Proceedings of the Eastern Asia Society for Transportation Studies, Vol.11,2017

Investigation of the Factors Affecting Mode Choice Preferences of Public Transport Users: A Case Study of Main North Road, Adelaide

Munshi NAWAZ [a], Sekhar SOMENAHALLI [b], Andrew ALLAN [c]

[a,b] *School of Natural and Built Environments, University of South Australia, Mawson Lakes, SA 5095.*

[a] *Email: munshi.nawaz@mymail.unisa.edu.au*

[b] *Email: sekhar.somenahalli@unisa.edu.au*

[c] *School of Art, Architecture and Design, University of South Australia, Adelaide, SA 5000.*

[c] *Email: Andrew.Allan@unisa.edu.au*

Abstract: In the context Australian cities, an integrated transport and land use planning approach is considered essential to serve the outer low-medium density suburban communities. The concept of bus based transit oriented developments (BTODs) offers a similar approach, where bus based transport routes can be transformed into bus rapid transit (BRT) corridors and can coordinate land use planning for BTODs alongside.

This paper investigates an existing transit passenger survey dataset for a selected corridor in metropolitan Adelaide, which examines the demographic, social and activity based data in understanding their level of influence in determining passengers' mode choices for public transport. A range of descriptive and statistical analysis underpinned this study, which helps understanding a number of independent variables that are perceived to have strong influence on public transport patronage and usage frequencies. The statistical analysis has identified some categorical variables, including the 'the purpose of journey' as a significant variable, which can help in formulating policies that aim at achieving an increased number of public transport patronage.

Keywords: Transit oriented development (TOD), Bus rapid transit (BRT), Multinomial logistic regression, Binomial logistic regression, Multiple Imputation.

1. INTRODUCTION

Urban sprawl is considered to be one of the key reasons behind the significant increase in private car ownership. During the period when cars were becoming more affordable to the general public, up until the 1960s, cities around the world developed to cater to cars as a solution to cities' transportation needs, which led to the development of planning policies that directly and indirectly facilitated sprawl. Public transportation was given less of a focus in overall town planning designs and has not coped with the pace of the rapid developments in the ever expanding outer suburbs, away from the places of regular community services and facilities. An alternative approach to reducing dependency on private vehicles would be to offer policies and strategies that result in the formulation of sustainable transport and land use plans. The concept of Transit Oriented Development (TOD) evolved from a principle that can be explained as the design of a mixed use relatively higher density development emphasizing pedestrian accessibility to a major transit route and reinforcing the use of public transportation (Parker, 2002).

Except for the cities of Sydney and Melbourne, the outer areas of all other Australian metropolitan and regional cities and towns, share a common characteristic of low density residential areas that have expanded outwards horizontally. The rail based transit corridors are considered to be the primary features of Sydney and Melbourne's transport network, while bus services are found to form the dominant features of the public transport networks of Brisbane, Perth and Adelaide. Bus based transit oriented developments (BTODs) can, therefore, be a potential alternative to RTODs in these cities, particularly to service the outer suburban areas. The existing bus routes can be effectively redesigned (i.e. aligned with the Bus Rapid Transit-BRT principles) to provide rail-like services, where although vehicle speed may not be same as rail (subject to specific designs etc.), frequencies can be similar. Buses can also provide the flexibility in serving car dependent outer communities, where rail may not be a viable option from planning, economic or political considerations. Adam et al. (2005) have analysed built environment characteristics and travel behaviour in support of bus based transit oriented development to examine the functionality and performance of such services and have suggested a number of design recommendations (i.e. regarding infrastructure design, service quality and frequency, etc.) to achieve increased public accessibility.

The concept of bus based transit oriented developments (BTODs) along developed bus transit corridors was considered by Cervero (2000) to be more viable options for the lower density outer suburban areas, which are located beyond the catchments of the existing rail corridors and are currently serviced only by conventional public bus services. Later, Cervero and Dai (2014) have closely looked into the concepts of TODs and examined how existing BRTs can be converted into transit corridors supporting TODs. Their work considered the disadvantages of a bus based TOD (i.e. low density, low speed and poor service frequency, lack of magnitude, unknown market implications, etc.), as identified by Currie (2005) and discussed relevant solutions in their paper (Cervero and Dai, 2014). Cervero and Dai (2014) developed design principles necessary to attain the required density and accessibility and argued that these could be implemented effectively in the context of low density suburban areas of Australia.

Ho and Mulley (2014) also suggested that traditional bus rapid transit (BRT) style bus services are essential for the fringe areas of metropolitan cities, where access to direct public transport routes (to the inner cities) are limited, which can be considered as one of the key causes that these communities became dependent on private cars. Taylor et al. (2011) have identified a number of social and qualitative attributes that play a direct role in determining the success of such a service. The level of patronage is one of the key determinants of the success of a service, which is mostly dependent upon the qualities of the service such as: pedestrian accessibility to the stations/ bus stops; properly connected networks of pedestrian and cycling pathways; safety and security locally and when using the transits; and the frequency of services during the weekdays and in weekends, etc.

In light of the above, this paper aims to conduct an investigation of a bus corridor, which has been selected as a route to become potentially transformed into a BRT corridor to support transit oriented developments in adjacent areas. The focus of this study was on the identification of attributes of the existing bus services along the corridor that may or may not influence existing passengers' choice of public transportation. An existing set of data from the Public Transport Passenger Survey 2009 (for Metropolitan Adelaide) was analyzed to achieve an understanding of the attributes and variables from the responses received from existing transit passengers. The outcome of the analysis is expected to establish a link between a list of preferential variables and the choice that the respondents made in selecting public transport services for their journeys. The relationship pattern of the variables is expected to help future designing of the transit corridor (in support of TODs) resulting in attracting increased public transport patronage.

2. THE CONTEXT OF THE STUDY

As observed in metropolitan Adelaide, the majority of the outer suburban areas are located at least 1km away from the existing rail network corridors. This results in compromising accessibility (i.e. in relation to increased walking distances, increased travel time, reduced mode choice option etc.) of general public to a mass transit (rail) services. Although Adelaide has extensive bus networks covering the majority of these areas (The Government of South Australia, 2014), the existing bus services were found to have either deteriorated in quality or have not expanded with the pace of the population growth (Currie, 2014) and thus have failed to attract people to shift away from their usual dependency on private cars to public transportation services.

