Euclidean distance measures the straight-line distance between points in a Euclidean space, cosine similarity calculates the cosine of the angle between two vectors, Jaccard similarity computes the intersection over the union of sets, and Pearson correlation coefficient assesses the linear correlation between two variables. Each method has its own advantages and limitations, and the choice of similarity measure depends on the specific context and nature of the data being analysed. In this report we select Euclidean distance measures as our main proxy for similarity measurements due to its interpretability and wide applicability.

Euclidean distance, rooted in the Pythagorean theorem, finds its application in contexts where data can be represented as points within a Cartesian coordinate system. Mathematically, it is computed as the square root of the sum of squared differences between corresponding coordinates of two points. Consider points A and B in an n-dimensional space, each characterized by their own set of coordinates (x1, x2, ..., xn) and (y1, y2, ..., yn) respectively. The Euclidean distance between these points is then calculated as follows:

$$\text{Euclidean distance} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \qquad (2)$$

Euclidean distance is a critical metric for identifying comparable trips, as it incorporates both the population and employment characteristics of the origin and destination zones, along with the distances between them. Consequently, the Euclidean distance for each pair of zones must be calculated based on these five key characteristics, encompassing trip data recorded in the given Household Travel Survey (HTS) and all potential pairs of zones in the target jurisdiction. Subsequently, the pair of zones yielding the minimum Euclidean distance can be selected as the most suitable match. Therefore, Equation 2 can be reformulated as follows:

$$\text{Match pair of zones} = \min_{n} \sqrt{\sum_{i=1}^{5}\left(x_i - y_{i,k}\right)^2} \qquad (3)$$

[3] Before the matching process, we maintain all assumptions and considerations related to cleaning and preparing the household travel surveys (HTS), as adopted in Stage II. For example, with Greater Melbourne, trips made by individuals from households located outside of Greater Melbourne are removed as part of the cleaning process to ensure data accuracy and relevance to the specific geographic area under study. Or only keeping trips recorded during 2018-2020 period, etc.

Where,

- x_i is the set of characteristics of the actual trips recorded, which are,
    - the population of the origin,
    - the employment of the origin,
    - the population of the destination,
    - the employment of the destination,
    - and the distance between the origin and destination

- $y_{i,k}$ is the set of "*i*" characteristics of the *"k"-th* pairs of zones of the target jurisdiction.

The number "n" of all possible pairs of zones is determined by the combination of zones within the target jurisdiction. For example, if the target jurisdiction comprises 100 zones, the total number of pairs of zones is calculated as Combination - C (100, 2), resulting in 4,950 pairs of zones.

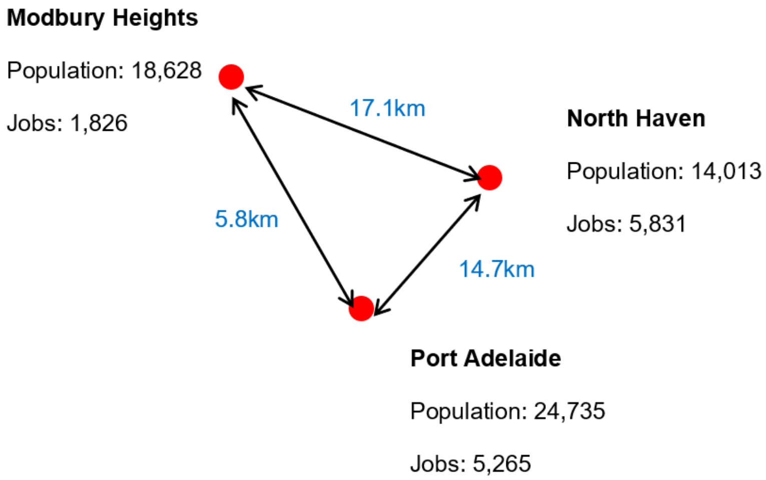
Building upon the previous example of actual trips made by a 30-year-old male residing in Greater Melbourne, the matched zones are provided at the conclusion of step 2 in Figure 3. Ideally, for each zone visited by the individual in Greater Melbourne, we aim to identify zones in Adelaide with similar characteristics in terms of population, employment, and distances between them. Consequently, the individual would have made three trips, beginning from "Modbury Heights" to "North Haven," then from "North Haven" to "Port Adelaide," before returning to "Modbury Heights."

The characteristics of each zone after being transferred to the Greater Adelaide (GA) dataset are only relatively comparable to zones in Greater Melbourne (GM) in terms of normalized figures. Illustrated in Table 3, "Modbury Heights" has a population of 18,628 and employment of 1,826, which normalize to 0.3226 and 0.0082, respectively. Similarly, "North Haven" has a population of 14,013 and employment of 5,831, which normalize to 0.2427 and 0.0263, respectively. The distance between "Modbury Heights" and "North Haven" is 17.1 km or 0.1069 in normalized distance. With this information, the Euclidean distance between the pair "Pakenham – North" and "Narre Warren - South West" in GM and the pair "Modbury Heights" and "North Haven" in GA is approximately 0.0243, which represents the lowest Euclidean distance when comparing the GM pair with all possible pairs generated from all GA zones.

Table 3. Characteristics of the zone that matched the actual trip from GM

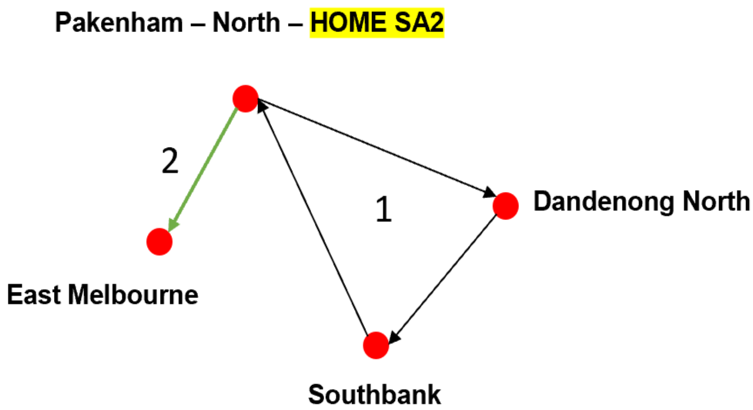
Modbury Heights		**North Haven**		
Population	Employment	Population	Employment	Distance
18,628	1,826	14,013	5,831	17.1km
Normalized variable				
Origin-Pop	Origin-Emp	Dest-Pop	Dest-Emp	N_distance
0.3226	0.0082	0.2427	0.0263	0.1069

Figure 3. Visualization of trips after transferring to Greater Adelaide



However, transferring trips from given jurisdictions can present significant challenges. Figure 4 illustrates an example of actual trips made by another individual recorded in the Greater Melbourne Household Travel Survey (HTS). The individual exhibits two distinct tours in the record. The first tour involves trips made to three different locations: "Pakenham – North," "Dandenong North," and "Southbank." After visiting the last zone, the individual returns to their home SA2 zone, "Pakenham – North." Subsequently, the individual embarks on their second tour, traveling between their home SA2 zone, "Pakenham – North," and "East Melbourne." In practice, observed travel behaviours can become much more complex. This complexity presents the first dilemma in the process of transferring trips to the target jurisdiction."

Figure 4: Example of multiple tours made by an individual in GM HTS



Previously, we outlined the method of selecting pairs of origin and destination zones based on the minimum value of Euclidean distance. While this approach works effectively for transferring individual pairs of zones, it becomes more complex when dealing with tours consisting of multiple interconnected trips. Continuing with the example provided earlier, suppose an individual's first trip within a tour is from "Pakenham – North" to "Dandenong North," followed by a second trip from "Dandenong North" to "Southbank." In this scenario, the transferred trips should maintain a similar sequential pattern. That is, if the first trip starts from location "A" to "B," then "B" becomes the starting point of the second trip to another destination "C". Hence, it is plausible that the destination for one pair of zones could serve as the origin for another pair. As a result, the matching process becomes considerably more intricate, as we need to match not just one pair of zones, but multiple pairs of zones simultaneously, considering the sequential nature of tours and potential overlaps between destinations and origins. Moreover, to add further complexity, we may need to consider whether the individual's home SA2 zone remains consistent across different tours.

Further, as reported in Table 4, we first analyse the number of distinct locations that individuals within each of the five Household Travel Surveys (HTS) may travel between. This count represents the number of unique locations visited by individuals, irrespective of the specific routes taken. We specifically separate tours[4] that a person may travel. Then for each tour that a person made, we count how many distinct zones that he/she travel between. We then present the distribution of the number of unique locations that an individual may travel between. The data indicates that approximately 80% to 90% of individuals travel between two to three distinct locations, while it is less common for individuals to travel among more than five different zones.

