

Crowding Costs and Expansion Factors for Sydney's Heavy Rail Network

Alexander J. Svanberg, Senior Economics Manager, Transport for NSW¹

¹231 Elizabeth Street, Sydney 2000

Email for correspondence: alex.svanberg@transport.nsw.gov.au

Abstract

Sitting and standing on trains in crowded conditions creates discomfort and anxiety for rail customers, a cost which is measured and monetized in cost-benefit analysis (CBA). This paper expands on the crowding cost methodology and parameters estimated in Douglas (2006) and the Transport for NSW (TfNSW) *Economic Parameter Values* (EPV) to calculate the average daily and total crowding costs across Sydney's heavy rail network using pre-COVID-19 Opal ticketing data. Given transport demand modelling is often only undertaken for the AM or PM peak hour period, this paper also provides cost expansion factors for Sydney, which allows CBA practitioners to estimate full-day and annual crowding costs from peak hour modelling. This paper finds that crowding costs on the Sydney heavy rail network exceeded a hundred million dollars in 2019, and that expansion factors varied significantly across different heavy rail lines.

1. Introduction

Crowding is a significant issue for public transport networks within many large urban centres, with Sydney cited as one of several large cities that faces issues with congestion across the transport network (L Svanberg & R Pydokke 2020). The effects of crowding are multi-faceted, and include increases in wait times, reduced reliability of services, increased in-vehicle time, and wellbeing impacts for customers, among others (Tirachini et al 2013).

To address public transport crowding issues, transport agencies can adopt a range of different policy responses. Increases in frequency or capacity of services, investments in additional network infrastructure, and pricing schemes to encourage travel outside of peak have all been implemented in Sydney in one form or another. In order to identify which policy interventions are likely to generate the best returns for the NSW community, Transport for NSW (TfNSW) and NSW Treasury recommend the use of cost-benefit analysis (CBA) to support informed investment decision-making (TfNSW 2019).

In NSW, transport CBAs often include the benefit (disbenefit) of a reduction (increase) in crowding costs that result from an investment or initiative. This specific benefit – generally labelled 'reduced crowding costs' or 'de-crowding benefits' – captures the discomfort and anxiety experienced by rail customers when they are forced to sit or stand in crowded conditions on their journey (Douglas 2006). Importantly, studies on perceived crowding costs undertaken in Sydney reveal that the impact of crowding increases exponentially with increases in passenger density (Douglas 2006, and Douglas & Jones 2016). While crowding can lead to

flow-on impacts that are also considered in CBA, such as increased wait time, suppressed demand for travel, or reduced reliability, those impacts are not the focus of this paper.

CBA practitioners looking to estimate the crowding costs for public transport customers in Sydney have access to approaches and parameter values outlined in TfNSW guidance (TfNSW 2020), but also face challenges in terms of data limitations. Sydney-specific studies (e.g. Douglas & Jones 2016) can be used to estimate the cost of crowding, provided comprehensive and detailed patronage data is available. CBA undertaken for NSW Transport cluster projects often use demand models to produce this patronage data. The most commonly used demand models for estimating rail patronage and crowding are the Public Transport Project Model (PTPM) or the Enhanced Train Crowding Model (ETCM). However, PTPM and ETCM provide forecasts of patronage on the Sydney rail network for the one-hour AM peak period only. To estimate annual patronage and crowding levels for the remainder of the day, and over a full year, crowding-specific expansion factors are needed to ‘scale up’ the one-hour AM results.

While expansion factors for patronage are commonly available, separate expansion factors for crowding are needed to accurately estimate annual crowding cost impacts, due to the exponential relationship between passenger density and crowding cost. In essence, the relationship between peak patronage and interpeak patronage is not equivalent to the relationship between peak crowding cost and interpeak crowding cost. Crowding-specific or ‘cost’ expansion factors are more difficult and data-intensive to estimate, because additional passenger-per-m² data is required for the crowding cost functions used in CBA.

This analysis builds on the existing crowding cost formulas reported in the TfNSW EPV (2020) based on Douglas (2006) to estimate total crowding costs per day, per line, on Sydney’s heavy rail network. In addition, estimating crowding costs for all services across a sample of full weekdays also allows for the calculation of 1hr, 2hr, and 3.5hr peak to daily cost expansion factors, per line, on Sydney’s heavy rail network. Currently, the TfNSW EPV does not report detailed crowding expansion factors – instead, project teams must undertake costly and time-consuming analysis to estimate these factors on a project-by-project basis. The information presented in this paper can be used to inform more accurate and consistent CBA and investment decision-making on NSW Transport cluster projects.

