



**LOW CARBON LIVING  
CRC**

## **RP 2021 Greening Suburban Travel**

**Smartcard data analysis**



Authors	Mrs. Mona.Mosallanejad, Dr. Sekhar Somenahalli and Mr. Callum Sleep
Title	Smartcard data analysis
ISBN	
Format	
Keywords	Low Carbon Living, Smartcard data, Adelaide
Editor	
Publisher	CRC for Low Carbon Living
Series	
ISSN	
Preferred citation	



**Australian Government**  
**Department of Industry,  
Innovation and Science**

**Business**  
Cooperative Research  
Centres Programme



**LOW CARBON LIVING  
CRC**

## Acknowledgements

This research is funded by the CRC for Low Carbon Living Ltd supported by the Cooperative Research Centres program, an Australian Government initiative.

## Disclaimer

Any opinions expressed in this document are those of the authors. They do not purport to reflect the opinions or views of the CRCLCL or its partners, agents or employees.

The CRCLCL gives no warranty or assurance and makes no representation as to the accuracy or reliability of any information or advice contained in this document, or that it is suitable for any intended use. The CRCLCL, its partners, agents and employees, disclaim any and all liability for any errors or omissions or in respect of anything or the consequences of anything done or omitted to be done in reliance upon the whole or any part of this document.

## Peer Review Statement

The CRCLCL recognises the value of knowledge exchange and the importance of objective peer review. It is committed to encouraging and supporting its research teams in this regard.

The author(s) confirm(s) that this document has been reviewed and approved by the project's steering committee and by its program leader. These reviewers evaluated its:

- originality
- methodology
- rigour
- compliance with ethical guidelines
- conclusions against results
- conformity with the principles of the [Australian Code for the Responsible Conduct of Research](#) (NHMRC 2007),

and provided constructive feedback which was considered and addressed by the author(s).

© 2018 Cooperative Research for Low Carbon Living

# Greening Suburban Travel

## Smartcard data analysis

### Table of Contents

Acknowledgements .....	3
Disclaimer.....	3
Peer Review Statement.....	3
Background .....	2
Data Structure .....	3
Extraction of OD matrices .....	5
Origin and Destination analysis .....	6
On-Time Performance of Bus Services using smartcard data.....	10
Assessing bus arrival time reliability.....	12
Processing data for reliability .....	13
Analysis of bus arrival time reliability.....	14
Conclusions .....	17
References.....	18

## List of Figures

Figure 1. Origin and Destination counts for each suburb (bus users).....	6
Figure 2 Suburbs with a high number of transfers (morning peak and weekdays) .....	8
Figure 3 Transfer type during weekends and weekdays.....	9
Figure 4 Comparison of weekday and weekend route lateness frequency distribution.....	14
Figure 5 The distribution of lateness at boarding stops .....	15

## List of Tables

Table 1 Individual MetroCard information .....	3
Table 2 Initial boarding data structure .....	13
Table 3 Descriptive statistics of route groupings .....	15
Table 4 Statistics for passenger observation data .....	16

## Background

As public transport agencies increasingly adopt the use of automatic data collection systems, a significant amount of boarding data becomes available, providing an excellent opportunity for transit planners to access spatial-temporal data (Rahbar et al. 2017; Tao 2018) which can be used for a better understanding of human mobility and the performance of a transit system (Mahrsi et al. 2017). Smartcard data can be used to examine a whole network regularly, and to make practical estimates of passenger origin-destination (OD) patterns and is a great asset in understanding public transport reliability issues. Having knowledge of public transport travel

demand and its reliability will facilitate the design of appropriate public transport routes, and increase the efficiency which will, in turn, enhance public transport patronage. The data used in this research was provided by the Department of Planning Transport and Infrastructure (DPTI) in Adelaide, South Australia for the month of May 2017. A methodology was developed, using SQL software and based on the trip chain model, to create an OD matrix for Adelaide's bus users, and from it to estimate the demand for the system. Adelaide was chosen for this study because unlike in other cities, commuters scan their smartcard upon boarding but not on alighting. This allows the algorithm to be generic and therefore applicable elsewhere.

## Data Structure

The primary function of the smartcard is to collect a fare, but it can also be utilised for finding passengers' travel patterns. Usually, smartcard data does not directly provide the information required for planners (Kurauchi & Schmöcker 2016). The flat fare policy and some zonal fare policies require commuters to tap once after boarding and record only a single transaction. However, in some cities, an exit reader is available as well if the fare is based on distance or zones. In such a system, each trip generates two records: for boarding and for alighting (Kurauchi & Schmöcker 2016). Each MetroCard contains spatial and temporal

information. In Adelaide, where a flat fare policy operates, commuters validate their cards when they board a public vehicle but not on alighting. Three modes of transport are available: bus, train and tram. The information for each smartcard transaction contains card identification, time, date, transport mode used, fare type, stop code, stop label, route code and validation type (see Table 1). When passengers swipe their card and pay an initial transaction, the fare is valid for two hours, and passengers can use any public transport within this time without incurring further costs.