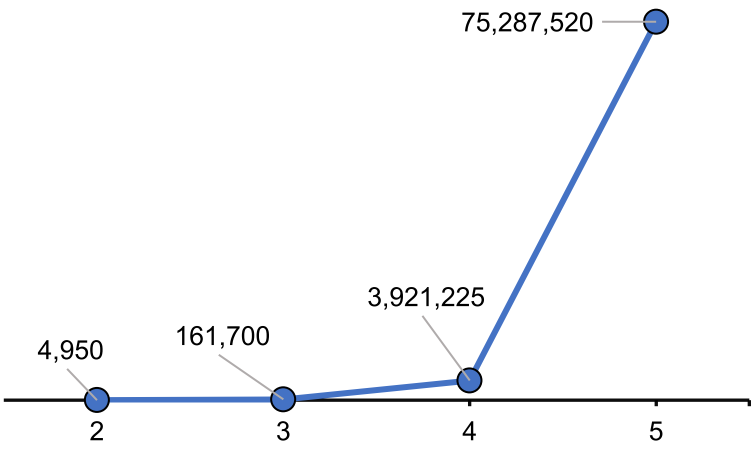
Table 4. Summary of the proportion of number of different zones that individuals travel across jurisdictions.

Number of different zones	**GM**	**SEQ**	**WA**	**ACT**	**TAS**
2	77%	80%	70%	69%	57%
3	18%	16%	21%	22%	25%
4	4%	3%	6%	7%	12%
>=5	1%	1%	3%	3%	6%

As previously discussed, the matching process may need to be executed concurrently for multiple pairs of zones. As the number of zones requiring matching increases, the number of calculations grows significantly. For example, if the target jurisdiction comprises 100 zones, and we need to transfer complex tours involving multiple zones, the number of calculations would be a combination of C (100, r), where r is the number of different zones in the tour that need to be transferred. Figure 5 illustrates the relationship between the number of combinations (y-axis) for a set of 100 zones and the increasing number of selected zones from the set (x-axis).

[4] A tour is defined as a series of trips that are completed by a person, with each trip ending back at their initial starting zone or home zone.

Figure 5. Number of combinations with increasing number of choosing objects from the set



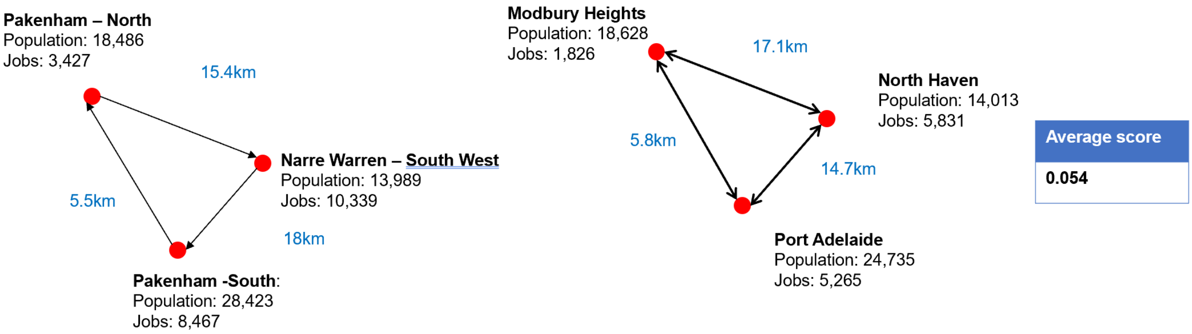
Given the limitations of our report, we propose several solutions to address the challenges outlined above. First, we opt for a naive approach where each trip is transferred independently, without considering potential connections to other trips within an individual's tour. Throughout the remainder of the report, we refer to the transferred Household Travel Surveys (HTS) resulting from this method as "Pseudo-HTS transfers at the trip level"[5].

Secondly, we introduce constraints at the tour level. Specifically, we analyse each individual's travel behaviour and divide it into distinct tours. We ensure that trips within each tour are connected, replicating the structure observed in the given HTS. We refer to this approach as "Pseudo-HTS transfers at the tour level" from this point onward. The most suitable tour is determined as the one experiencing the smallest average Euclidean distance among all origin-destination pairs observed in the actual HTS. Figure 6 illustrates the optimal matching at the tour level, enforcing constraints on connected trips. In this example, the matching tour comprises three distinct locations: "Modbury Heights," "North Haven," and "Port Adelaide," corresponding to the translated zones from "Pakenham – North," "Narre Warren – South West," and "Pakenham – South," respectively. Additionally, the three pairs of zones are connected in a manner consistent with the observed tour. The resulting average Euclidean score is reported as 0.054, calculated as the average score from each pair of zones.

To mitigate the exponential increase in calculations that could become infeasible in the short term, we purposefully exclude any tours consisting of five or more locations from the Pseudo-HTS transfers at the tour level. This decision is guided by the smaller proportion of individuals observed to travel between five or more unique locations, as illustrated in Table 4. By focusing on tours with a smaller number of locations, we aim to streamline the matching process and prioritize computational efficiency without sacrificing the representativeness of the transferred data.

[5] We use the terms “Pseudo-HTS” and “transferred HTS” interchangeably throughout the report.

Figure 6. Example of matching at tour level



2.3 Validation of the Pseudo Household Travel Survey

In this section, we delve into the calibration process of the synthesized HTS data from step 2. We emphasize the paramount importance of justifying the chosen method by thoroughly assessing whether the translated HTS effectively mirrors the travel behaviour patterns within the target jurisdiction. This assessment holds significant weight as it directly impacts the accuracy and relevance of the synthesized data. Ultimately, the goal is to enhance the utility of the synthesized data for subsequent analysis and decision-making processes concerning policy, planning, and resource allocation within the target jurisdiction.

We then propose the strategy of utilizing a known HTS as a base for comparison. By estimating the traditional four-step travel demand model using both the known HTS and the synthesized HTS, we can compare the resulting parameters to determine if the synthesized HTS can produce predictions similar to those derived from the known HTS. This comparative analysis serves as a robust validation method, allowing us to assess the effectiveness and reliability of the synthesized data in capturing the intricacies of travel behaviour within the target jurisdiction. We will employ the Greater Perth Household Travel Survey (HTS) as the reference dataset for contrasting with the Pseudo-HTS. This choice stems from the perception that the transportation patterns and characteristics in Western Australia (WA) bear greater resemblance to those of Greater Adelaide.

The current design of SAM is developed from the traditional four-step travel demand model. This adaptation involves the utilization of ten sub-models to forecast output flows across the transport network. Parameters within each component are updated based on demographic, land use, and transport network data specific to the Greater Adelaide (GA) metropolitan region for the year 2021. However, for the purpose of evaluating the suitability of the synthesized Household Travel Survey (HTS), a streamlined approach is adopted. This involves the adoption of a much simpler four-step travel model with eliminating trip purpose segmentation and reducing the number of model components. As a result, the analysis can prioritize key variables and facilitate a more efficient evaluation of the Pseudo-HTS's performance in estimating transport network flow. Specifically, a reduced version of trip production, trip attraction, trip distribution, and mode choice components are employed in subsequent analyses.

3 SUMMARY OF PSEUDO HOUSEHOLD TRAVEL SURVEYS

This section describes the Pseudo-HTSs that we comprise after the second step. In Table 5 illustrates the number of trips transferred from each given HTS. Overall, the Pseudo-HTS at trip and tour level comprises 126,941 and 122,919 trips, respectively[6]. Southern Queensland and Greater Melbourne account for more than 70% of all of the transferred trips, while Greater Hobart and Australian Capital Territory make up the rest of around 26-27%. As previously discussed, at tour level, we do not translate tours that travel from 5 or more places. As a results, the number of trips in the Pseudo-HTS at tour level is slightly lower than that at trip level. In details, the total number of trips removed in the HTS tour level is 4,022, which represents 2%, 3%, 7% and 4% of the figures reported at trip level for SEQ, GM, TAS and ACT, respectively. The proportions of removed trips is comparable to the proportions of number of different zones that are reported in each tour made by individuals (Table 4). For instance, in Greater Hobart, 6% of trips made belong to tours that comprise of trips made between 5 or more locations. Hence, we experience 7% of trips being remove in the Pseudo-HTS at tour level when we translate trips from Greater Hobart. The Pseudo-HTS at trip level, however, does not suffer from trip lost since we transfer each trip separately.