The rail volume and crowding expansion factors reported in this paper are calculated based on the TfNSW electronic ticketing ‘Opal’ system. Opal data is collected at the entry and exit station for all paid customer journeys, but does not capture the specific service that passengers use. For this analysis, data was extracted from the Rail Opal Assignment Model (ROAM), a model that assigns Opal journeys to services based on the rail daily working timetable and train punctuality data. More information on the ROAM data set used in this analysis can be found in Section 2.1.

ROAM was created by TfNSW to estimate loads on trains based on Opal ticketing data. It works by assigning Opal journeys to services based on a rail daily working timetable that has been offset to reflect any recorded delay on any given day. The customer load is aggregated to train services and assigned to rail lines to model train capacity, entries, and exits at each station on a train’s route. The ROAM data used for this analysis is based on a sample of six weekdays from March 2019, chosen to align with input data used by Sydney’s main public transport demand models. As Opal journeys includes fare-paying passengers only, customer loads may be understated relative to actual patronage – as a result, the crowding costs calculated using this approach may also be understated, though the impact is not expected to be material.

2. Approach to calculate crowding cost

The crowding cost is calculated using the following equation:

$$Crowding_{ijt} = \sum_{i=1}^x DL_{ijt} \times T_{ijt} \times M_{ijt} \times VTT$$

Where:

i = departure station

j = arrival station

t = rail service departure time

DL_{ijt} = departure load for each train departing station *i* towards station *j* at time *t*

T_{ijt} = in-vehicle travel time (IVT) between station *i* and station *j* for train departing at time *t*

M_{ijt} = crowding cost multiplier for journey segment *ijt*, for the percentage seated capacity

VTT = value of travel time

The crowding cost calculation uses parameters from the TfNSW EPV (2020) for both the crowding cost multiplier (derived from Douglas 2006, which estimates crowding cost multipliers for heavy rail) and the value of travel time, as presented in Table 1. Effectively, Table 1 shows the *M_{it} × VTT* component of the crowding cost calculation as the ‘Crowding cost per hour’, which would be applied to the passenger hours (*DL_{ijt} × T_{ijt}*) experienced on a service given the seated capacity (the number of passengers on the train relative to the number of seats available).

Prices are shown in 2019 dollars, matching the year of the ROAM sample data.

Table 1: Parameters used for calculating crowding impact (2019 dollars)

Item		Parameter Value
Travel time		Value of travel time per hour (VTT)
Value of travel time		\$17.72
Seated Capacity	Crowding cost multiplier (M)	Crowding cost per hour (M x VTT)
Under 80%	-	-
80% - 90%	0.01	\$0.18
90% - 100%	0.02	\$0.35
100% - 110%	0.05	\$0.89
110% - 120%	0.09	\$1.59
120% - 130%	0.15	\$2.66
130% - 140%	0.21	\$3.72
140% - 150%	0.29	\$5.14
150% - 160%	0.38	\$6.73
160% - 170%	0.48	\$8.51
170% - 180%	0.60	\$10.63
180% - 190%	0.72	\$12.76
190% - 200%	0.86	\$15.24
Over 200% (crush capacity)	1.25 ⁽¹⁾	\$22.15

Source: TfNSW EPV (2020) based on Douglas (2006). Note (1) The TfNSW EPV allows for a range of alternative crush capacity values to be used in CBA - this paper uses a 1.25 cost multiplier.

As noted in the TfNSW EPV (2020), alternative methodologies are available for estimating the cost of crowding beyond crush capacity.¹ Outside of Australia, other jurisdictions such as New Zealand use a fixed crowding cost multiplier that does not vary with seated capacity, or may only apply to certain trip types (Waka Kotahi NZ Transport Agency, 2021). Estimated expansion factors and total crowding would differ if costs were estimated using a different approach to that used in this paper.