Table 1. Individual MetroCard information

Media code	Fare type	Transport mode	Date & time	Stop code	Latitude	Longitude	Route code	Direction
807***CB	SV	4	2017-05-01 09:49:35	8089	- 34.979759	138.525912	Tram	1
94E***FB	TICKETS	1	2017-05-01 10:39:15	3351	- 34.924343	138.598468	251	1
11C***89	28DAY	1	2017-05-05 10:46:32	3285	- 34.920343	138.607313	271	1
707***27	OTHER	1	2017-05-01 11:04:05	2072	- 34.870071	138.638452	H22	1
584***97	SV	5	2017-05-08 11:06:36	1852	- 34.860916	138.650472	GWC	1

Note: Transport mode: 1 = Bus, 3 = Station, 4 = Tram, 5 = Train

There are some deviations from the one-swipe rule: railway stations in Adelaide operate under a closed system, and swiping is required for both boarding and alighting, and various systemic and user issues mean that transfers between the train and other modes cannot be estimated directly from the MetroCard. In addition, there is a free tram zone in Adelaide where passengers do not need to swipe their cards; this means that the tram boarding point is not available. Given these limitations, this study focuses on bus users.

## Extraction of OD matrices

Knowledge of transit demand plays a decisive role in public transport plans to improve the performance of the system. One common method for estimating the destination is the trip chain model. As mentioned previously, each smartcard can provide the boarding location and time of each bus trip but not the alighting location, so the trip chain model assumes the alighting stop is located within acceptable walking distance of the next stop. Some assumptions considered in this algorithm are:

- The initial boarding location of a trip leg is the 'origin'.
- A passenger's alighting point is assumed to be within walking distance of the next boarding stop
- Passengers return to the place where they first boarded that day, or to some other nearby station.
- Commuters take the first available service after arriving at a boarding place.
- Each smartcard is used by a single commuter and cannot be used by multiple passengers.
- Commuters who use the public transport system do not use any other mode of transport on that same day

## Origin and Destination analysis

One of the critical considerations when planning transit services is estimating the demand for each route, to determine the frequency and capacity of the vehicles (Tamblay, Muñoz & Ortúzar 2018). An OD matrix provides critical information for transit planners by estimating the number of journeys between different zones, information which can be used in transportation planning, design and management. After analysing the data based on

the trip chain model, bus users' origins and destination counts during the morning peak were derived for each suburb (Figure 1). Most trips originated from Paradise, Modbury, Adelaide, and Klemzig suburbs. Adelaide, Bedford and Modbury suburbs were the destinations for most journeys during the day. Suburbs with the highest origins and destinations were shortlisted and analysed. Finally, the data was validated using other sources from the Department of Transport and a limited survey.