In Table 6, we represent more summary statistics about the Pseudo-HTS and compare with the original HTS from Greater Perth. In Panel A, the number of trips and number of persons record in the new HTS are similar with HTS at tour level is slightly lower than that at trip level. Hence, the trip rates per person are 3.64 and 3.54, which is quantitatively equivalent. Nonetheless, the trip rate per person recorded in WA HTS is 4.48 which is noticeably higher than the number reported earlier. In Panel B and C, we break down the HTSs based on the assigned mode of transportation and number of car ownership as further comparisons. The distribution of trips based these categories are almost identical regardless of whether it is origin or Pseudo-HTS. Particularly, Car is the dominance choice with around 80% of trips reported using this mean of transportation, followed by Walk, PT and Bicycle. Moreover, most individuals in these HTSs reported to have at least one car. Opposing to just a fraction of 1-2% possess no car. Further descriptions of the Pseudo-HTSs will be discussed in the next few sections where it is applicable.

Table 5. Number of trips transferred from each given HTS.

Jurisdiction	Trip level		Tour level			Trip removed in HTS Tour level	% drop compare with trip level
	N	%	N	%			
SEQ	54,548	43%	53,589	44%		959	2%
GM	38,206	30%	37,067	30%		1,139	3%
TAS	20,075	16%	18,763	15%		1,312	7%
ACT	14,112	11%	13,500	11%		612	4%
Total	126,941		122,919			4,022	

Table 6. Summary of HTS from WA and the Pseudo-HTSs

Panel A. Number of trips and persons			
	Number of trips	Number of persons	Trip rate per person
WA HTS	28,727	6,413	4.48
Pseudo-HTS - trip level	126,941	34,869	3.64
Pseudo-HTS - tour level	122,919	34,690	3.54

[6] The numbers are total number of trips counted for all individuals.

Panel B. Assigned mode summary						
	WA HTS		Pseudo-HTS - trip level		Pseudo-HTS - tour level	
	N	%	N	%	N	%
Bicycle	435	2%	1,918	2%	1,676	2%
Car	23,250	81%	100,148	79%	85,012	78%
PT	1,272	4%	8,110	6%	6,367	5%
Walk	3,770	13%	16,765	13%	15,855	15%

Panel C. Car ownership summary						
	WA HTS		Pseudo-HTS - trip level		Pseudo-HTS - tour level	
	N	%	N	%	N	%
0 CAR / HH	388	1%	2,543	2%	2,253	2%
1 CAR / HH	6,618	23%	34,566	27%	30,047	27%
2 CARS / HH	14,015	49%	60,547	48%	51,979	48%
3+ CARS / HH	7,706	27%	29,285	23%	24,631	23%

4 TRIP PRODUCTION

In the conventional four-stage modelling approach, the initial phase involves trip end generation, which calculates the number of origins and destinations within each Traffic Analysis Zone (TAZ). In prior sections of this report, we used stratified household trip characteristics along with zoning land-use data to estimate trip origins and destinations for individual trips. These trips can be classified in various ways, considering factors such as trip purpose, frequency, timing, distance, spatial separation between origin and destination (O-D), and mode of travel. In the SAM model, trips are categorized based on their purposes. However, for the purpose of validating the Pseudo-HTS, the model is simplified to predict the total number of trips generated at the individual level, regardless of their purposes.

4.1 Model structure

For modelling trip production, a regression function with the following was defined:

$$y_h = \boldsymbol{\alpha} + \boldsymbol{\beta} w_h + \boldsymbol{\gamma} r_h + \sum_{k=1}^{3} \boldsymbol{\pi_k} c_i \tag{3}$$

The variables used in this equation were defined as following:

y_h	Number of trips produced by individual h
w_h	Dummy variable equals to one for working individual (age of 18 – 64), zero otherwise.
r_h	Dummy variable equals to one for retired person (age of 65+), zero otherwise.
c_i	Number of cars owned by household i that the person h belong to
α, β, γ, and π_k	Model parameters to be estimated

4.2 Results and discussion

Table 7 presents a comparative analysis of trip production model results derived from different HTS. In dissecting the parameters, several key observations emerge. First and foremost, the constant terms across the models exhibit slight variations, indicating nuanced differences from the baseline level of trip production captured by the original WA HTS. Moving beyond the constants, the estimated coefficients for variables such as working age group, retired age group, and the number of cars per household (1 car, 2 cars, 3+ cars) vary between the three HTSs. The analysis reveals subtle variations in the coefficients of different age groups and levels of car ownership between the Original Work Activity-based Household Travel Survey (WA HTS) and the Pseudo Household Travel Survey (HTS) at both trip and tour levels. For instance, the coefficient for the working age group is slightly higher in the Pseudo-HTS at the trip level (1.47) compared to the Original WA HTS, but marginally lower at the tour level (1.35). Similarly, the coefficient for the retired age group is 1.18 in the Original WA HTS, lower than the Pseudo-HTS at the trip level (1.32) but higher than the corresponding coefficient at the tour level (1.05). This trend is mirrored in the coefficients related to levels of car ownership.

From these observations, it can be inferred that the transferred HTS data at the trip level may overestimate trip generation by individuals, while at the tour level, it tends to underestimate it. However, despite these quantitative differences, the scales of variation are relatively small. Therefore, it is hypothesized that the transferred HTS data, despite the discrepancies, could serve as a viable alternative source of information for trip generation sub-models.

In conclusion, the comparative analysis of trip production model results provides valuable insights into the nuances of transportation modeling approaches. By examining the variations in estimated coefficients, goodness-of-fit measures, and sample sizes, stakeholders in urban planning and transportation management can make informed decisions regarding infrastructure investments, policy interventions, and mobility solutions, ultimately contributing to the development of more efficient, sustainable, and resilient transportation systems.

Table 7. Trip production model results

Variable	**Original WA HTS**		**Pseudo-HTS trip level**		**Pseudo-HTS tour level**	
	est	**t-stat**	**est**	**t-stat**	**est**	**t-stat**
Constant	2.16	8.65	2.37	17.05	2.32	38.90
Work	1.41	18.12	1.47	34.10	1.35	37.13
Retired	1.18	10.82	1.32	23.20	1.05	22.15
1 CAR / HH	1.27	5.08	1.52	10.96	1.18	19.32
2 CARS / HH	1.43	5.80	1.71	12.50	1.55	26.86
3+ CARS / HH	1.20	4.82	1.54	11.07	1.18	18.70
Adjusted R2	0.05		0.04		0.04	
No. obs.	6,413		34,870		34,690	

5 TRIP ATTRACTION

Analysing the trip attraction model presents greater complexity compared to estimating the trip production model. This complexity stems from the fact that the sampling strategy in household travel surveys does not specifically target trip attraction locations. Consequently, it is common practice to estimate the parameters at an aggregated level, such as Local Government Areas (LGA) or Statistical Area Level 2 (SA2), rather than at a disaggregated spatial level like travel area zones. Typically, the model is estimated using a linear function incorporating variables that represent the level of activity in a zone, such as employment types, student enrolment at educational institutions, and households or population counts for various trip purposes. In this version, we employ three key variables which are population, employment and total enrolment at each SA2 level.

5.1 Model structure

In the validation step, trip attraction sub-model is using the following linear regression equation based on population, employment and number of enrolments:

$$y_j = \boldsymbol{\alpha} + \boldsymbol{\beta} population_j + \boldsymbol{\gamma} emp_j + \boldsymbol{\pi} enrol_i \quad (4)$$

The variables used in this equation were defined as following:

y_j	Number of trips attracted[7] by zone *j*
$population_j$	Population living in zone *j*
emp_j	Employment working in zone *j*
$enrol_i$	Number of enrolments in zone *j*
α, β, γ, and π	Model parameters to be estimated

5.2 Results and discussion

The results of a trip attraction sub-model are presented in Table 8, comparing estimates derived from the Original WA Household Travel Survey (WA HTS) with those obtained from the Pseudo Household Travel Survey (HTS) at both trip and tour levels.

Firstly, significant differences are evident across the models. The constant term in the Pseudo-HTS at the trip level is notably higher (88,710) compared to the Original WA HTS (5,313) and the Pseudo-HTS at the tour level (83,460). This substantial variation suggests differing baseline levels of trip attraction captured by each model, which could be attributed to methodological discrepancies or variations in sample representation.

Regarding the independent variables, the coefficients associated with population, employment, and enrolments demonstrate distinct patterns across the models. In the Original WA HTS, population and employment exhibit substantial positive coefficients, indicating strong associations with trip attraction. However, in the Pseudo-HTS at both trip and tour levels, the coefficients for population and employment are substantially lower, with some even displaying negative values. This suggests divergent interpretations of the

[7] The number of trips attracted by each zone is calculated as the sum of the total number of trips multiplied by their corresponding trip weights for each zone.

impacts of population and employment on trip attraction between the Original WA HTS and the Pseudo-HTS models.