In the context of metropolitan Adelaide, the 30 Year Plan for Greater Adelaide (30YPGA) 2010 aimed to create 14 new TODs and 20 more sites that are proposed to be developed on the basis of TOD concept and design characteristics. The Plan targets that 60% of all new housing in Metropolitan Adelaide (over the next 30 years) will be located within 800m of transit corridors (TC), however, the plan fails to provide policy directions for the 80% of the existing metropolitan built up areas, which will remain largely unchanged as a result of the Plan (Department of Planning and Local Government - South Australia). The Integrated Transport and Land Use Plan (ITLUP) 2015, had focused on the coordination of the various components of the public transportation system having regard to the existing accessibility and transport disadvantage experienced in the outer metropolitan areas. Considering the fact that the existing bus network supports around 80% of public transport journeys in Adelaide, the Plan has incorporated a number of priority actions in regard to bus based services, which are generally aligned with the basic principles of transit oriented development and with the key goal of the 30YPGA (The Government of South Australia, 2015).

Cervero and Dai (2014) indicated in their paper that there is potential for existing BRTs to be converted into transit corridors supporting TODs. The increased frequency of services is one of the attributes that BRT offers, which can be linked to an increased number of patronage as Rose et al. (2013) have found that travellers would be prepared to walk additional distances if bus frequencies were improved. The Adelaide OBahn, as a form of BRT, was considered as a great success by many for its unique infrastructure and its ability to attract more trips to and from Adelaide CBD (Bray and Scrafton, 2000, Currie, 2006). It was also regarded as a highly significant infrastructure in the Planning Strategy of Metropolitan Adelaide, which laid out a strategy for the future to develop a mass transit corridor from the Adelaide CBD to Salisbury via OBahn (Allan and Fielke, 2015). The OBahn also has the potential to facilitate the transformation of the corridor adjacent areas into TODs; particularly at the locations of the Klemzig and Paradise interchanges (Allan et al., 2015), which is aligned with suggestions made by Cervero and Dai (2014) as mentioned earlier. However, it needs to be acknowledged that a service/ infrastructure similar to OBahn would be very unlikely to be replicated elsewhere within the metropolitan Adelaide area because of the lack of similar settings and for the lack of political willingness for such a project. It was therefore considered very important that conventional bus routes are evaluated and are selected to be transformed as BRT corridors that can not only be developed to cater for rail-like high frequency and high speed reliable services to communities but also that has the potential to be linked within appropriate land use planning practices.

3. METHODOLOGY

3.1 Selection of the Study Area and Study Route

This study, as part of a broader research, identified north-eastern parts of the metropolitan Adelaide as a very important region from a major transport connection point of view. A number of major roads (i.e. Churchill Road, Port Wakefield Road, Main North Road, North-East Road, etc.), railways (i.e. Adelaide's northern rail corridor connecting Gawler) and a dedicated bus corridor (i.e. the OBahn) are situated within this region. As part of the research, five of the key bus routes and roadways within this region were considered as initial candidate corridors. A two-step process was undertaken to compare the selected candidate corridors' qualities to be transformed into BRT corridors. Firstly, the candidate corridors were independently analysed based on nine selected parameters/ variables to determine the suitability scores of the initial corridors and secondly, a multi criterial decision analysis was undertaken to determine the level of significance (in terms of relative weights on a scale of 0 to 1) of these parameters. The collective results of the whole process revealed that 'Main North Road' would be the most suitable corridor for transformation into a transit corridor for the BTOD model. Main North Road was hence selected as the case study corridor for this research.

Starting from North Adelaide, the first 21km length of Main North Road was chosen as a study route for this analysis as this segment services a number of neighbourhoods, including the far northern suburbs of Para Hill West, Salisbury South, Salisbury East, Elizabeth, Elizabeth North, Munno Para, etc. Residential use was found to be dominant within 1.0km of the roadways although, along with the road frontages, commercial, retail and services predominate. The population and employment density were identified to be the highest within the road segment between Nottage Terrace and Grand Junction Road, while densities appear to be dispersed much more weakly along the northern segment of the corridor.

This paper looks into this route specific transport passengers' personal information and preferential selections, collected as part of the 2009 survey, to understand their mode choice considerations.

3.2 Collection and Re-organisation of Relevant Data

The original Public Transport Customer Survey 2009 contains approximately 10,884 passenger data records. The sampling of the survey was done based on the public transport services (Somenahalli et al., 2013). The study corridor, Main North Road, is one of the busiest bus corridors within metropolitan Adelaide and is presently used by a number of bus routes to different destinations. For the purpose of developing a dataset for this study, all bus services were identified that use Main North Road either along the entire length or part of it and were separated from the primary dataset (i.e. from the Survey 2009). As a result, data records of a total of 213 respondents were selected for the study dataset. Although the dataset is around 7 years old, the characteristics and the service length of the corridor has not changed very much during this period. Therefore, the data collected in 2009 is considered to be still valid for the corridor. As part of a broader research, it is also intended to undertake a similar transit passenger survey for the corridor (i.e. Main North Road), which is expected to further validate the outcomes of this analysis.

The original questionnaire was developed to collect a wide range of data against approximately 25 different variables. For the purpose of this analysis, data collected under those initial variables were rearranged and refined to define the scope of the expected result areas. For example, data collected under both questions 2a (i.e. "will not" need to transfer to another service) and 3e (i.e. "will" need to transfer to another service) were amalgamated to develop a variable to provide data on "the primary purpose" of the passengers' journey. After the data set was refined, a total of 13 variables were finally derived for the purpose of this analysis.

3.3 Descriptive and Statistical Analysis

The refined set of 13 variables asked questions covering a range issues based on the respondents' regular activity, personal choices, their socio-economic and demographic conditions (partly). Although the respondents were likely to have answered these question independently, it was very likely that their answers had interconnected links and influences on their selection of public buses. As part of the descriptive analysis, this paper has attempted to look into the types and extent of the different variables and to investigate one variable's potential relationship with others.

As part of the statistical analysis, appropriate statistical methods were applied to identify correlations amongst the variables. A logistic regression analysis was undertaken using the existing set of primary data and a new set of imputed data, to achieve a deep understanding of the independent variables.

The missing data management formed an integral part of the statistical analysis (and of this study), where the Multiple Imputation method was applied to impute a completely new dataset to compare results with the outcomes of the primary dataset.

4. ANALYSIS AND RESULTS

4.1 Descriptive Analysis

Within the Public Transport Customer Survey 2009 dataset, a total of 213 transit passengers were identified using the existing bus services along the Main North Road corridor. On the basis of gender distribution, 46.48% of the respondents were identified to be male and 44.13% were female, while 9.39% of this corridor specific data did not record any gender description, which can otherwise be considered as missing data. According to the age distribution, the 15-24 age group was identified to constitute the largest share of 31%, while the age Group of '15 and below' formed the lowest share of 1.41%. The total survey data 2009 identified the same age group forming the highest and lowest contributions, but at 39.8% and 2.9% respectively (Somenahalli et al., 2013). The general age distribution is demonstrated in Figure 1.