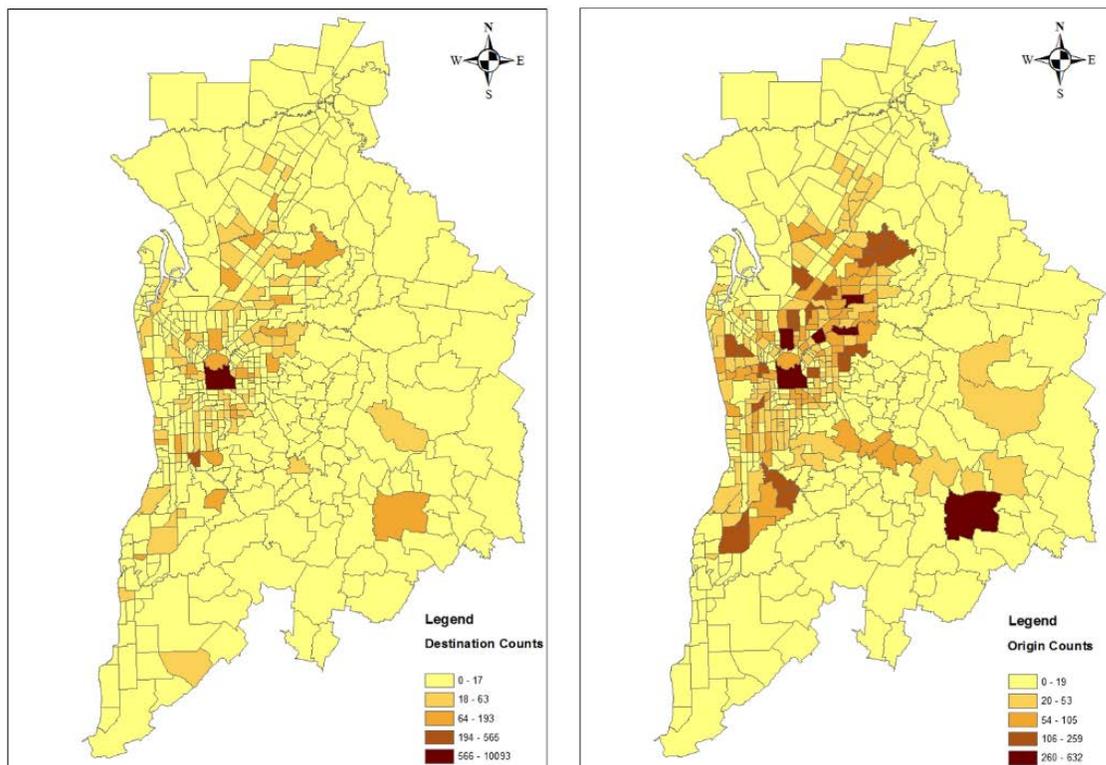


Figure 1. Origin and Destination counts for each suburb (bus users)

The origin-destination analysis showed that bus movements were radial, and most trips during the morning peak ended in the CBD. These movements were further explored to rationalise the existing routes. The information from an OD analysis that was used to identify new bus routes to optimize the bus routes and reduce congestion within the Central Business District (CBD) of Adelaide. Few examples of suggested new routes are given below:

- Modbury–Bedford Park: the OD analysis showed high demand from Modbury to Flinders University during the morning peak, but just one route (G40) runs between the suburbs, going through the CBD. The results indicate that providing direct routes from Modbury interchange to Bedford park will reduce overall travel time to passengers and also reduce congestion in the Adelaide CBD.
- Paradise–Bedford Park: there are two bus routes between these two suburbs (W90 and G40), and both pass through the CBD, which is heavily congested during the morning peak. It is worth

exploring the option of a direct route from Paradise to Flinders University that avoids congested city links.

- Modbury–North Adelaide: bus routes between these two suburbs run through the CBD. As explained above, by introducing direct routes between these two suburbs will lead to reduced congestion in the city center.

The OD analysis also helped us in greater understanding about the mode transfers and the role played by the key interchanges. In addition, the results indicate that all the transfer points are the same for both weekdays and weekends, although destinations may change. Most transfers during weekdays occurred in three suburbs: Adelaide (CBD), Paradise, and Modbury and as expected most passengers travelled to Adelaide during the morning peak to start a daily activity (refer Figure 2) . Modbury and Paradise are identified as the busiest interchanges, and it is evident that most commuters use these locations for transfer.