Furthermore, the coefficient for enrolments, representing educational enrollments, shows inconsistency across the models. While the Original WA HTS presents a positive coefficient for enrolments, indicating a positive relationship with trip attraction, the coefficients in the Pseudo-HTS models are negative. This discrepancy highlights contrasting findings regarding the influence of educational enrollments on trip attraction between the Original WA HTS and the Pseudo-HTS models.

In terms of model fit, the adjusted R-squared values vary considerably between the models. The Original WA HTS model exhibits a relatively high adjusted R-squared value of 0.741, indicating a good fit of the model to the data. In contrast, the Pseudo-HTS models at both trip and tour levels display substantially lower adjusted R-squared values of 0.017 and 0.03, respectively, suggesting poorer model fits compared to the Original WA HTS model. In brief, the comparison of the trip attraction sub-model results reveals notable differences between the Original WA HTS and the Pseudo-HTS models at both trip and tour levels. These differences underscore the importance of the transferred datasets.

Table 8. Trip attraction model results

Variables	Original WA HTS		Pseudo-HTS trip level		Pseudo-HTS tour level	
	est	t-stat	est	t-stat	est	t-stat
Constant	5,313	0.88	88,710	14,400	83,460	12,200
Population	3.98	12.74	-1.10	0.74	-1.39	0.63
Employment	34.70	11.45	6.10	7.21	9.10	6.09
Enrolments	0.94	1.12	-1.43	2.00	-0.90	1.69
Adjusted R2	0.741		0.017		0.03	
No. obs.	170		170		170	

The investigation delves deeper into the significant disparities observed in the estimation of trip attraction between the Pseudo Household Travel Surveys (HTSs) and the original Perth Household Travel Surveys (HTS). Another approach is employed wherein the actual number of trips attracted by each zone is derived and aggregated for comparative analysis.[8]

In Table 9, zones such as "Perth City" emerge as the primary attractors of trips, aligning with expectations as these zones typically serve as centers for various activities including work, entertainment, and education. However, a profound shift in zonal distribution is observed when utilizing data from the Pseudo-HTS at the trip level, as evidenced in Table 10. Notably, zones such as "Winthrop," which originally attracted a mere 0.07% of total trips according to the original WA HTS, now ascend to the top position, capturing 4.41% of trips as per the Pseudo-HTS at the trip level. Similarly, other zones within the top 10 in Table 10 experience significantly elevated proportions of trips compared to the original WA HTS. Consequently, critical zones such as "Perth City" will exhibit lower numbers of trips relative to the total number of trips.

Similarly, Table 11 outlines the top 10 zones based on the number of attracted trips utilizing the Pseudo-HTS at the tour level. The observations echo the trends observed with the Pseudo-HTS at the trip level. Zones such as "Midland – Guildford" and "Bateman" attract 5,938 and 3,493 trips, respectively, constituting 4.83% and 2.84% of the total number of

[8] Full table with all Greater Perth zones is reported at Table 15 in the Appendix.

trips. However, when combined, these two zones account for merely 1.3% of the total number of attracted trips reported in the Original WA HTS.

To further identifying the issue, we subsequently endeavor to find the optimal match for each zone within the given jurisdiction to the target jurisdiction. In this analysis, we simply utilize the population and employment data of each SA2 zone only, apply the Euclidean distance method as previously discussed, and seek to ascertain the most suitable match within the Greater Perth zone[9]. The outcome indicates that zones such as "Bateman" or "Midland – Guildford" in Greater Perth are more inclined to be the most suitable zones following the matching process for all zones from all given jurisdictions. We then conduct a comparison between the zones more likely to be matched and the top zones with the highest number of attracted trips using the two Pseudo-HTS datasets. This analysis reveals numerous similarities. Hence, it suggests that certain zones are overrepresented following the matching process, while others are undervalued. We represent the results of this process at Table 15 in the Appendix.

One potential explanation for this discrepancy could be attributed to the uniform consideration of population, employment, and distance between zones through the Euclidean distance measure that we employ. When the first two criteria, namely population and employment, are closely aligned, the method systematically underplays the significance of distances between two zones. Consequently, the minimum similarity score, utilized as our criterion for selecting the most suitable matched pair of zones, may only reflect the suitability based on population and employment. This bias, then, overlook the significance of geographical distances in the matching process.

In summary, this exercise represents a fundamental issue pertaining to the matching process of trips from given jurisdictions to the target jurisdiction. The failure to accurately reflect travel behavior between SA2 zones raises questions regarding the efficacy of the current methodology.

Table 9: Top 10 zones of attracting trips sorted by original WA HTS

Perth zone	Original WA HTS		Pseudo-HTS trip level		Pseudo-HTS tour level	
	N	%	N	%	N	%
Perth City	1,526	5.31%	1,958	1.54%	2,564	2.09%
Subiaco - Shenton Park	591	2.06%	669	0.53%	398	0.32%
Ellenbrook	566	1.97%	408	0.32%	60	0.05%
Nedlands - Dalkeith - Crawley	487	1.70%	445	0.35%	603	0.49%
Baldivis	475	1.65%	390	0.31%	712	0.58%
Willetton	473	1.65%	811	0.64%	1206	0.98%
Bicton - Palmyra	470	1.64%	530	0.42%	491	0.40%
Fremantle	431	1.50%	21	0.02%	22	0.02%
Wembley - West Leederville - Glendalough	405	1.41%	1,066	0.84%	1,415	1.15%
Mindarie - Quinns Rocks - Jindalee	394	1.37%	293	0.23%	739	0.60%

[9] We do not consider the distance since it is simply a matching process of zone-to-zone.

Table 10. Top 10 zones of attracting trips sorted by Pseudo-HTS trip level.

Perth zone	Original WA HTS		Pseudo-HTS trip level		Pseudo-HTS tour level	
	N	%	N	%	N	%
Winthrop	21	0.07%	5,593	4.41%	2,949	2.40%
Midland - Guildford	340	1.18%	4,932	3.89%	5,938	4.83%
Scarborough	176	0.61%	3,520	2.77%	1,137	0.93%
Bateman	35	0.12%	3,518	2.77%	3,493	2.84%
Willagee	97	0.34%	3,331	2.62%	2,878	2.34%
Morley	368	1.28%	3,126	2.46%	1,850	1.51%
Floreat	121	0.42%	2,781	2.19%	1,138	0.93%
Wanneroo	163	0.57%	2,580	2.03%	1,089	0.89%
Claremont (WA)	217	0.76%	2,460	1.94%	3,029	2.46%
Swanbourne - Mount Claremont	228	0.79%	2,266	1.79%	718	0.58%

Table 11. Top 10 zones of attracting trips sorted by Pseudo-HTS tour level

Perth zone	Original WA HTS		Pseudo-HTS trip level		Pseudo-HTS tour level	
	N	%	N	%	N	%
Midland - Guildford	340	1.18%	4,932	3.89%	5,938	4.83%
Bateman	35	0.12%	3,518	2.77%	3,493	2.84%
Claremont (WA)	217	0.76%	2,460	1.94%	3,029	2.46%
Winthrop	21	0.07%	5,593	4.41%	2,949	2.40%
Willagee	97	0.34%	3,331	2.62%	2,878	2.34%
Perth City	1,526	5.31%	1,958	1.54%	2,564	2.09%
Middle Swan - Herne Hill	111	0.39%	1,974	1.56%	2,421	1.97%
South Lake - Cockburn Central	200	0.70%	1,706	1.34%	2,355	1.92%
Rivervale - Kewdale - Cloverdale	364	1.27%	1,758	1.39%	2,319	1.89%
Madeley - Darch - Landsdale	216	0.75%	1,793	1.41%	2,230	1.81%

6 TRIP DISTRIBUTION

The trip distribution component of the model offers insights into the flow of trips between geographic units (such as SA2 areas). It considers factors like travel time or cost, known as travel impedance, along with trip generation data to determine the distribution of trips. These variables play a crucial role in shaping the trip distribution function.