2.1. Area assessed

The sample ROAM data covers the nine suburban and five intercity lines in the Sydney heavy rail network. All 170 Sydney Trains stations are included in the sample data, as well as an additional 118 stations from the Intercity Trains Network. This includes the full Blue Mountains, Southern Highlands, South Coast, and Central Coast lines, as well as the Hunter line between Telarah and Newcastle Interchange. Given the very low crowding cost estimates for the section of the Hunter line included in the analysis, it is likely that the sample data includes all sections of the network in which crowding costs are incurred.

Crowding cost was calculated for each individual rail service, on each track segment between two stations, for every service in the sample ROAM data, covering six full weekdays. For example, crowding cost would be calculated for a T1 Service departing Redfern Station at

¹ Refer to page 51 of the *TfNSW Economic Parameter Values*.

9:02am towards Central Station, based on the departure load patronage at Redfern, and the duration of in-vehicle travel time between the two stations. Crowding cost would then be re-estimated for the service as it leaves Central Station, accounting for changes in the departure load from boardings and alightings, and so on for all services, on all track-segments, across the entire data sample.

The data sample includes over 300,000 unique data points – that is, over 300,000 station ‘departures’ as measured when a unique train service leaves any station across the suburban or intercity network (such as the Tuesday 9:02am T1 from Redfern Station, discussed above). The patronage and crowding costs are estimated at each departure point, breaking a single rail service down into multiple station-to-station components, each with a unique patronage load and associated crowding cost. Using data at this level of granularity removes the need to average patronage across a service.

For additional granularity, the crowding cost calculation could further be broken down to the individual carriage level, if it was expected that certain carriages on each service experienced higher or lower levels of crowding than the average for the service. As carriage-level data is not available, this analysis implicitly assumes that passengers spread out evenly across the available carriages on each service.

3. Crowding cost estimates

Crowding costs on the Sydney suburban and intercity heavy rail networks were estimated based on ROAM data from a sample of six weekdays. Crowding costs were estimated for the entire 24-hour time period, and then split by time of day to allow for expansion factors to be calculated. The average crowding costs from the ROAM data are presented in Table 2 below.

Table 2: Average crowding cost by heavy rail line (2019 dollars)

Rail line	1hr AM peak	1hr PM peak	Daily total	Annual total*
Central Coast	2,700	3,020	8,360	2,108,000
South Coast	5,110	1,700	10,940	2,757,000
Southern Highlands	-	-	20	5,000
Blue Mountains	250	260	1,360	342,000
Hunter	30	-	40	11,000
T1 West	49,080	24,210	124,680	31,419,000
T1 North Shore	25,770	13,310	58,440	14,728,000
T2 Inner West	2,330	1,050	6,440	1,623,000
T2 Leppington	13,640	7,940	31,450	7,926,000
T3 Bankstown	12,230	1,970	19,860	5,005,000
T4 Illawarra	36,240	7,000	55,430	13,968,000
T5 Cumberland	3,050	460	4,650	1,173,000
T8 Airport & South	28,160	3,760	47,480	11,965,000
T9 Northern	26,400	8,600	50,890	12,824,000
Total	204,990	73,280	420,020	105,849,000

Note values may not sum due to rounding. Annual total estimated using an annualisation factor of 252.

Table 2 shows that the impacts of overcrowding on passenger amenity creates \$106m of disbenefits per year, with the majority of that from the T1 and T4 suburban lines (2019 dollars).

Crowding levels high enough to generate economic disbenefits were not seen on the Southern Highlands line in the sample ROAM data. Given the low total volume of trips, and the low crowding cost incurred on the intercity heavy rail lines, these intercity expansion factors should be treated with caution.

Crowding impacts were estimated for a sample weekend day and no service was found to experience patronage high enough to generate crowding disbenefits. Therefore, an annualisation factor of 252 should be used when estimating crowding disbenefits for Sydney, rather than the higher volume expansion factors per the EPV.

Crowding costs have been estimated based on the total number of passenger hours travelled at different levels of crowding, measured in terms of patronage as a percentage of seated capacity. At 100 per cent seated capacity, there are as many passengers on a train as there are seats (though many passengers remain standing even when seats are available, based on personal preference). Initial stated preference surveys undertaken in Douglas (2006) found that the discomfort of crowding begins to be perceived by travelers at 80 per cent seated capacity and increases exponentially until maximum capacity is reached.