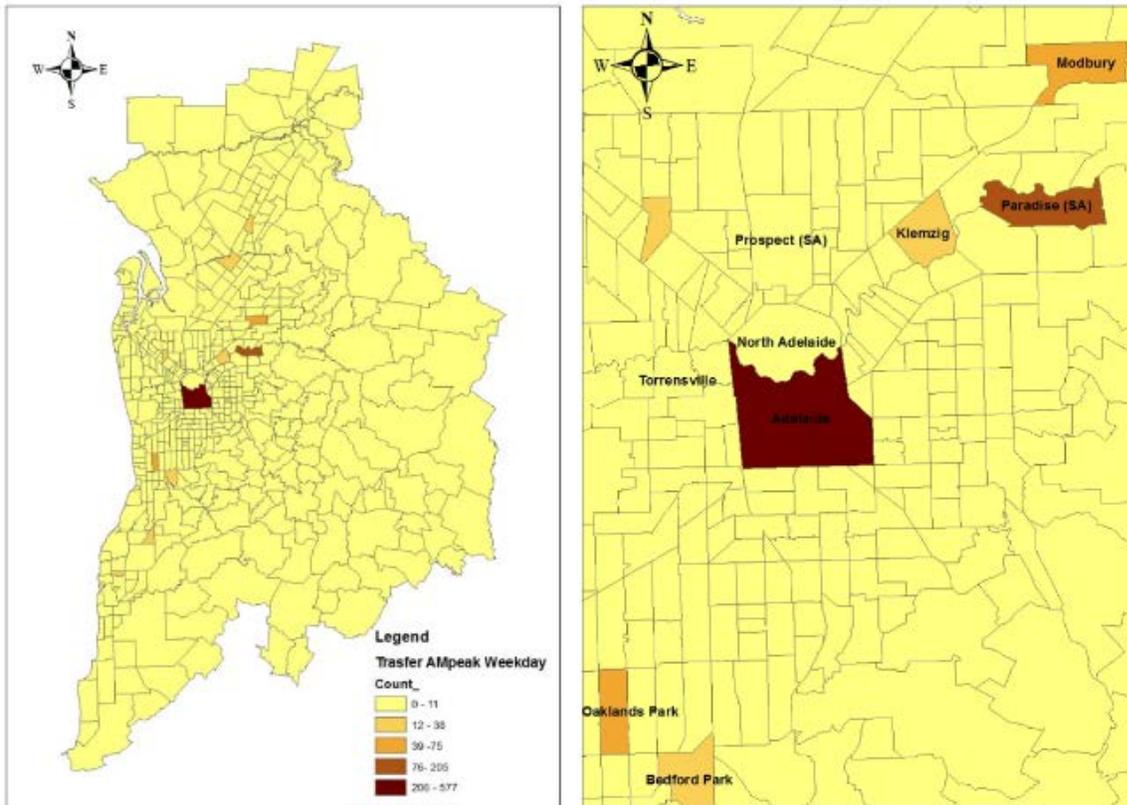


Figure 2 Suburbs with a high number of transfers (morning peak and weekdays)

As presented in Figure 3, the comparison of weekday and weekends illustrates that transfer between bus to bus and train to train are more popular than other kinds of transfers. While on the weekend, most commuters transferred from bus to bus (51%), and on the weekdays it decreased to 39.76%, which means that the number of passengers who transfer from train to

train on the weekdays is higher than on the weekend. It can be related to the reliability of trains in comparison with bus due to traffic congestion and less travel time, so people prefer to use a more reliable system on the weekdays to access to their destination. Travel pattern among bus to train transfers and vice versa is the same during weekdays and weekend.

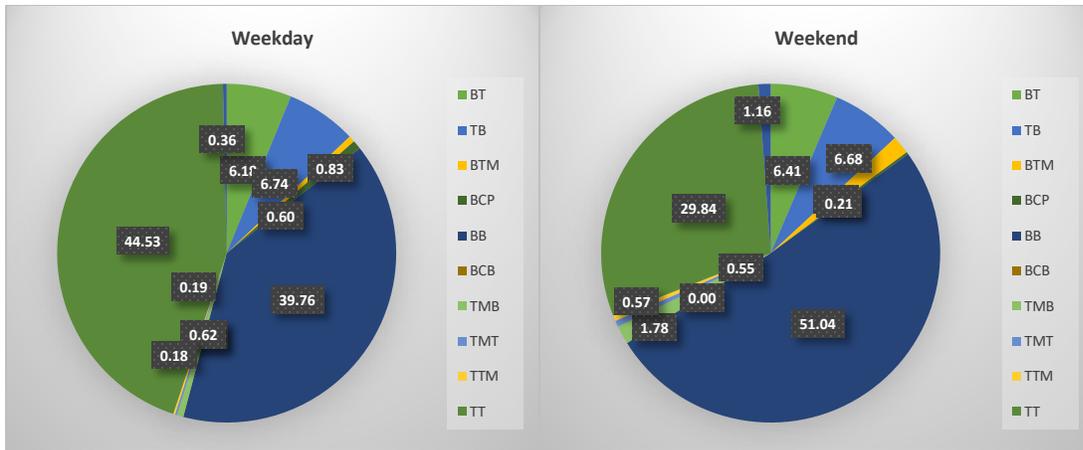


Figure 3: Transfer type during weekends and weekdays.

## On-Time Performance of Bus Services using smartcard data

Throughout the world, the reliability of public transport systems is constantly under review. In recent years, the widespread prevalence of privately owned motor vehicles and people's quickening pace of life has increased the importance of public transport service reliability and on-time performance. This is of potential concern for bus services as buses share road space with a growing number of other vehicles. In Adelaide, the capital city of South Australia, the public transport system has been plagued by concerns of unreliable services (Kelton, 2012a).

The South Australian community is encouraged, by the government, to use public transport especially for regular trips such as the daily commute. However, the South Australian public sector has found that many commuters are avoiding public bus services, reducing the total number of commuters using public transport (Kelton, 2012b). South Australia's initial boardings for metropolitan public transport rose each year incrementally between 2000 and 2009, reaching 52.4 million in the 2009–2010 financial year (DPTI, 2011). However, DPTI's Annual Report for 2010–2011 (Department of Transport Energy and Infrastructure, 2011) states that in 2010–2011, initial boardings reduced by 2.2 per cent to 51.25 million. One reason for this reduction is the perceived unreliability of services. (Nankervis, 2016) Often, buses do not meet the advertised service times, with many services running a quarter or even half an hour late—or, in some cases, not arriving at all (Kelton, 2012a).