6.1 Model structure

The trip distribution sub-model employs the conventional gravity model formulation to predict the movement of trips from one origin to another destination. This model characterizes the relationship between trip flows and the characteristics of both origin and destination locations:

$$y_{pcij} = a_{pci} b_{pcj} P_{pci} A_{pcj} f_{pc}\left(C_{ij}\right)$$

Where:

$$f_{pc}\left(C_{ij}\right) = C_{ij}^{\alpha_{pc}} e^{\beta_{pc} C_{ij}}$$

y_{pcij} = *Number of trips from origin zone i to destination zone j of purpose p made by households with car ownership level c*

a_{pci}, b_{pcj} = *Zone-specific row and column balancing factors*

P_{pci} = *Number of trips produced by zone j of purpose p made by households with car ownership level c*

A_{pcj} = *Number of trips attracted by zone j of purpose p made by households with car ownership level c*

$f_{pc}\left(C_{ij}\right)$ = *cost deterrence function for zone i to zone j*

C_{ij} = *Generalized cost of travel from zone i to zone j*

$\alpha_{pc}, \beta_{pc} = \ coefficients\ to\ be\ calibrated$

The generalized cost utilized in trip distribution is derived from the combined expenses associated with daily travel via both highway and public transportation modes. This blended cost is computed by integrating data from car and public transport skims, following a defined set of equations:

$$C_{ij} = S_{TD} = \frac{-1}{\lambda} ln\left(e^{-\lambda\, S_{PT}} + e^{-\lambda\, S_{HWAY}}\right)$$

$$S_{HWAY} = 100\beta(\alpha_T S_{HT} + \alpha_D S_{HD})$$

$$S_{PT} = 100\beta(S_{PS})$$

Where:

S_{TD} = *Output Blended Skims used in Trip Distribution (output)*

S_{HT} = *Highway Travel Time Skims (input)*

S_{HD} = *Highway Travel Distance Skims (input)*

S_{HWAY} = *Highway Skim Costs (working matrix)*

S_{PS} = *Public Transport Time Skims (input)*

S_{PT} = *Public Skim Costs (working matrix)*

α_T = *Assignment Time Factor* = 1

α_D = *Assignment Distance Factor* = 0.505

β = *Mode Choice Value of Time* = 0.188 ($/Min)

λ = *calibrated constant* = 0.05

Only the coefficients of the deterrence function α_{pc} and β_{pc} need to be estimated. Employing the blended cost methodology, cost bands are established in one-dollar increments, and the allocation of trips within each band is determined using data from household travel surveys. Subsequently, a model is developed with the subsequent general specification:

$$y_{pcn} = k_{pc} C_n^{\alpha_{pc}} e^{\beta_{pc} C_n}$$

Where:

y_{pcn} = *Number of trips of purpose p made by households with car ownership level c with generalised cost in cost band n*

C_n = *Mean generalized cost of travel for cost band n*

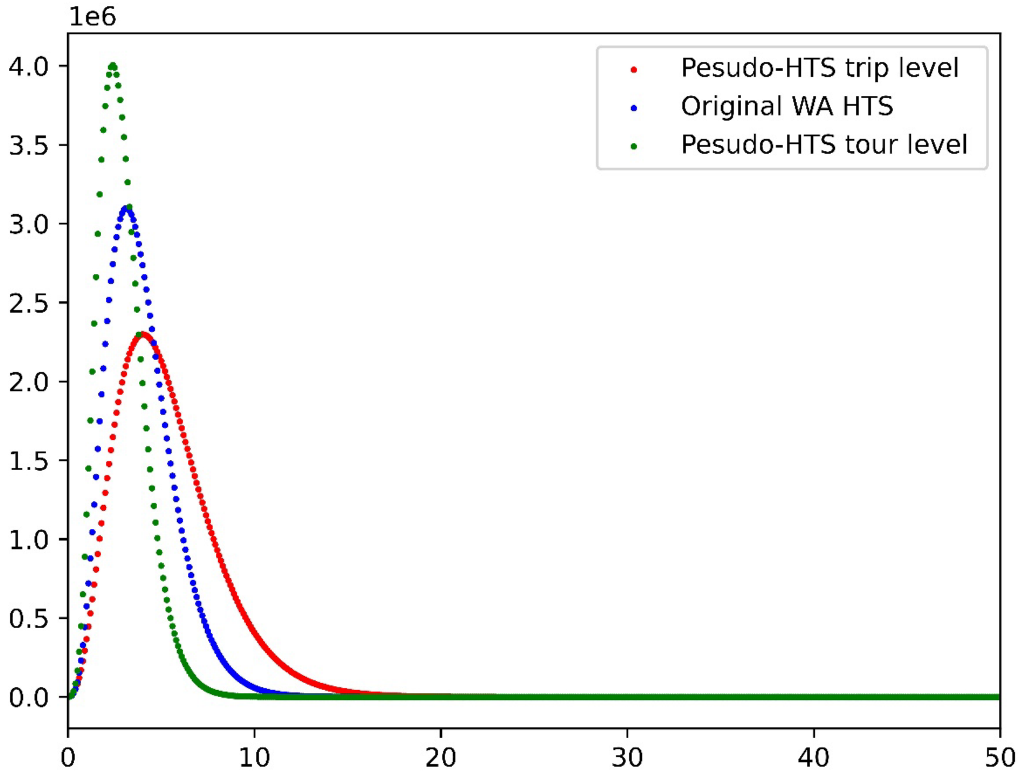
6.2 Results and discussion

Table 12 and Figure 7 present estimated values of alpha and beta coefficients derived from three HTS datasets. These coefficients are utilized in the deterrence function, which models the number of trips for a specific purpose made by households with varying car ownership levels and generalized travel costs. The alpha coefficient represents the sensitivity of trip numbers to changes in generalized travel cost, while the beta coefficient reflects the rate of decay in trip numbers as travel cost increases. In details, across the different HTS datasets, there are variations in the estimated values of both alpha and beta coefficients. For instance, comparing the alpha coefficients, we observe that the values range from 4.03 in the original WA HTS to 3.12 in the Pseudo-HTS at the trip level and 4.78 in the Pseudo-HTS at the tour level. Similarly, for the beta coefficients, the values vary from -1.25 in the original WA HTS to -0.76 in the Pseudo-HTS at the trip level and -1.91 in the Pseudo-HTS at the tour level. These variations indicate that the sensitivity of trip numbers to changes in generalized travel cost and the rate of decay in trip numbers with increasing travel cost differ across the HTS datasets. Nonetheless, these estimated parameters are quantitatively comparable across the different HTS datasets. This suggests that while there may be slight differences in the magnitude of the coefficients, the overall relationships between generalized travel cost and trip numbers remain consistent across the datasets. Such comparability strengthens the applicability of the Pseudo-HTSs.

Table 12. Trip distribution results

	Original WA HTS	Pseudo-HTS trip level	Pseudo-HTS tour level
Alpha	4.03	3.12	4.78
Beta	-1.25	-0.76	-1.91

Figure 7. Estimated deterrence functions.



7 MODE CHOICE

The mode choice sub-model serves to forecast the transportation mode preferences for individuals or groups. To accurately capture diverse transportation options, including both motorized and non-motorized alternatives, the mode choice splits should reflect a comprehensive range of options. This ensures a more realistic representation of multimodal transportation choices, accommodating varying preferences and needs within the transportation network.

7.1 Model structure

In SAM, the mode choice model comprises six alternatives, with three related to cars at different occupancy levels, and the remaining three representing public transport, cycling, and walking. Consistent with this approach, we maintain the same structure in this stage. However, unlike trip purpose segmentation, we do not differentiate between purposes and estimate a single mode choice model for each of the household travel surveys (HTSs).

Other assumptions, value of in-vehicle time, ratio of waiting time to in-vehicle time, etc.. implemented in the stage II remain unchanged.

The utility functions for each alternative has the following form:

$$u_{ntCAR} = \boldsymbol{\alpha}_{CAR} + \boldsymbol{\beta} ivt_{CAR} + \left(\frac{\boldsymbol{\beta} \times 0.6}{17.4}\right) cost_{CAR}$$

$$u_{ntPT} = \boldsymbol{\alpha}_{PT} + \boldsymbol{\beta}\left(ivt_{PT} + (\mathbf{1.5} \times wkt_{PT}) + (\mathbf{2} \times wtt_{PT}) + (\mathbf{5} \times trans_{PT})\right) + \left(\frac{\boldsymbol{\beta} \times 0.6}{17.4}\right) cost_{PT}$$

$$u_{ntACTIVE} = \boldsymbol{\alpha}_{ACTIVE} + \boldsymbol{\beta} t_{ACTIVE}$$

Where:

u_{ntj} = $Utility\ of\ alternative\ j\ for\ individual\ n\ and\ trip\ t$

$\alpha = Alternative\ specific\ constant$

$ivt_j = In - vehicle\ travel\ time\ for\ alternative\ J$

$cost_j = travel\ cost\ for\ alternative\ J$

$wkt_j = Walking\ time\ for\ alternative\ J$

$wtt_j = Wait\ time\ for\ alternative\ J$

$trans_j = Number\ of\ transfers\ for\ alternative\ J$

7.2 Results and discussion

We report the estimation results in Table 13. Comparing the coefficients estimated using the three datasets reveals minor differences in their magnitudes across various variables. For instance, the magnitude of the coefficient values for "BICYCLE_MC" differs slightly between the datasets, with the highest absolute value observed in the Original WA HTS (-3.136), followed by the Pseudo-HTS at the tour level (-3.000), and then the Pseudo-HTS at the trip level (-2.973).