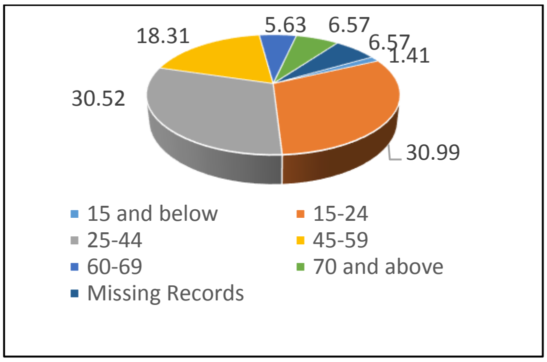

Figure 1: Respondents' Age Groups (in %)

The dataset shows that approximately 46% of the passengers use public buses every weekdays, while approximately 37% use buses from 1 to 4 days a week. Figure 2 demonstrates the passengers' usage frequency in weekdays (in blue).

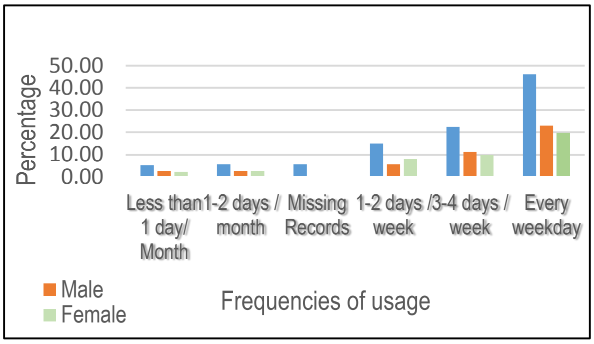

Figure 2: Usage Frequency on Weekdays

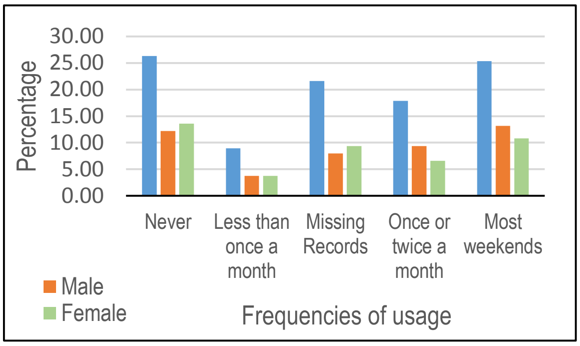

Figure 3: Usage Frequency on Weekends

The chart also compares the usage distribution based on the passengers' gender, where male passengers appear to use public buses slightly more frequently than female passengers. This is quite opposite to what the 'travel to work' data of the Census 2006 suggests, where female bus users were found to use buses more than male users (Somenahalli et al., 2013). Usage frequency during weekends appear to provide interesting results, where the percentage of those respondents who never use public buses during weekends were found (Figure 3) to be approximately the same (around 25%) as those who use it.

In considering the primary purposes of the journeys, the highest two distributions were found for the trips to work and for the trips to home, which were approximately 31% and 21% respectively. Another significant purpose was found to be education, which constitutes travels to schools and to tertiary education facilities and, was identified as a purpose of the journey by approximately 20% of the respondents. The remainder purpose types included shopping, recreation, medical and personal services etc. Refer to Figure 4 for the percentage distribution of the selected purposes.

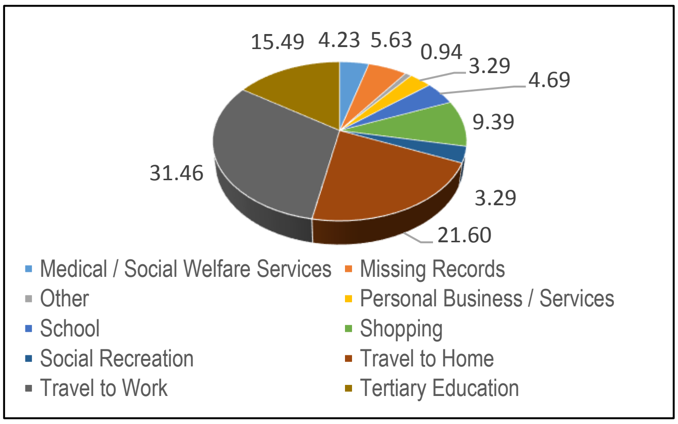

Figure 4: Primary Purpose of Journey

In using public transportation, the respondents identified a number of modes that they use to travel to the bus stops and to the destinations. In the questionnaire, 6 types of modes were listed for the trips to bus stops, while 4 modes were listed for the trips to destination. Both for the trips to bus stops and to the destinations, walking was found overwhelmingly to be the predominant mode identified by the respondents, i.e. constituting approximately 72% and 85% respectively. Refer to Figure 5 for the distribution of the types of travel modes to bus stops.

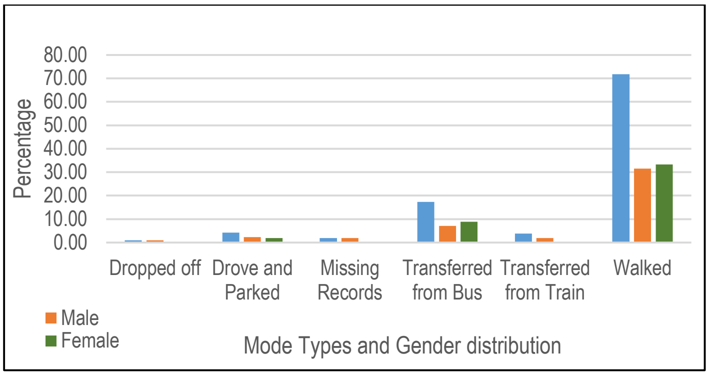

Figure 5: Used Travel Mode to Bus Stops

The nearest contribution was found by the option ‘transfer from other buses’, which constituted approximately 17%. Mode choices, based on gender, do not appear to be very significant. Female respondents were found to be slightly more inclined to walk to bus stops than their male counterparts, but walking to the destinations (from bus stops) show the opposite results.

When the respondents were asked about the reasons as to why they chose public transport, the majority (approximately 41%) of them mentioned more convenience for choosing transport compared to any other modes. Approximately 12% identified ‘car parking costs’ as a reason for using public buses, while around 7% of the respondents identified both fuel prices and environmental reasons as their prime consideration in selecting public buses. Figure 6 represents the percentage distribution of passengers’ stated reasons for selecting public buses for their trips. Approximately 45% of the respondents mentioned that they did not have access to private vehicles for this trip, which clearly indicates that the passengers had no alternative available to them other than using public transport. By contrast, approximately 30% of the respondents mentioned that despite having access to private vehicles they chose public transport, which might be a result of many other factors, as have been discussed here, in making a well-informed mode choice decision.