South Australia's Public transport system is operating well below its full potential. According to the Australian Bureau of Statistics (Australian Bureau of Statistics, 2009), 14.4 per cent of adults across Australia were using public transport for their trip to work or study in 2006, while in Adelaide this figure was less than 10%. The use of public transport between 1996 and 2006 increased by only 18 per cent in Adelaide, dwarfed by increases of 35 per cent and 22 per cent in Melbourne and Brisbane respectively (Australian Bureau of Statistics, 2009).

According to the Adelaide Metro website (Adelaide Metro, 2012), the quality of South Australian public transport needs minor improvement. DPTI monitors the performance of the bus contractors to make sure that the service quality (on-time running and reliability) meets community needs and demands. DPTI defines service as 'on-time' and 'reliable' if the vehicle departs no more than 59 seconds before and no more than 4:59 minutes (i.e., 4 minutes 59 seconds) after the time published in the timetable (Adelaide Metro, 2012). It must be noted that not all stops appear on the timetables; at these locations, estimated times are provided to the travelling public. Even with 6 minutes' flexibility, a large proportion of services are failing to meet targets. This lack of reliability for public transport services is a significant concern for the community.

In the past, several attempts have been made to improve the reliability of bus services in Adelaide, including: fining the contractors operating the bus services when they fail to meet targets (Bray & Wallis, 2008); continuously changing and reviewing timetables to suit changing road conditions; fitting buses with

Global Positioning System (GPS) devices; and auditing buses to determine which bus routes require attention. Automated Vehicle Location (AVL) systems are helping public transport agencies all over the world to improve their performance. However, there is a difference between the performance at the vehicle level and what the passenger experiences, often at the stop level (Chen et al., 2009) so it is important to collect and interpret the data accordingly. This study seeks to investigate travel time reliability as seen by the passenger using data collected automatically.

## Assessing bus arrival time reliability

Using boarding data to assess travel time has the advantage that these records directly relate to passenger experiences. Furthermore in Adelaide boarding data is recorded at the stop locations, as are the timetables, eliminating the need to process and compare the datasets geographically. Bus services were separated from other route services offered by Adelaide Metro for analysis primarily because they form the bulk of the network and are most affected by travel time variability. In the Adelaide network, the bus driver is also the ticket salesman and therefore must wait until he/she is satisfied no further passengers need to buy a ticket before departing the stop. Compare this to the rail services where fare payment is collected by an onboard vending machine; a passenger could conceivably buy a ticket in transit and validate it as the vehicle is about to reach the next stop. Time spent selling tickets has been previously attributed exclusively to dwell time (time spent stationary at a stop) (Dorbritz et al., 2009), this helped shape the approach for the investigation, giving confidence that the last boarding would reflect bus departure time. As expected similarity was observed when the estimations of departure time from boarding data were compared to those obtained from Automatic Vehicle location records. These records were obtained previously covering a small selection of routes.

There are however some limitations of this method. Primarily, we only have data on a bus's location when a person, passenger or staff, boards and validates a ticket. This makes it difficult to track a bus's progress in the afternoon heading

away from Adelaide and impossible to ascertain arrival time at terminals. A more detailed study to address these concerns could be undertaken with AVL data.

## Processing data for reliability

The busses true departure time was estimated from the last validation at a particular stop. This is deemed valid for assessing the lateness for bus services as there is only one boarding door at the front of the bus and the driver's presence helps enforce fare payment. For example, the records in Table 2 below are those showing the progress of bus 1125 along route 503. Those records highlighted in the darkest grey will be retained for further processing.

The raw data as shown in Table2 has three distinct sections of information. Firstly there's the identifying information specific to each record in the form of an ID and a timestamp. Next there's the geographical information identifying the boarding location, and finally, there's the service information relating to the vehicle's operation.