Since crowding costs are based on the amount of time spent in crowded conditions, crowding costs are highest on the parts of the network leading into and out of the Sydney CBD. Crowding levels are highest within the CBD, although travelers generally spend only a short proportion of their total journey there. In contrast, travel between Strathfield and Redfern Station on the T1 line has a lower *level* of crowding, but contributes to a greater proportion of network crowding *costs* because of the longer station-to-station duration.

Variability between days in the ROAM sample data was low on the T1, T4, T8 and T9 lines, but higher on the T2 Leppington and T3 Bankstown reporting lines that have lower average patronage and crowding costs (Figure 1).

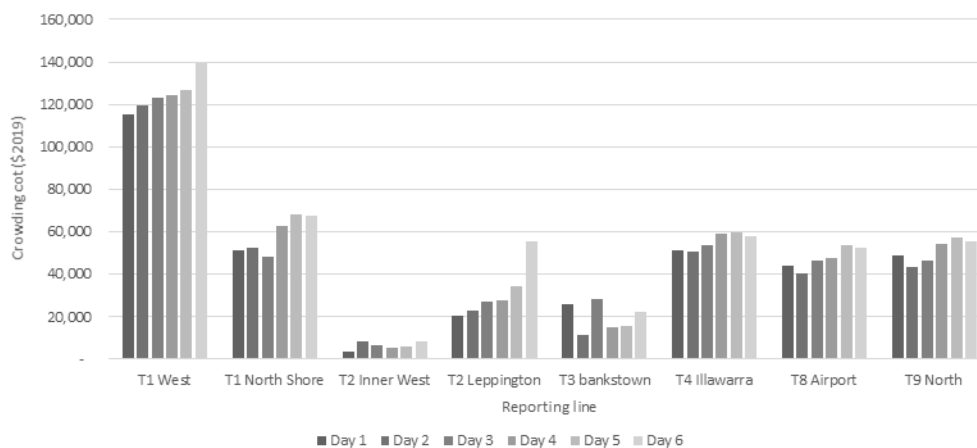


Figure 1: Crowding disbenefits by day, March 2019

It is important to note that this analysis only captures the discomfort to passengers directly impacted by crowded conditions. Crowding on rail lines often has a flow on impact to reliability or operating costs, and may influence some travelers to switch to other modes, increasing congestion elsewhere on the transport network (Tirachini 2013). These impacts are

not captured in the crowding cost estimates in Table 2, which as a result would understate the full cost of overcrowding to the NSW community.

3.1. Comparison to previous studies

This analysis of crowding costs differs to earlier cost estimates calculated in Wang & Legaspi (2012). In that analysis, crowding costs for 2011 across the Sydney CityRail network were estimated at \$85m in 2011 dollars, approximately equal to \$122m in June 2019 dollars. This is not directly comparable to the \$105m estimated for the suburban and intercity fleet estimated from ROAM data in this paper, based on several differences in source data and calculation approach.

Table 3: Parameters used for calculating crowding impact

Category	Wang and Legaspi (2012)	This paper
Rail network and timetable	2011 rail network and timetable, featuring different service frequencies and stopping patterns	March 2019 suburban and intercity heavy rail network and timetable (Metro services excluded)
Data source	Passenger counts undertaken as part of CityRail surveys, surveyed at CBD cordon stations and other selected stations from 19 October 2010 to 11 May 2011. Passenger loads at other stations interpolated or extrapolated from survey data	Sample of six weekdays from March 2019, extracted from ROAM, which assigns passengers to train services based on Opal ticketing 'tap on / tap off' data
Crowding cost calculation approach	Based on Douglas (2004) value of time estimates and Douglas (2006) crowding cost functions, including alternate costs per hour for passengers standing for longer than 10, 15 or 20 minutes	Based on simplified crowding cost functions derived from Douglas (2006) and reported in Transport for NSW <i>Economic Parameter Values</i> (2020). This approach uses a single, weighted cost per hour function, derived from the 10, 15, and 20 minute functions in Douglas (2006). The EPV cost function also has a higher disbenefit value for time spent at 'crush capacity'

Source: Wang, B. and Legaspi, J., 2012. Developing a train crowding economic costing model and estimating passenger crowding cost of Sydney CityRail network.

As a result of these differences, the crowding costs reported in this paper are likely to understate crowding costs relative to the calculation approach used in Wang and Legaspi (2012). However, the costs also reflect increased patronage and additional investment into the heavy rail network between 2010 and 2019, as well as differences in data collection approaches.