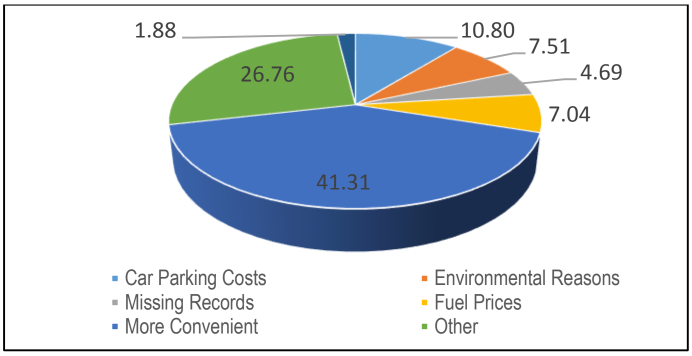

Figure 6: Stated Reasons for using Public Transport

Considering the data on concession card availability among the respondents, an interesting result was found when this dataset was compared with the frequency of public transport usage.

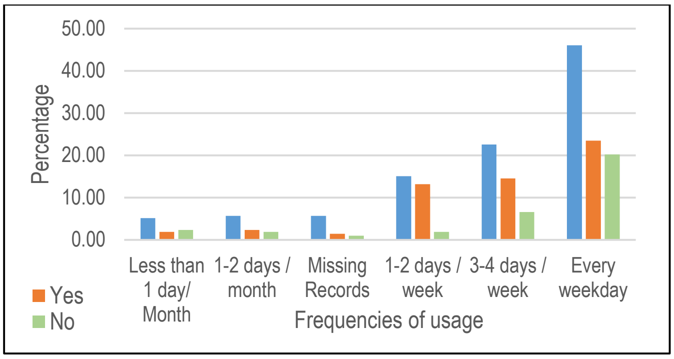

Figure 7: Relationship Between Usage Frequency on Weekdays and Concession Cards Availability

As Figure 7 shows, the frequency of usage of public transport appeared to have a relation with the availability of concession cards. The higher frequency of public transport usage appeared to have a strong relationship with the availability of the concession cards. A similar result was also found for the data available for the weekend travel frequencies. In both scenarios, however, it was found that there is not much difference of influence of having a concession card when the frequency of public transport usage is less. The result may provide a general indication that reduction in the price of public transport tickets may encourage more frequent usage, both on weekdays and on weekends.

The comparison of private car drivers' data, household car ownership data and the frequency of public transport usage provides interesting results, where all three variables appear to be interrelated.

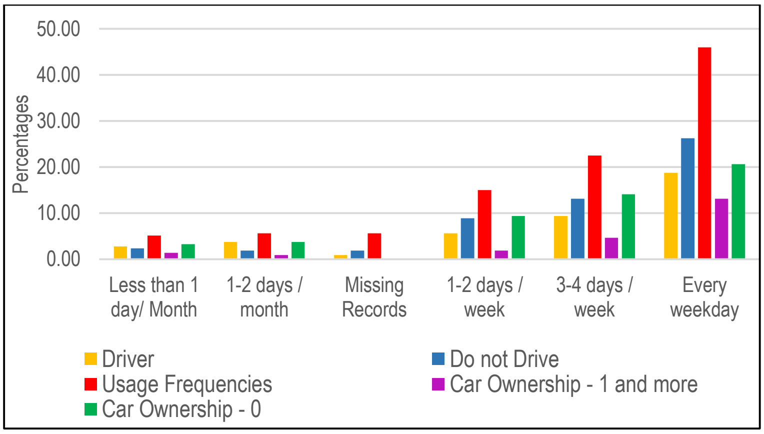

Figure 8: Relationship Between Usage Frequency on Weekdays comparing Car Drivers and Car Ownership

Figure 8 shows that not being able to drive a private vehicle appeared to have a positive relationship with a higher frequency of usage of public transport. The chart also shows that the stated household ownership of 'no' vehicle appeared to have a greater influence on the usage of public transport than the ownership of '1 & more' vehicles. Similar results were found for the dataset for weekend trips, where non-drivers and households with no car ownership tend to result in a higher frequency of weekend trips using public transport. In both scenarios, it was found that dependency on driving and higher car ownership appeared to have slightly more influence when compared against lesser frequency of usage of public transportation.

4.2 Statistical Analysis

4.2.1 Correlation

Based on the above analysis it can be assumed that there are some potential relationships among the selected variables. The analysis was primarily aimed at identification of the variables that might or might not have an impact on public transport patronage. Based on this criteria, the “weekend and weekdays usage frequencies” were considered to be the dependent variables, which were subject to change based upon their relationship with other variables. A correlation test was undertaken to understand such relationships, which was followed by a regression analysis to develop a model for future estimation of public transport usage number.

The correlation analysis was undertaken using MS Excel, where all variables were considered. The variable “ticket type” was excluded from this analysis as this was considered to be dependent itself on the already identified dependent variables. The correlation results are presented in Table 3, which appeared to have not reflected the findings of the previous analysis in terms of potential relationships among variables. The “weekday usage frequencies” was found to have a correlation coefficient of -0.26 with the variable “the purpose of the journey”, which means that those have a negative relationships of lower strength (Taylor, 1990, Miles and Shevlin, 2000), while the next relationship was found with the variable “age group” (-0.23), which is of similar strength. No other variable were found to have correlations as stronger as expected from the previous descriptive analysis. In contrast, the correlation results for the “weekend usage frequencies” were found more consistent as it negatively relates to the variables describing the respondents’ driving capabilities and the availability of a car for the trip. The results showed correlation coefficients of -0.38 and -0.27 for those variables respectively, which were considered to represent reasonable relationships among the variables (Miles and Shevlin, 2000, Taylor, 1990). The respondents’ entitlements to concession cards was found to score a coefficient of 0.23, which means that the corresponding relationship is not significant. The correlation results between independent variables also provided an understanding of the multicollinearity, when two or more independent variables are found to have strong correlation with each other (Laerd Statistics, 2017).