Table 2: Initial boarding data structure

1	VALIDATION_ID	VALIDATION Time	STOP_CODE	STOP_LABEL	STOP_LATITUDE	STOP_LONGITUDE	ROUTE_CODE	TRIP_DIRECTION	Vehicle_NUMBER
196028	1609919983	8:26:36	6931	Zone A Klemzig Interchange	-34.886524	138.639269	C1		2 1125
196029	1609919985	8:26:50	6931	Zone A Klemzig Interchange	-34.886524	138.639269	C1		2 1125
196030	1609920004	9:13:02	1217	Zone A Tea Tree Plaza Intercha	-34.831325	138.694308	503		2 1125
196031	1609920006	9:13:05	1217	Zone A Tea Tree Plaza Intercha	-34.831325	138.694308	503		2 1125
196032	1609920027	9:29:05	1656	34D Valiant Rd	-34.85004	138.67158	503		2 1125
196033	1609920029	9:29:08	1656	34D Valiant Rd	-34.85004	138.67158	503		2 1125
196034	1609920031	9:29:10	1656	34D Valiant Rd	-34.85004	138.67158	503		2 1125
196035	1609920034	9:30:01	1706	34C Valiant Rd	-34.853742	138.671835	503		2 1125
196036	1609920039	9:31:46	1762	34 Lyons Rd	-34.857462	138.667326	503		2 1125
196037	1609920041	9:31:48	1762	34 Lyons Rd	-34.857462	138.667326	503		2 1125
196038	1609920047	9:34:09	1869	29B Sudholz Rd	-34.861157	138.661002	503		2 1125
196039	1610698084	15:33:03	7735	School Urrbrae College Cross R	-34.964703	138.625472	990		1 1125
196040	1610698086	15:33:05	7735	School Urrbrae College Cross R	-34.964703	138.625472	990		1 1125

## Analysis of bus arrival time reliability

The data was first aggregated by route, and while there was some inconsistency in average lateness observed across the three days individually, there is a clear trend observed towards consistency across days. On a scatter plot all the routes and their average likenesses were plotted for each day separately. Each route was treated equally spaced one unit apart along the X-axis. Regardless of the order of the routes presented the trend lines for the two days are in high agreement both showing little to no variance across bus services (Figure

4). Statistically, the routes on Wednesday and Thursday have an almost standard distribution of lateness with mean and median values within 20 seconds. Their average lateness was within 21 seconds through the variance as measured by standard deviation was almost 40 seconds higher on Thursday. As might be expected bus services showed less variability on Saturday with average lateness of 55 seconds and a standard deviation of fewer than two minutes (Table 3). The figure depicts this difference between route lateness distributions on a weekday vs a weekend. Because there were no routes in every time category a moving average (Mov.Avg.) has been used to smooth the distribution curve.

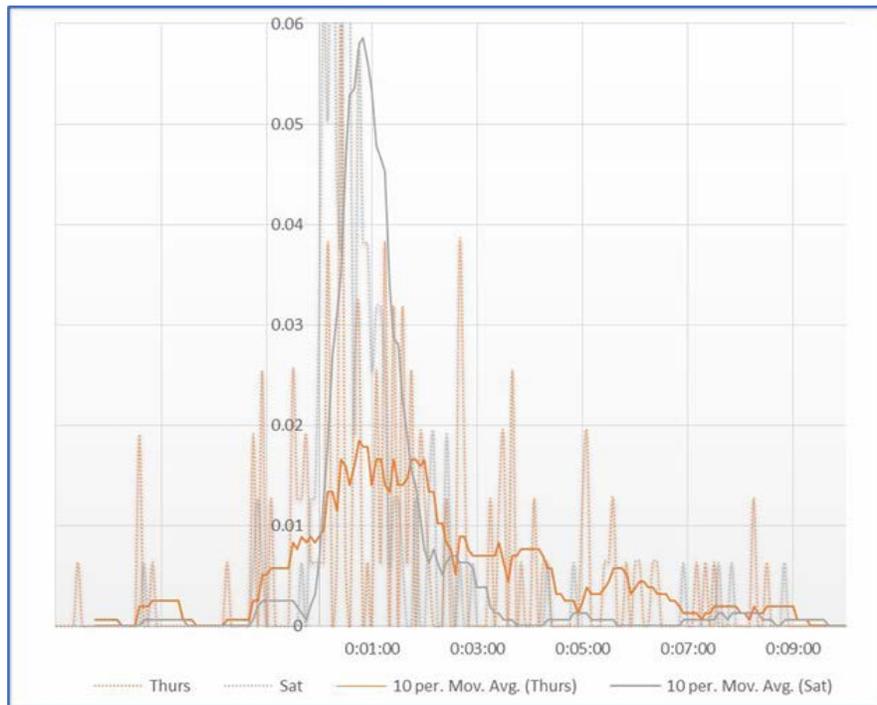


Figure 4: Comparison of weekday and weekend route lateness frequency distribution.