It is not fully known which of these factors contributes the most to the difference in crowding cost estimates. However, given the significant increase in patronage on the Sydney rail network between 2011 and 2019, it does support the hypothesis that the investments made into the rail network in that time have decreased or offset any increase in crowding costs.

4. Crowding cost expansion factors

Cost-benefit analysis of transport projects in Greater Sydney generally use either PTPM or ETCM to estimate the change in the amount of travel occurring at different levels of crowding from an investment or initiative. As current versions of these models only estimate impacts for peak periods, crowding costs are often quantified based on 1 hour or 2 hour AM peak outputs. Expansion factors are then used to estimate weekday and annual traffic based on modelled results for the AM peak period.

The main source of parameters for CBA in NSW, the TfNSW *Economic Parameter Values* (2020), does not currently report crowding expansion factors but recommends these factors to be calculated on a project-specific basis. This paper provides general cost expansion factors that can be used in CBA where project-specific information is not available.

The rail crowding expansion factors in Table 4 have been derived from the whole-of-network crowding costs calculated in this paper and reported in Table 2. Cost expansion factors are based on the relationship between the AM peak costs and the all day costs. For example, if the cost of crowding in the 1 hour AM peak was known to be half the total daily crowding cost, then an expansion factor of 2.0 could be used to expand 1 hour ETCM cost outputs to the all-day values.

Table 4 reports the daily and annual crowding costs for Sydney from 1 hour, 2 hour, or 3.5 hour AM peak periods estimated from the sample ROAM data.

Table 4: Rail crowding expansion factors by reporting line

Rail line	Cost expansion factor			
	1hr to daily	2hr to daily	3.5hr to daily	Daily to annual
Suburban network				
T1 West	2.54	1.81	1.49	252
T1 North Shore	2.27	1.62	1.49	252
T2 Inner West	2.77	1.55	1.42	252
T2 Leppington	2.31	1.66	1.53	252
T3 Bankstown	1.62	1.25	1.21	252
T4 Illawarra	1.53	1.26	1.25	252
T5 Cumberland	1.52	1.21	1.21	252
T8 Airport & South	1.69	1.23	1.18	252
T9 Northern	1.93	1.49	1.43	252
Total	2.03	1.51	1.39	252
Intercity network	1hr to daily	2hr to daily	3.5hr to daily	
Central Coast	3.09	2.34	2.26	252
South Coast	2.14	1.23	1.21	252
Southern Highlands				<i>No crowding cost</i>
Blue Mountains	5.46	3.90	2.72	252
Hunter	1.35	1.06	1.06	252
Total	2.56	1.71	1.66	252
Suburban and Intercity	1hr to daily	2hr to daily	3.5hr to daily	
Total	2.05	1.51	1.39	252

Note that Intercity expansion factors are higher in some instances because peak travel may occur outside the traditional 7:30am to 8:30am ‘peak’ 1 hour period used for consistency across the whole network. For example, crowding costs on the South Coast line was higher between 8:00am and 9:00am than it was for the 1 hour peak.

The expansion factor is calculated based on the relationship between the crowding cost in the peak time period and the all-day crowding cost. The 7:30am to 8:30am period is used for the 1

hour AM peak in the analysis - this definition differs slightly from other peak definitions, such as those used in ETCM, which measures all travel that occurs on rail services *arriving at central station* between 7:30am and 8:30am.

For this analysis, the 1 hour AM peak is measured from 7:30am to 8:30am, and the 1 hour PM peak is measured from 5pm to 6pm. As a result of peak spreading on the heavy rail and broader transport network, crowding disbenefits are also incurred outside of the 1 hour AM and PM peak periods.

Figure 2 shows the profile of crowding costs incurred by passengers across the suburban and intercity fleet, by time of day. Of note is that 34.5 per cent of all crowding-based discomfort experienced by Sydney Trains passengers occurs within the 8:00am to 8:30am window – more than the total cost of all of crowding that occurs after midday.