4.2.2 Regression

Given the majority of the primary data collected were categorical in nature, a logistic regression analysis was undertaken using SPSS in order to identify the independent variables that best fit with the dependent variables. The subject dataset (for the study corridor only), which represented a small fraction of the overall primary dataset (i.e. for public transport routes within metropolitan Adelaide), was carefully examined and it was considered that a rearrangement of the primary categories would provide a better result. The multinomial logistic regression analysis was undertaken as a first step of the study, given that the ‘weekdays usage frequencies’, as the dependent variable, had five (5) primary categories. For the convenience of interpreting the analysis outcomes the categories of the dependent variable were re-coded and were reduced to three (3) categoris as shown in Table 1 below:

Table 1: Rearranged codes for the dependent variable

Weekdays usage frequencies	
Primary Categories	Re-coded Categories
Every weekday	Every Weekdays
3-4 days / week	3-4 days per week
1-2 days / week	Less than monthly to 1-2 days a week
1-2 days / month	
Less than 1 day/ Month	

As part of a two step process for analysis, each of the independent variables was run with the dependent variable separately as a first step to identify the individual statistical results for the variables. The relevant results in the SPSS generated outputs (i.e. ‘Model Fitting Information’, ‘Goodness-of-fit’, ‘Pseudo R-square’ and ‘Liklihood Ratio Tests’) were examined and compared to determine the individual independent variables that generated an acceptable level of significance (*p*-value of 0.05 or less). The ‘Primary purpose of Journey’, ‘reason for catching transport’ (partly), ‘concession card entitlements (partly)’ and ‘no household car ownership’ were identified to be the four (4) of the ten (10) independent variables, which were found to have the statistical significance needed to individually explain the predictions about the passengers’ public transport usage frequencies during weekdays.

As part of the second step, when the logistic regression was undertaken with a number of different combinations of these shortlisted independent variables, the combined results did not reflect the outcomes of the models with single independent variables. The best results were obtained for the combination of the variables ‘primary purpose of journey’ and the ‘concession card entitlements’ based on the model fitness tests and the statistical indicators. The ‘purpose of the journey’ performed well in explaining the likelihood of the higher usage frequencies for weekdays. The categories of ‘journey to home’ and ‘journey to work’ results showed significance values of 0.022 and 0.019 respectively in explaining the logarithmic odds for the ‘every weekdays’ frequency, while ‘journey to work’ was the only category found for the next category of frequency. However, the ‘concession card entitlements’ did not have any relationship significance with either of the dependent variable categories (i.e. found *p*-values of 0153 and 0.549 for the dependent categories of ‘Every Weekdays’ and ‘3-4 days per week’ respectively), whereas the variable showed a significace level of 0.02 when it performed independently.

Although the dependent variable was designed to contain the nominal variables as categories, the gradual nature of the frequencies could also be considered as ordinal variables. The categories of the key dependent variable, ‘weekdays usage frequencies’ hence were further re-coded and were reduced to two categories, i.e. category 1 was representing the passengers’ usage frequencies of ‘1-5 weekdays’, while the second category included data reflecting either ‘no use’ of public transport at all or ‘negligible/ very infrequent’ level of usage. Based on the newly developed categories, a further test was undertaken using the binomial logistic regression for comparing the outcomes of results with that of the multinominal logistic regression. Interestingly, the analysis produced a different result.

Similar to the two step process as used in the multinomial logistic regression, each of the independent variable were run in the regression model exclusively with the dependent variable and the results generated only two (2) independent variables meeting the statistical checklist to qualify for regression equation. The ‘primary purpose of journey’ was found to be the independent variable common in the results of both logistic analyses. The passengers’ ‘driving ability’ was identified to be a new independent variable resulted from the binomial logistic regression analysis. Unlike the previous analysis, the binary regression model also produced a combined output with consistent results, where both independent variables were found to obtain satisfactory *p*-values. The model also passed all initial fitness tests such as; ‘Omnibus Tests for Model Coefficients’ and ‘Hosmer and Lemeshow Tests’. The category ‘journey to work’ within the ‘primary purpsoe of journey’ variable was again found to be the only performer with a significance values of 0.01, while the passengers’ ‘driving ability’ was found to have a p-value of 0.013. Refer to Table 4 for the comparative analysis of the results for the multinomial and binomial logistic regression analysis.

4.3 Managing Missing Data- by Multiple Imputation

The dataset used for this study has a range of missing data, which shows that data is missing within the range 1.88% to 20% depending on the types of the variables in the dataset. To understand the current status of the missing data, the original dataset was analysed using SPSS and it was found that approximately 9.97% of data was missing from the overall dataset as shown in the Figure 9.

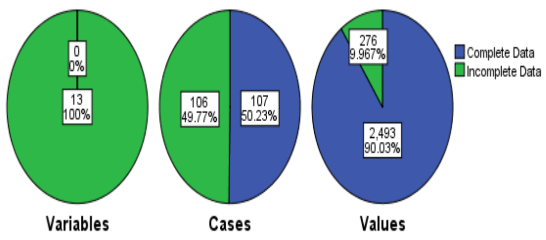

Figure 9: Description of Identified Missing Values

An extent of 10% missing data in a dataset is considered to be quite significant and such a level of missing data may bias statistical results (Schlomer et al., 2010). The experts have developed a number of methods in dealing with missing data, systematic deletion is one of them (Olinsky et al., 2003, Tsikriktsis, 2005). Considering the size of the missing data (approximately 10%) in the dataset, this study took an approach to replacing all missing data from an assumption that recovery/ replacement might provide different results than the results obtained from the original dataset. This step can also be considered as an additional method of testing the results that were already obtained from the original dataset.

There are a number of statistical methods available for replacing missing data from a dataset, which are widely recognised and used by experts based on the types and the forms of both the datasets and the outcomes. The "Expectation of Maximization – EM" and "Multiple Imputation –MI" are two of those which are widely recognised for having the ability to deal with relatively large datasets (Olinsky et al., 2003, Schlomer et al., 2010, Vriens and Melton, 2002). Apart from the missing data information that is already known for the original dataset (10%), an additional test was undertaken to understand the randomness of the missing data from the used dataset. The results of the randomness test are shown below in Table 2.

Table 2: Results of the Little's MCAR Test

EM Means[a]

Travel Mode to Stop	Travel Mode To Destination	Primary Purpose Of Journey	Weekday Usage Frequency	Weekend Usage Frequency	Reason For Catching Transport	Ticket Type	Concession Card	Private Car Driver	Car Availability for this Trip	Household Car Ownership	Age Group	Gender
3.40	2.89	3.09	4.04	2.55	3.07	1.29	.63	.43	.33	1.35	3.17	1.49

a. Little's MCAR test: Chi-Square = 475.401, DF = 403, Sig. = .007

The above 'Missing Completely At Random – MCAR' results found that the significance level (i.e. 'P-value') was less than 0.05%, which indicated that the data were missing in a systematic/ biased way (Van Ness et al., 2007). There can be a number of reasons behind a systematic missing of data. With one established reason being that of placing questions in an initial questionnaire that the respondents might have preferred not to answer to, either because they did not have a relevant answer or they might have found some of the questions personal / sensitive in nature (Schafer and Graham, 2002). Based on these two criteria (i.e. relatively a higher missing percentage and a systematic nature of the missing data), the Multiple Imputation (MI) method was considered to be the most appropriate option for the management of the study dataset (Schlomer et al., 2010, Vriens and Melton, 2002).