Table 3 : Descriptive statistics of route groupings

	Wed (Time: hh:mm:ss)	Thurs (Time: hh:mm:ss)	Sat (Time: hh:mm:ss)
<b>Average</b>	0:01:40	0:01:22	<b>0:00:55</b>
<b>St dev</b>	0:02:27	0:03:05	<b>0:01:53</b>
<b>Median</b>	<b>0:01:21</b>	<b>0:01:10</b>	<b>0:00:30</b>

The average observed lateness was calculated as well as the percentage of boarding locations where the service exceeded the five-minute tolerance for lateness. The bus route 747 stood out here as being on average between 7.5 and 9.25 minutes late across both days. Furthermore, buses at 17% of boarding locations were reported as late by the Adelaide metro standard of arrival 5 minutes or more after the published time.

The 747 route is a feeder service linking the Seaford and Noarlunga interchanges in Adelaide's

far south in a clockwise loop. Interestingly the 745 route which follows the same streets but in an anticlockwise direction showed much more variability with average lateness approaching 12 minutes on Wednesday and only 3.5 minutes on Thursday. Perhaps this is due to a disproportionate number of un-signalised right turns required across traffic for the 745 services. When considering the whole dataset of passenger observations shown in Figure rather than route groups, further differences were observed.

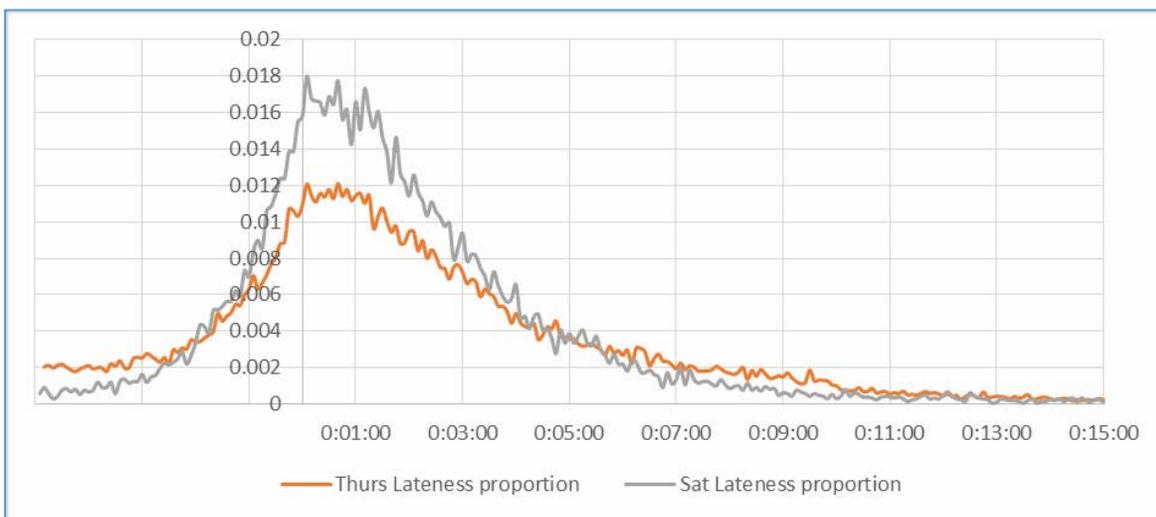


Figure 5 The distribution of lateness at boarding stops

Table 4 Statistics for passenger observation data

AVERAGE HR:MIN:SEC	0:00:36 (Thursday)	0:02:13 Saturday
MEDIAN HR:MIN:SEC	0:01:03	<b>0:01:12</b>
ST DEV HR:MIN:SEC	0:13:03	<b>0:08:26</b>
% ONTIME	<b>56.332%</b>	<b>75.451%</b>

Of the weekdays the services on Thursday performed considerably poorer with a standard deviation of 13 minutes. This higher spread means the buses were within the acceptable limits of +1 and -5 minutes for only 56% of observations compared with 75.5% of those on Saturday. Where there were multiple boardings of bus service at the same stop the time between boardings could be found. A smartcard, boarding took place on average 12 seconds after the previous boarding, whereas boarding with a magnetic ticket (tickets sold inside the bus) took twice as long as 24 seconds. It was also found that these magnetic ticket boardings were over-represented in the database of final boardings at each stop. Across a typical week, single trip tickets make up 8% of all boardings while these same tickets are the final boarding at a stop 17% of the time. This does confirm that the sales and validation process is prolonging the time that buses stand at some stops. When the effect was investigated at a network level, it was found that there is no relationship between the percentage of departures where the last recorded boarding was with a paper ticket and how late the buses

became. This indicates that the sale of tickets inside the bus, is at most a minor cause of travel time unreliability. The presence of extra time taken by paper/magnetic ticket purchase having no effect on reliability indicates that the distribution of such events either spatially or temporarily is captured in the timetable. This implies that there are travel time savings if not reliability improvements to be gained through off-board ticket sales or prepaid only services such as those used in Sydney (Byatt, Oscuro & Rookes, 2008).