Similarly, coefficients associated with car-related variables such as "B_COST" (car cost) and "B_TIME" (car travel time), also exhibit consistent negative values across all datasets, indicating decreases in utility with increases in these factors. Again, there are minimal variations in the magnitude of these coefficients across the datasets. On the contrary, estimated parameters for “B_WAIT_TIME” and “B_WALK_TIME” show distinguished differences between the original HTS of WA and our Pseudo-HTSs. In details, “B_WAIT_TIME” and “B_WALK_TIME” of the original HTS of WA are -0.465 and -2.564, which are five times larger in absolute term on average. The discrepancies are also the results from the higher concentration of certain zones after the matching process, which brings bias to the wait time and walk time for public transport.

For variables related to public transport (PT_MC) and walking (WALK_MC), the coefficients show mixed patterns across the datasets. While the coefficients for public transport are negative across all datasets, suggesting decreases in utility for this mode choice. The coefficient for walking is positive values observed in the Original WA HTS (0.406) but negative in the Pseudo-HTS at both trip level (-1.149) and tour level (-0.746). Overall, while the general trends in coefficient signs are consistent across the datasets with less variations in the magnitudes of these coefficients.

Table 13. Mode choice results

Variables	**Original WA HTS**		**Pseudo-HTS trip level**		**Pseudo-HTS tour level**	
	Value	**Rob. t-test**	**Value**	**Rob. t-test**	**Value**	**Rob. t-test**
BICYCLE_MC	-3.136	-33.163	-2.973	-70.252	-3.000	-64.379
B_COST	-0.483	-18.123	-0.361	-41.405	-0.331	-29.591
B_TIME	-2.916	-8.856	-2.150	-18.495	-2.145	-16.159
B_WAIT_TIME	-0.465	-4.487	-0.177	-3.721	0.097	1.958
B_WALK_TIME	-2.564	-12.009	-0.196	-4.321	-0.598	-11.280
OCC2_MC	-1.018	-58.564	-1.057	-112.617	-1.019	-100.622
OCC3p_MC	-1.367	-78.902	-1.333	-145.802	-1.268	-130.671
PT_MC	-0.704	-6.131	-1.472	-37.510	-1.730	-37.818
WALK_MC	0.406	3.430	-1.149	-30.477	-0.746	-20.852

8 Conclusions

This study highlights the obstacles inherent in the calibration and validation of travel demand models (TDMs), particularly for smaller urban areas with limited resources. Conventional methodologies that require the collection of primary data can often be prohibitively expensive, and have led planners to seek alternative strategies, such as leveraging datasets from comparable jurisdictions to estimate model parameters and/or adopting parameters from established literature values. The transferred model parameters can subsequently be used to generate a pseudo-HTS for the target region using microsimulation methods (e.g. Greaves, 2000; Stopher et al., 2003; Greaves, 2006). In general, findings from these and other related studies in the broader literature on the transferability of travel demand models indicate that the method is sufficiently accurate in most cases at generating travel behaviours for the target region. However, the method can only generate behaviours specified explicitly as dependent variables in the travel demand models. For example, if the mode choice sub-model excludes modes such as taxis, ridesharing and e-scooters, then the model cannot predict these mode choice behaviours, and consequently, use of these modes is not captured by a pseudo-HTS generated using this method.

To address these challenges, our research introduces a novel non-parametric technique for transferring individual trips between HTSs. We hypothesise that a significant proportion of variance in travel behaviour decisions within an urban area can be explained by the relative location of people and jobs within that area. For example, imagine an individual living in a low-density and sparsely populated suburb in Tasmania, 10 km away from the Hobart CBD, who commutes to the CBD every day by car. If this individual were transferred to the Perth metropolitan area, such that they lived roughly 10 km away from the Perth CBD, our non-parametric framework would predict that the individual commutes to the Perth CBD every day by car. Travel behaviours observed for other individuals belonging to other jurisdictions can similarly be transferred to the target region, and the collection of trips thus transferred can be treated as a ‘pseudo-HTS’ for the target region.

This is admittedly a naïve assumption that does not control for differences in other mitigating factors between the source and target regions, such as public transport network, road congestion and climate. In principle, one could make the non-parametric framework more sophisticated to account for differences in each of these variables. However, that might necessitate using a parametric framework of sorts similar to the original approach used by, among others, Stopher et al. (2003), as a parsimonious solution to the problem of how to incorporate a large number of explanatory variables. The benefit of the non-parametric framework is that the transferred behaviours are not constrained by model assumptions. If the individual in our example above took a taxi to work instead in Hobart, then they would be assumed to take a taxi to work if they were living in Perth at a comparable distance away from the CBD as they were in Hobart.

We applied the framework to generate a pseudo-HTS for Greater Perth (GP), using HTS data collected in other jurisdictions. We compare inferences from the inferred pseudo-HTS with those from the actual HTS conducted in GP to validate the methodology. We use GP as our target region instead of GA because recent HTS data is available for the former (which is used to validate the methodology).

For TDM sub-models that are estimated at a trip or individual-level, our approach appears to perform reasonably well. For example, for the trip production sub-model, the pseudo-HTS tends to exhibit slight variations in constant terms and estimated coefficients for age groups and car ownership levels between the original WA HTS and the pseudo-HTS. Despite these quantitative differences suggesting potential overestimation at the trip level and underestimation at the tour level in the transferred HTS data, they remain relatively small, indicating the viability of the pseudo-HTS as an alternative data source for trip generation sub-models. Variations are also observed in coefficients of the trip distribution sub-model across the datasets. Again, despite these variations, the estimated parameters are quantitatively comparable across the datasets, suggesting consistent relationships between generalized travel cost and trip numbers. Next, mode choice model reveal minor variations in coefficient magnitudes across variables among the three datasets, also suggesting that the pseudo-HTS confidently predicts the utility of transport mode in the target jurisdiction.

However, sub-models estimated at a more aggregated level may potentially be biased due to limitations inherent in our approach. Firstly, transferring trips at the individual level without carefully considering the relationship between zones could result in over-matching for certain zones and under-matching for others. Secondly, this report relies solely on one measure of similarity – Euclidean distance, which may diminish the importance of distance measures while overemphasizing zonal characteristics such as population and employment. Lastly, our approach is constrained to using weights to assess the impact of each trip and tour made by individuals, as it is challenging to simultaneously incorporate varying weights for persons, trips, tours, zones, and similarities.

9 APPENDIXES

Table 14. Count of trips attracted to each of Greater Perth zones from the three HTSs.

Perth zone	Original WA HTS		Pseudo-HTS trip level		Pseudo-HTS tour level	
	N	%	N	%	N	%
Alexander Heights - Koondoola	48	0.17%	275	0.22%	314	0.26%
Alkimos - Eglinton	42	0.15%	57	0.04%	62	0.05%
Applecross - Ardross	271	0.94%	1698	1.34%	1626	1.32%
Armadale - Wungong - Brookdale	304	1.06%	159	0.13%	243	0.20%
Balcatta - Hamersley	242	0.84%	642	0.51%	813	0.66%
Baldivis	475	1.65%	390	0.31%	712	0.58%
Balga - Mirrabooka	172	0.60%	431	0.34%	622	0.51%
Ballajura	54	0.19%	757	0.60%	603	0.49%
Banjup	182	0.63%	420	0.33%	601	0.49%
Bassendean - Eden Hill - Ashfield	126	0.44%	527	0.42%	798	0.65%
Bateman	35	0.12%	3518	2.77%	3493	2.84%
Bayswater - Embleton - Bedford	283	0.99%	1316	1.04%	2142	1.74%
Beckenham - Kenwick - Langford	224	0.78%	1125	0.89%	503	0.41%
Beechboro	252	0.88%	1318	1.04%	487	0.40%
Beeliar - Wattleup	53	0.18%	400	0.32%	413	0.34%
Belmont - Ascot - Redcliffe	99	0.34%	1623	1.28%	238	0.19%
Bentley - Wilson - St James	255	0.89%	1209	0.95%	1326	1.08%
Bertram - Wellard (West)	148	0.52%	107	0.08%	135	0.11%
Bibra Industrial	83	0.29%	88	0.07%	133	0.11%
Bibra Lake	9	0.03%	162	0.13%	150	0.12%
Bicton - Palmyra	470	1.64%	530	0.42%	491	0.40%
Booragoon	191	0.66%	798	0.63%	1241	1.01%
Bull Creek	117	0.41%	1051	0.83%	1354	1.10%
Bullsbrook	45	0.16%	109	0.09%	156	0.13%
Butler - Merriwa - Ridgewood	188	0.65%	234	0.18%	330	0.27%
Byford	128	0.45%	232	0.18%	283	0.23%
Calista	89	0.31%	184	0.14%	288	0.23%
Camillo - Champion Lakes	122	0.42%	697	0.55%	881	0.72%
Canning Vale - East	235	0.82%	465	0.37%	747	0.61%
Canning Vale - West	165	0.57%	557	0.44%	724	0.59%
Canning Vale Commercial	105	0.37%	121	0.10%	193	0.16%
Cannington - Queens Park	234	0.81%	858	0.68%	738	0.60%
Carabooda - Pinjar	7	0.02%	14	0.01%	5	0.00%
Carramar	308	1.07%	270	0.21%	740	0.60%
Casuarina - Wandi	65	0.23%	525	0.41%	601	0.49%