As all 13 variables were identified to have missing data to some extent, all variables were selected for the imputation model in SPSS, where 5 new sets of data were instructed to be produced as part of the multiple imputation process. The minimum and maximum values were carefully selected as constraints matching with the types and range of relevant attribute values of the original dataset. All relevant variables were independently selected to use the dataset as a predictor within imputing the missing values. SPSS was also used to display an imputed dataset along with the descriptive statistics of the different variables with imputed data. The process resulted in the creation of a new dataset, which displayed 5 new imputed datasets along with the original dataset for comparison.

The statistical summaries (i.e. the relative mean and standard deviations) of all new imputed datasets were carefully compared with those of the original dataset to ensure that the MI process produced a statistically acceptable outcome. After having this comparison done for all selected variables, it was found that the Imputed Dataset 1 was the closest match with the original dataset.

4.4 Analysis of the Imputed Dataset 1

A new multinomial logistic regression analysis was undertaken for the Imputed Dataset 1 following the same approach adopted for the original datasets as discussed earlier in Section 4.2. A similar two-step process was followed as was used for the analysis of the primary data. As a result of step 1, the 'primary purpose of journey' was again found to be the dominant variable, while the 'concession card entitlement' was found to have limited significance in explaining the dependent variable. The 'age group' was identified to be the third independent variable, which has a very strong significance level (p-value of 0.000) for its categorical value of 1, which is 'under 24' years.

Similar to the outcomes of the primary dataset analysis, the multinomial logistic regression of the Imputed Dataset 1 had also produced a combined output of two independent variables, where both independent variables were found to otain satisfactory *p*-values. The model also passed all initial fitness tests such as; 'Model Fitting Information', 'Goodness-of-fit', 'Pseudo R-square' and 'Liklihood Ratio Tests'. The category 'journey to work' within the 'primary purpose of journey' variable was again found to be the strong performer in both public transport usage categories, while the 'age group' of 'under 24' was found to have a strong significance value of 0.000. Table 5 represents the analysis results of the imputed dataset as generated by SPSS, which can be compared to the results of the primary datasets in Table 4.

5. DISCUSSION

Although the usage frequencies were considered separately for weekdays and weekends in the primary dataset, public transport usage on weekdays was considered to be the dominant feature of the two in shaping overall patronage for the week. The outcomes of the discussed logistic regression analysis can be interpreted as the total number of passengers, with the probability of using public buses at a higher level of frequencies, which are subject to the factors related to the independent variables such as the primary purpose of journey, age group, dependency on private vehicles as drivers and the respondents' entitlements to concession cards.

The primary purpose of journey had a list of purposes, the survey results were found strongly dominated by the work trips, which is considered to have played a key role in generating this relationship with the usage frequencies. For the ‘age group’ variable, the lower and the higher cohorts were found to be more dependent on public transportation than the other cohorts, whereas the logistic regression analysis found that the age group of ‘under 24’ had strong logarithmic odds to increase the probability of high frequency public transport usage. So, for planning future public transport design, the working population would need to be focused on while the age group of ‘under 24’ (potentially future working population) needs to be further targeted to make the services sustainable.

The relationships with the respondents’ driving ability of private vehicles and their entitlements to concession cards were found to be relatively straight forward. This can be interpreted as the lower number of respondents with driving ability may transform into a higher level of patronage. Within the state of South Australia, concession cards mean reduced fares for transportation and reduced prices for many other services, which can therefore be interpreted as resulting in an increased level of patronage.

As with the abovementioned independent variables, at least two other variables were expected to have a strong influence on the number of usage frequency variable (as had shown to have potential in the descriptive analysis section of this paper). The modes of travel to the stops or to the destination can be considered as important variables, which given collectively, could reveal travel distances to and from the stops via any of the modes identified. Another surprising result was that household car ownership was found to have no significant influence at all on the estimation of the usage frequencies. Firstly, the higher number of household private car ownership may reflect higher dependency on private cars hence, it can be expected that higher ownership may translate into a lower usage of public transport services and vice-verse.

The correlation results between independent variables also provided an understanding of the multicollinearity, when two or more independent variables are found to have strong correlation with each other. This often creates a problem of identifying the right independent variable/s to explain the nature of the dependent variable (Laerd Statistics, 2017). Table 3 shows that some of the variables have correlation with each other of varied strengths. It was very difficult to ascertain whether this relationship could have influenced the results of the regression analysis, where all these variables were tested both independently and collectively but the results did not appear as expected. For example, the variables “household car ownership’ and ‘passengers’ driving ability’ were found to have a strong correlation (coefficient of 0.62). This may have provided the discussed multicollinearity effect on the outcome of the logistic regression analysis as none of these variables were found to have enough supporting evidence in explaining the dependent variable.

Table 3: Correlation results among selected variables (from original dataset).

	Mode to Stop	Mode to Destination	Journey Purpose	Reason For Catching Transport	Concession Card	Car Driver	Car Availability For this Trip	Car Ownership Household	Age Group	Gender	Weekday Usage Frequency	Weekend Usage Frequency
Mode to Stop	1											
Mode to Destination	0.02	1										
Journey Purpose	0.03	-0.08	1									
Reason for Catching Transport	0.08	0.00	0.04	1								
Concession Card	0.15	0.03	0.30	0.06	1							
Car Driver	-0.16	-0.09	-0.17	-0.12	-0.27	1						
Car Availability for this Trip	-0.09	-0.11	-0.21	-0.22	-0.24	0.62	1					
Car Ownership Household	-0.04	-0.13	-0.24	-0.10	-0.14	0.31	0.34	1				
Age Group	-0.01	-0.06	0.05	-0.02	-0.01	-0.01	0.03	-0.25	1			
Gender	-0.05	-0.10	0.07	-0.08	-0.10	0.02	0.14	-0.04	0.32	1		
Weekday Usage Frequency	0.08	0.13	-0.26	0.05	-0.10	-0.08	-0.06	0.13	-0.23	-0.04	1	
Weekend Usage Frequency	0.08	0.19	-0.09	0.12	0.23	-0.38	-0.27	-0.04	-0.15	-0.08	0.31	1

Correlation within independent variables, higher strength may cause multicollinearity

Correlation between dependent and independent variables.