## Conclusions

The public transport OD matrix is a useful prerequisite for planners to optimise public transport systems. The reliability of the system is an important criterion to encourage people to leave their vehicles at home and take public transport instead. This research presented an overview of ridership patterns using one-month MetroCard data in Adelaide. An accurate estimation of public transport OD will be a significant help to public agencies involved in route rationalisation, which will lead to higher public transport patronage. In further studies, census data could be used to validate this

algorithm, and sensitivity analysis could also be carried out for various assumptions. Adelaide's bus services show less variability of lateness on the weekends although there is a lower percentage of prepaid tickets used. Ticket sales inside the bus are increasing the travel time of Adelaide's public transport bus services. However, they are not contributing to travel time unreliability. Removing the cash ticket sales from Adelaide's bus network will not improve reliability. however, travel time savings could be achieved

## References

- Adelaide Metro 2012, *On time running performance archive*, Department of Planning Infrastructure and Transport, Adelaide. Available from: <https://www.adelaidemetro.com.au/About-us/On-Time-Running> [July 2018].
- Australian Bureau of Statistics 2009, *Public transport use for work and study*. Available from: <http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/4102.0Chapter10102008>>
- Bray, D & Wallis, I 2008, 'Adelaide bus service reform: impacts, achievements and lessons', *Research in Transportation Economics*, vol. 22, no. 1, pp. 126
- Byatt, M, Oscuro, G & Rookes, M 2008 'Improving efficiency: an evaluation of Sydney buses', *International Conference Series on Competition and Ownership in Land Passenger Transport*.
- Chen, X, Yu, L, Zhang, Y & Guo, J 2009, Analyzing urban bus service reliability at the stop, route, and network levels, *Transportation Research Part A: Policy and Practice*, vol. 43, no. 8, pp. 722–734.
- Department for Transport Energy and Infrastructure (DPTI) 2011 *Annual Report 2010–11*. Available from: [http://www.dpti.sa.gov.au/data/assets/pdf\\_file/0020/75035/Annual\\_report\\_2010\\_2011\\_online.pdf](http://www.dpti.sa.gov.au/data/assets/pdf_file/0020/75035/Annual_report_2010_2011_online.pdf)>
- Dorbritz, R, Luthi, M, Weidmann, AU & Nash, A 2009, 'Effects of Onboard Ticket Sales on Public Transport Reliability', in *Transportation Research Record Journal of the Transportation Research Board 2110*, pp. 112–119.
- Kelton, S 2012a, 'More commuters avoiding Adelaide buses', *The Advertiser*. Available from: <https://www.adelaidenow.com.au/news/south-australia/more-commuters-avoiding-adelaides-buses/news-story/5ed76b4ae4d8e28bc14028a1e76b50a5>>
- Kelton, S 2012b, 'One in three Adelaide buses are running late, State Government data shows', *The Advertiser*. Available from: <http://www.adelaidenow.com.au/news/south-australia/one-in-three-buses-miss-deadlines/story-e6frea83-1226413857865>
- Kurauchi, F & Schmöcker, J-D 2016, *Public transport planning with smart card data*, CRC Press, Boca Raton, FL.
- Mahrsi, MKE, Come, E, Oukhellou, L & Verleysen, M 2017, 'Clustering smart card data for urban mobility analysis per cent', *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 3, pp. 712–728.
- Nankervis, D 2016, 'Adelaide Metro survey: public transport users complain about frequency, reliability of bus, train and tram services', updated July, 2018, *Advertiser*, Adelaide. Available from: <https://www.adelaidenow.com.au/news/south-australia/adelaide-metro-survey-public-transport-users-complain-about-frequency-reliability-of-bus-train-and-tram-services/news-story/8ad8276a86f5b93aad9f8cc56d1fb3f9>>
- Rahbar, MM, Hickman, M & Tavassoli, A 2017, 'Determining route-choice behaviour of public transport passengers using Bayesian statistical inference', *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice*, vol. 26, no. 1.
- Tamblay, S, Munoz, J-C & de Dios Ortuzar, J 2018, *Extended methodology for the estimation of a zonal origin-destination matrix: a planning software application based on smart card trip data*, Transport Research Board, Washington, DC.
- Tao, S 2018, 'Samrtcard data evaluation', in F Kurauchi & J-D Schmöcker (eds), *Public transport planning with smart card data*, CRC Press, Boca Raton, FL.