Chidlow	9	0.03%	164	0.13%	171	0.14%
City Beach	35	0.12%	1750	1.38%	2076	1.69%
Claremont (WA)	217	0.76%	2460	1.94%	3029	2.46%
Clarkson	303	1.05%	261	0.21%	451	0.37%
Como	205	0.71%	874	0.69%	1069	0.87%
Coogee	59	0.21%	296	0.23%	328	0.27%
Coolbellup	144	0.50%	424	0.33%	502	0.41%
Cooloongup	168	0.58%	340	0.27%	735	0.60%
Cottesloe	282	0.98%	1295	1.02%	1078	0.88%
Craigie - Beldon	92	0.32%	286	0.23%	259	0.21%
Currambine - Kinross	157	0.55%	185	0.15%	264	0.21%
Dawesville - Bouvard	100	0.35%	21	0.02%	30	0.02%
Dianella	308	1.07%	614	0.48%	869	0.71%
Duncraig	279	0.97%	920	0.72%	1345	1.09%
East Fremantle	131	0.46%	817	0.64%	1202	0.98%
East Victoria Park - Carlisle	281	0.98%	505	0.40%	718	0.58%
Ellenbrook	566	1.97%	408	0.32%	60	0.05%
Falcon - Wannanup	139	0.48%	88	0.07%	140	0.11%
Floreat	121	0.42%	2781	2.19%	1138	0.93%
Forrestdale - Harrisdale - Piara Waters	175	0.61%	248	0.20%	390	0.32%
Forrestfield - Wattle Grove	211	0.73%	215	0.17%	261	0.21%
Fremantle	431	1.50%	21	0.02%	22	0.02%
Fremantle - South	356	1.24%	344	0.27%	471	0.38%
Gidgegannup	114	0.40%	204	0.16%	233	0.19%
Girrawheen	52	0.18%	556	0.44%	688	0.56%
Glen Forrest - Darlington	58	0.20%	538	0.42%	609	0.50%
Gosnells	151	0.53%	446	0.35%	581	0.47%
Greenfields	139	0.48%	241	0.19%	196	0.16%
Greenwood - Warwick	116	0.40%	509	0.40%	607	0.49%
Halls Head - Erskine	236	0.82%	209	0.16%	469	0.38%
Hamilton Hill	82	0.29%	302	0.24%	393	0.32%
Hazelmere - Guildford	57	0.20%	1122	0.88%	1186	0.97%
Heathridge - Connolly	24	0.08%	105	0.08%	214	0.17%
Helena Valley - Koongamia	15	0.05%	1592	1.25%	1801	1.47%
Henderson	28	0.10%	115	0.09%	133	0.11%
Herdsman	5	0.02%	104	0.08%	73	0.06%
High Wycombe	164	0.57%	290	0.23%	495	0.40%
Hillarys	94	0.33%	467	0.37%	472	0.38%
Hope Valley - Postans	2	0.01%	41	0.03%	45	0.04%
Huntingdale - Southern River	205	0.71%	1563	1.23%	483	0.39%
Iluka - Burns Beach	42	0.15%	157	0.12%	201	0.16%
Innaloo - Doubleview	228	0.79%	1007	0.79%	968	0.79%
Jandakot	108	0.38%	1499	1.18%	1651	1.34%
Joondalup - Edgewater	345	1.20%	250	0.20%	298	0.24%

Kalamunda - Maida Vale - Gooseberry Hill	147	0.51%	373	0.29%	394	0.32%
Karrinyup - Gwelup - Carine	391	1.36%	724	0.57%	1214	0.99%
Kelmscott	188	0.65%	269	0.21%	326	0.27%
Kewdale Commercial	26	0.09%	172	0.14%	121	0.10%
Kings Park (WA)	33	0.11%	131	0.10%	127	0.10%
Kingsley	59	0.21%	1836	1.45%	368	0.30%
Kwinana Industrial	41	0.14%	121	0.10%	89	0.07%
Leeming	50	0.17%	327	0.26%	482	0.39%
Lesmurdie - Bickley - Carmel	236	0.82%	146	0.12%	282	0.23%
Lockridge - Kiara	130	0.45%	668	0.53%	950	0.77%
Maddington - Orange Grove - Martin	331	1.15%	199	0.16%	285	0.23%
Madeley - Darch - Landsdale	216	0.75%	1793	1.41%	2230	1.81%
Malaga	89	0.31%	112	0.09%	113	0.09%
Malmalling - Reservoir	2	0.01%		0.00%	1	0.00%
Mandurah	352	1.23%	1033	0.81%	622	0.51%
Mandurah - East	50	0.17%	324	0.26%	431	0.35%
Mandurah - North	174	0.61%	145	0.11%	194	0.16%
Mandurah - South	48	0.17%	281	0.22%	403	0.33%
Manning - Waterford	69	0.24%	989	0.78%	1282	1.04%
Marangaroo	42	0.15%	430	0.34%	425	0.35%
Maylands	160	0.56%	853	0.67%	1153	0.94%
Melaleuca - Lexia	1	0.00%	16	0.01%	16	0.01%
Melville	286	1.00%	465	0.37%	683	0.56%
Middle Swan - Herne Hill	111	0.39%	1974	1.56%	2421	1.97%
Midland - Guildford	340	1.18%	4932	3.89%	5938	4.83%
Mindarie - Quinns Rocks - Jindalee	394	1.37%	293	0.23%	739	0.60%
Morley	368	1.28%	3126	2.46%	1850	1.51%
Mosman Park - Peppermint Grove	87	0.30%	645	0.51%	877	0.71%
Mount Hawthorn - Leederville	173	0.60%	1426	1.12%	1718	1.40%
Mount Lawley - Inglewood	311	1.08%	480	0.38%	691	0.56%
Mount Nasura - Mount Richon - Bedfordale	77	0.27%	235	0.19%	305	0.25%
Mullaloo - Kallaroo	82	0.29%	1996	1.57%	427	0.35%
Mundaring	333	1.16%	144	0.11%	212	0.17%
Mundijong	165	0.57%	254	0.20%	414	0.34%
Murdoch - Kardinya	305	1.06%	777	0.61%	865	0.70%
Nedlands - Dalkeith - Crawley	487	1.70%	445	0.35%	603	0.49%
Neerabup National Park	3	0.01%	11	0.01%	10	0.01%
Nollamara - Westminster	88	0.31%	672	0.53%	957	0.78%
Noranda	46	0.16%	680	0.54%	810	0.66%
North Coogee	57	0.20%	1107	0.87%	1207	0.98%
North Perth	57	0.20%	926	0.73%	1353	1.10%
Ocean Reef	67	0.23%	679	0.54%	456	0.37%