Table 4: The comparison of the logistic regression results (of primary data)

Multinomial

Model Fitting Information

Model	Model Fitting Criteria: -2 Log Likelihood	Likelihood Ratio Tests: Chi-Square	df	Sig.
Intercept Only	78.145			
Final	57.831	20.314	10	.026

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	8.703	8	.368
Deviance	10.691	8	.220

Pseudo R-Square

Cox and Snell	.106
Nagelkerke	.121
McFadden	.054

Parameter Estimates

New Weekdays Usage Frequencies[a]		B	SE	Wald	Sig.	Exp(B)
Every Weekdays	Intercept	-.584	.442	1.748	.186	
	[PrimaryPurposeOfJourney=1]	1.345	.586	5.273	.022	3.837
	[PrimaryPurposeOfJourney=2]	1.378	.587	5.519	.019	3.967
	[PrimaryPurposeOfJourney=3]	.988	.560	3.117	.077	2.687
	[PrimaryPurposeOfJourney=4]	-.664	1.239	.287	.592	.515
	[PrimaryPurposeOfJourney=5]	0[b]	.	.	.	.
	[ConcessionCard=0]	.623	.435	2.047	.153	1.864
	[ConcessionCard=1]	0[b]	.	.	.	.
3-4 days per week	Intercept	-.645	.470	1.884	.170	
	[PrimaryPurposeOfJourney=1]	.741	.666	1.237	.266	2.099
	[PrimaryPurposeOfJourney=2]	1.397	.641	4.757	.029	4.043
	[PrimaryPurposeOfJourney=3]	.184	.647	.081	.776	1.202
	[PrimaryPurposeOfJourney=4]	-.403	1.245	.105	.746	.668
	[PrimaryPurposeOfJourney=5]	0[b]	.	.	.	.
	[ConcessionCard=0]	-.310	.517	.359	.549	.734
	[ConcessionCard=1]	0[b]	.	.	.	.

The reference category is: Less than monthly to 1-2 days a week[a]
This parameter is set to zero because it is redundant.[b]

Binomial

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	16.248	5	.006
	Block	16.248	5	.006
	Model	16.248	5	.006

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	123.719[a]	.082	.157

a. Estimation terminated at iteration number 6 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	7.671	6	.263

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	PrimaryPurposeOfJourney			12.618	4	.013	
	PrimaryPurposeOfJourney(1)	-.965	.667	2.095	1	.148	.381
	PrimaryPurposeOfJourney(2)	-1.856	.720	6.636	1	.010	.156
	PrimaryPurposeOfJourney(3)	-1.270	.711	3.193	1	.074	.281
	PrimaryPurposeOfJourney(4)	1.322	.959	1.900	1	.168	3.750
	Private_Car_Driver(1)	-1.276	.514	6.155	1	.013	.279
	Constant	-.458	.547	.703	1	.402	.632

a. Variable(s) entered on step 1: PrimaryPurposeOfJourney, Private_Car_Driver.

Table 5: Multinomial logistic Regression analysis results for the Imputed Dataset 1

Model Fitting Information

Model	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	107.134			
Final	53.753	53.380	12	.000

Likelihood Ratio Tests

Effect	Model Fitting Criteria	Likelihood Ratio Tests		
	-2 Log Likelihood of Reduced Model	Chi-Square	df	Sig.
Intercept	53.753[a]	.000	0	.
PrimaryPurposeOfJourney	75.887	22.134	8	.005
Age_Group	83.576	29.822	4	.000

The chi-square statistic is the difference in -2 log-likelihoods between the final model and a reduced model. The reduced model is formed by omitting an effect from the final model. The null hypothesis is that all parameters of that effect are 0.

a. This reduced model is equivalent to the final model because omitting the effect does not increase the degrees of freedom.

Pseudo R-Square

Cox and Snell	.222
Nagelkerke	.252
McFadden	.118

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	6.012	8	.646
Deviance	7.570	8	.477

Parameter Estimates

Usage Frequency- Weekdays[a]		B	Std. Error	Wald	df	Sig.	Exp(B)	95% Confidence Interval for Exp (B) Lower Bound	Upper Bound
Every Weekdays	Intercept	-16.124	3171.479	.000	1	.996			
	[PrimaryPurposeOfJourney=1]	1.271	.565	5.064	1	.024	3.564	1.178	10.782
	[PrimaryPurposeOfJourney=2]	1.567	.545	8.263	1	.004	4.792	1.646	13.948
	[PrimaryPurposeOfJourney=3]	.484	.534	.822	1	.365	1.622	.570	4.618
	[PrimaryPurposeOfJourney=4]	-1.587	1.176	1.822	1	.177	.205	.020	2.049
	[PrimaryPurposeOfJourney=5]	0[b]	.	.	0	.	.	.	.
	[Age_Group=1]	16.239	3171.479	.000	1	.996	11279795.20	.000	.[c]
	[Age_Group=2]	13.330	3171.479	.000	1	.997	615193.923	.000	.[c]
	[Age_Group=3]	0[b]	.	.	0	.	.	.	.
3-4 days per week	Intercept	-16.749	.510	1079.498	1	.000			
	[PrimaryPurposeOfJourney=1]	.418	.605	.477	1	.490	1.519	.464	4.977
	[PrimaryPurposeOfJourney=2]	1.189	.551	4.665	1	.031	3.285	1.116	9.665
	[PrimaryPurposeOfJourney=3]	-.047	.598	.006	1	.938	.954	.295	3.084
	[PrimaryPurposeOfJourney=4]	.415	.760	.299	1	.584	1.515	.342	6.712
	[PrimaryPurposeOfJourney=5]	0[b]	.	.	0	.	.	.	.
	[Age_Group=1]	16.414	.475	1196.071	1	.000	13445709.04	5303855.358	34085976.96
	[Age_Group=2]	15.564	.000	.	1	.	5746309.094	5746309.094	5746309.094
	[Age_Group=3]	0[b]	.	.	0	.	.	.	.

a. The reference category is: Less than monthly to 1-2 days a week.

b. This parameter is set to zero because it is redundant.

c. Floating point overflow occurred while computing this statistic. Its value is therefore set to system missing.

6. CONCLUSION

It needs to be noted that the results were obtained from a very small size database of 213 respondents compared to the vast original dataset of 10884 respondents. The analysis was focused on the small dataset only because these provided the responses from the users along the study corridor, which is considered to be statistically valid. Also it meant that the demographic characteristics and the types and preferences of mode choices (and their underlying reasons) would be similar and consistent for the study corridor, which is the key focus of the overall research. However, it is worth consideration of the whole set of data to find out if the list of independent variables and their relative level of influence would be different to this analysis.

In accordance with the TOD principles, this paper has investigated and analysed a conventional public bus route as an alternative to rail transit corridors with an overarching goal of increasing the number of public transport patronage. The study had analysed an existing dataset to identify and to discuss factors/ variables that may or may not influence the decision making of a potential passenger to select public transport. The outcome of this study has helped identification of some independent variables, which will become a key consideration in future transit services design with a focus on identified target groups of the general population such as: working population; population in the higher and lower age groups; population of the lower income groups; and population that has not yet developed the driving skills etc. There will always be a considerable challenge in convincing the general population to give up the freedom of using private modes of transport unless public transport options are presented to them as a genuine and compelling alternative. From the results of this analysis, it can be concluded that a well-designed (i.e. with consideration of the relevant variables as discussed in this paper) bus based transit corridor with coordinated land use planning policies may likely to achieve an increasing public transport patronage within the context of a metropolitan Adelaide and for the cities of similar characteristics.

ACKNOWLEDGMENTS

We are grateful to the Government of South Australia's Department of Planning, Transport and Infrastructure (DPTI), for allowing the University of South Australia to use the Public Transport Customer Survey 2009 data for academic and research purposes.

REFERENCES

Adam, W., Fleming, D. & Parsons Brinckerhoff Australia 2005. Effective bus-based transit oriented development. *Transit Oriented Development: Making It Happen, 2007, Perth, Western Australia, Australia.* Perth, Western Australia, Australia.

Allan, A. & Fielke, M. 2015. Modeling Local Accessibility Networks to Optimize the Planning of Transit Oriented Developments in Adelaide. *Computers in Urban Planning and Urban Management (CUPUM).* MIT, Boston, USA.

Allan, A., Nawaz, M. & Fielke, M. 2015. Transforming Adelaide into a city of networked TODs using buses: case study of the Adelaide OBahn. *Australasian Transport Research Forum.* Sydney, Australia.

Bray, D. & Scrafton, D. 2000. The Adelaide O'Bahn: Ten Years On. *8th Joint Conference on Light Rail Transit.* Dallas, Texas, 2000: American Public Transportation Association and the Transportation Research Board.

Cervero, R. 2000. Transport and land use: Key issues in metropolitan planning and smart growth. *University of California Transportation Center*.

Cervero, R. & Dai, D. 2014. BRT TOD: Leveraging transit oriented development with bus rapid transit investments. *Transport Policy,* 36**,** 127-138.

Currie, G. 2005. Strengths and weakness of bus in relation to transit oriented development. *Transit Oriented Development–Making it Happen Conference.* Perth, Western Australia.

Currie, G. 2006. Bus rapid transit in Australasia: Performance, lessons learned and futures. *Journal of Public Transportation,* 9**,** 1.

Currie, G. 2014. Public transport progress and failure-Keeping up with growth in Australian cities. *Infrastructure for 21st Century Australian Cities: Papers from the ADC Forum National Infrastructure and Cities Summit March 2014.* Sydney, Australia: ADC Forum.

Department of Planning and Local Government - South Australia. *30 Year Plan for Greater Adelaide (30YPGA)* [Online]. Available: http://www.dpti.sa.gov.au/__data/assets/pdf_file/0006/132828/The_30-Year_Plan_for_Greater_Adelaide_compressed.pdf [Accessed 20 July 2013].

Ho, C. & Mulley, C. 2014. Metrobuses in Sydney: How high capacity and high frequency services are benefiting the Metropolitan fringe. *Research in Transportation Economics,* 48**,** 339-348.

Laerd Statistics. 2017. *Multinomial Logistic Regression using SPSS Statistics* [Online]. Lund Research Ltd. Available: https://statistics.laerd.com/spss-tutorials/multinomial-logistic-regression-using-spss-statistics.php [Accessed 14 January 2017.

Miles, J. & Shevlin, M. 2000. *Applying Regression and Correlation: A Guide for Students and Researchers*, SAGE Publications.

Olinsky, A., Chen, S. & Harlow, L. 2003. The comparative efficacy of imputation methods for missing data in structural equation modeling. *European Journal of Operational Research,* 151**,** 53-79.

Parker, T. 2002. Statewide Transit-Oriented Development Study: Factors for Success in California: Final Report. California Department of Transportation.

Rose, J., Mulley, C., Tsai, C.-H. P. & Hensher, D. 2013. Will bus travellers walk further for a more frequent service? A Stated preference investigation. *Australasian Transport Research Forum (ATRF), 36th, 2013, Brisbane, Queensland, Australia.*

Schafer, J. L. & Graham, J. W. 2002. Missing data: our view of the state of the art. *Psychological Methods,* 7**,** 147.

Schlomer, G. L., Bauman, S. & Card, N. A. 2010. Best practices for missing data management in counseling psychology. *Journal of Counseling Psychology,* 57**,** 1.

Somenahalli, S., Sleep, C., Primerano, F., Wadduwage, R. & Mayer, C. 2013. Public Transport Usage in Adelaide. *Procedia - Social and Behavioral Sciences,* 104**,** 855-864.

Taylor, M. a. P., Holyoak, N. & Meng, L. 2011. Increasing the patronage of Adelaide's Northern Rail Corridor. *Australian Transport Research Forum* Adelaide.

Taylor, R. 1990. Interpretation of the correlation coefficient: a basic review. *Journal of Diagnostic Medical Sonography,* 6**,** 35-39.

The Government of South Australia. 2014. *"New trams, trains and buses"* [Online]. Available: https://www.sa.gov.au/topics/transport-travel-and-motoring/public-transport-and-travel/new-trams-trains-and-buses [Accessed 29 May 2014].

The Government of South Australia. 2015. *The Integrated Transport and Land Use Plan, July 2015* [Online]. Available: http://www.transportplan.sa.gov.au/ [Accessed 23 December 2015].

Tsikriktsis, N. 2005. A review of techniques for treating missing data in OM survey research. *Journal of Operations Management,* 24**,** 53-62.

Van Ness, P. H., Murphy, T. E., Araujo, K. L. B., Pisani, M. A. & Allore, H. G. 2007. The use of missingness screens in clinical epidemiologic research has implications for regression modeling. *Journal of Clinical Epidemiology,* 60**,** 1239-1245.

Vriens, M. & Melton, E. 2002. Managing missing data. *Marketing Research,* 14**,** 12.