O'Connor (WA)	48	0.17%	252	0.20%	193	0.16%
Osborne Park Industrial	194	0.68%	789	0.62%	851	0.69%
Padbury	109	0.38%	392	0.31%	530	0.43%
Parkwood - Ferndale - Lynwood	147	0.51%	220	0.17%	308	0.25%
Parmelia - Orelia	74	0.26%	264	0.21%	431	0.35%
Perth Airport	136	0.47%	244	0.19%	257	0.21%
Perth City	1526	5.31%	1958	1.54%	2564	2.09%
Pinjarra	55	0.19%	53	0.04%	97	0.08%
Port Kennedy	95	0.33%	106	0.08%	170	0.14%
Riverton - Shelley - Rossmoyne	199	0.69%	819	0.65%	530	0.43%
Rivervale - Kewdale - Cloverdale	364	1.27%	1758	1.39%	2319	1.89%
Rockingham	341	1.19%	98	0.08%	132	0.11%
Rockingham Lakes	11	0.04%	9	0.01%	7	0.01%
Roleystone	45	0.16%	167	0.13%	224	0.18%
Safety Bay - Shoalwater	114	0.40%	134	0.11%	109	0.09%
Scarborough	176	0.61%	3520	2.77%	1137	0.93%
Serpentine - Jarrahdale	66	0.23%	73	0.06%	87	0.07%
Seville Grove	60	0.21%	276	0.22%	323	0.26%
Singleton - Golden Bay - Secret Harbour	84	0.29%	199	0.16%	249	0.20%
Sorrento - Marmion	81	0.28%	411	0.32%	587	0.48%
South Lake - Cockburn Central	200	0.70%	1706	1.34%	2355	1.92%
South Perth - Kensington	262	0.91%	1166	0.92%	363	0.30%
Spearwood	192	0.67%	2029	1.60%	427	0.35%
Stirling - Osborne Park	183	0.64%	1053	0.83%	365	0.30%
Stratton - Jane Brook	106	0.37%	591	0.47%	739	0.60%
Subiaco - Shenton Park	591	2.06%	669	0.53%	398	0.32%
Success - Hammond Park	109	0.38%	687	0.54%	1300	1.06%
Swan View - Greenmount - Midvale	158	0.55%	1258	0.99%	1598	1.30%
Swanbourne - Mount Claremont	228	0.79%	2266	1.79%	718	0.58%
Tapping - Ashby - Sinagra	42	0.15%	200	0.16%	272	0.22%
The Vines	94	0.33%	197	0.16%	285	0.23%
Thornlie	350	1.22%	1212	0.96%	563	0.46%
Trigg - North Beach - Watermans Bay	88	0.31%	339	0.27%	506	0.41%
Tuart Hill - Joondanna	123	0.43%	839	0.66%	870	0.71%
Two Rocks	78	0.27%	24	0.02%	124	0.10%
Victoria Park - Lathlain - Burswood	223	0.78%	462	0.36%	340	0.28%
Waikiki	229	0.80%	470	0.37%	174	0.14%
Wanneroo	163	0.57%	2580	2.03%	1089	0.89%
Warnbro	66	0.23%	89	0.07%	172	0.14%
Welshpool	95	0.33%	99	0.08%	95	0.08%

Wembley - West Leederville - Glendalough	405	1.41%	1066	0.84%	1415	1.15%
Wembley Downs - Churchlands - Woodlands	214	0.74%	1439	1.13%	905	0.74%
Willagee	97	0.34%	3331	2.62%	2878	2.34%
Willetton	473	1.65%	811	0.64%	1206	0.98%
Winthrop	21	0.07%	5593	4.41%	2949	2.40%
Woodvale	196	0.68%	1244	0.98%	607	0.49%
Yanchep	127	0.44%	29	0.02%	36	0.03%
Yangebup	15	0.05%	691	0.54%	928	0.76%
Yokine - Coolbinia - Menora	142	0.49%	1305	1.03%	1470	1.20%
Total	**28,727**		**126,897**		**122,886**	

Table 15. Number of zones in Greater Perth after transferring zones in other jurisdictions using only population and employment.

Matched zone	GM	SEQ	TAS	ACT	Total
Two Rocks	2	2	2	22	28
Bateman	4	8	4	5	21
Morley	13	2			15
Woodvale	7	7			14
City Beach	4	9	1		14
Claremont (WA)	4	7		2	13
Midland - Guildford	4	7	1		12
Serpentine - Jarrahdale	1	5	1	5	12
South Lake - Cockburn Central	7	4			11
Success - Hammond Park	5	5		1	11
Mundijong	1	6	1	3	11
Rivervale - Kewdale - Cloverdale	8	2			10
Middle Swan - Herne Hill	2	6	2		10
Mandurah	5	4			9
Willagee	1	3	2	3	9
Madeley - Darch - Landsdale	7	1			8
Halls Head - Erskine	7	1			8
Karrinyup - Gwelup - Carine	7	1			8
Cottesloe	5	2	1		8
Helena Valley - Koongamia	2	1		5	8
Bayswater - Embleton - Bedford	7				7
Bentley - Wilson - St James	7				7
Willetton	6	1			7
Sorrento - Marmion	4	3			7
Swan View - Greenmount - Midvale	3	3	1		7
Cooloongup	2	4	1		7
Chidlow	1	4	1	1	7
Mandurah - East	1	3		3	7
Jandakot	1	1		5	7
Tapping - Ashby - Sinagra	5	1			6
Gosnells	5	1			6
Yokine - Coolbinia - Menora	4	2			6
Carramar	4	2			6
Calista	3	2	1		6
Winthrop	1	3	1	1	6
Trigg - North Beach - Watermans Bay	1	3	1	1	6
Bullsbrook	1	3		2	6
Bassendean - Eden Hill - Ashfield	4	1			5
North Perth	4	1			5

Scarborough	3	2			5
Mindarie - Quinns Rocks - Jindalee	3	2			5
Clarkson	3	2			5
Greenwood - Warwick	3	2			5
O'Connor (WA)	3			2	5
Parmelia - Orelia	2	3			5
Mount Hawthorn - Leederville	1	4			5
Innaloo - Doubleview	4				4
Canning Vale - West	3		1		4
Maddington - Orange Grove - Martin	3	1			4
Manning - Waterford	3	1			4
Wembley Downs - Churchlands - Woodlands	3	1			4
Canning Vale - East	3	1			4
Dianella	3	1			4
High Wycombe	3	1			4
East Victoria Park - Carlisle	3	1			4
Kingsley	2	2			4
Falcon - Wannanup	2	2			4
Duncraig	2	2			4
Warnbro	2	2			4
Lesmurdie - Bickley - Carmel	2	2			4
Booragoon	2	2			4
Como	2	2			4
Mount Nasura - Mount Richon - Bedfordale	1	3			4
Melville	1	3			4
Hope Valley - Postans	1			3	4
Butler - Merriwa - Ridgewood	3				3
Port Kennedy	3				3
Nedlands - Dalkeith - Crawley	3				3
Wanneroo	3				3
Hillarys	3				3
Balga - Mirrabooka	3				3
Cannington - Queens Park	3				3
Rockingham	3				3
Mandurah - South	2		1		3
Huntingdale - Southern River	2	1			3
Tuart Hill - Joondanna	2	1			3
Waikiki	2	1			3
Fremantle - South	2	1			3
Maylands	2	1			3
Swanbourne - Mount Claremont	1	2			3
Greenfields	1	2			3
Girrawheen	1	2			3

Safety Bay - Shoalwater	1	2			3
Roleystone	1	2			3
Mullaloo - Kallaroo	1	2			3
Osborne Park Industrial	1	1		1	3
Perth Airport	1	1		1	3
Forrestdale - Harrisdale - Piara Waters	2				2
Riverton - Shelley - Rossmoyne	2				2
Thornlie	2				2
Alexander Heights - Koondoola	2				2
Armadale - Wungong - Brookdale	2				2
Padbury	2				2
Beechboro	2				2
Victoria Park - Lathlain - Burswood	2				2
Nollamara - Westminster	2				2
Bull Creek	2				2
Spearwood	2				2
Parkwood - Ferndale - Lynwood	2				2
Subiaco - Shenton Park	2				2
Mandurah - North	1	1			2
Perth City	1	1			2
Wembley - West Leederville - Glendalough	1	1			2
Forrestfield - Wattle Grove	1	1			2
Bicton - Palmyra	1	1			2
Byford	1	1			2
Murdoch - Kardinya	1	1			2
Singleton - Golden Bay - Secret Harbour	1				1
Hamilton Hill	1				1
Belmont - Ascot - Redcliffe	1				1
Bertram - Wellard (West)	1				1
Ellenbrook	1				1
Joondalup - Edgewater	1				1
Malaga	1				1
Balcatta - Hamersley	1				1
Mount Lawley - Inglewood	1				1

10 REFERENCES

Cambridge Systematics (2012). Travel demand forecasting: Parameters and techniques, Transportation Research Board.

Greaves, S. P. (2000). Simulating household travel survey data in metropolitan areas. Louisiana State University and Agricultural & Mechanical College.

Greaves, S. P. (2006). Simulating household travel survey data. In Travel Survey Methods: Quality and Future Directions (pp. 625-650). Emerald Group Publishing Limited.

Stopher, P. R., Bullock, P., Rose, J. M., & Pointer, G. (2003). Simulating household travel survey data in Australia: Adelaide case study. Road & Transport Research Journal, 12(3), 29-44.

Stopher, P. R., Greaves, S., & Bullock, P. (2003). Simulating household travel survey data: application to two urban areas. 82nd Annual Meeting of the Transportation Research Board, Washington, DC,