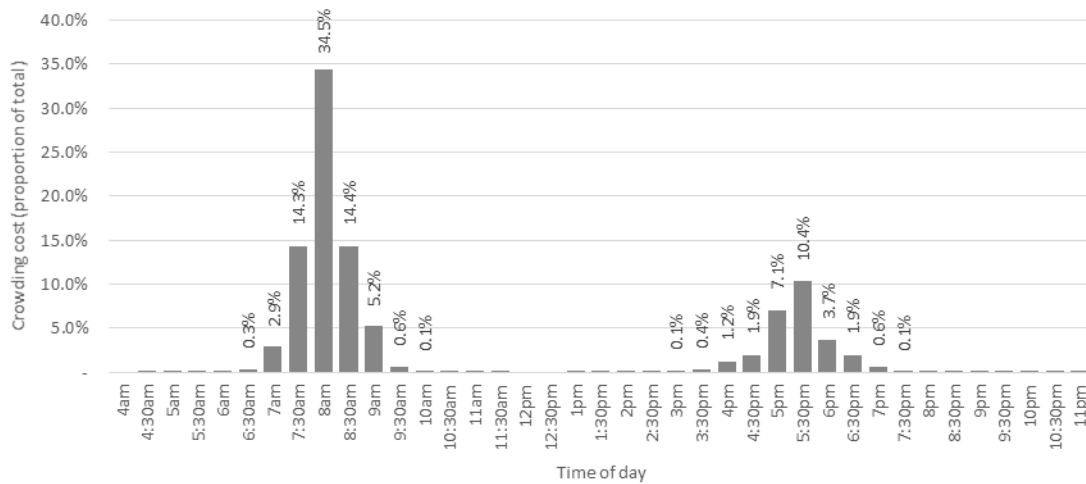


Figure 2: crowding costs per half-hour time period, March 2019

For comparison, Figure 3 shows the profile of heavy rail *boardings* across the suburban and intercity fleet, by time of day. While the 1 hour AM and the 1 hour PM peak account for 25.8 per cent of total boardings, they account for 66.3 per cent of network crowding costs.

This shows the level of service that can be supported across the Sydney heavy rail network before crowding costs begin to be incurred. In addition, it shows that increases in patronage above a certain level result in exponential growth in crowding costs, consistent with the much higher crowding costs at Fruin levels D and F found in Douglas (2006).

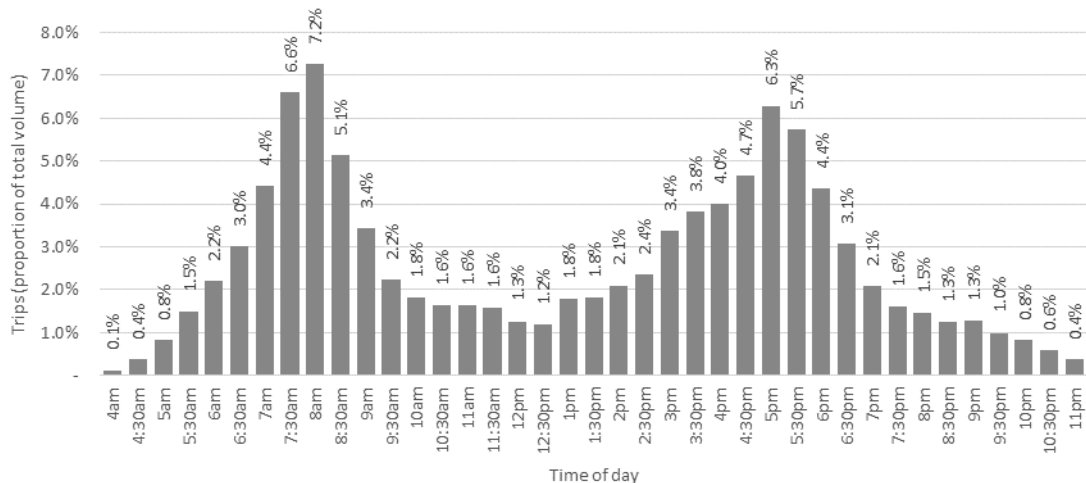


Figure 3: boardings per half-hour time period, March 2019

4.1. Analysis of results

As noted in Svanberg & Pydokke (2020), many of the studies that present define or measure crowding levels do not typically present solutions or policies to manage crowding. Conversely, studies and policies aimed at mitigating or preventing crowding impacts sometimes do not attempt to monetise these impacts, instead using ‘rule of thumb’ outcomes such as changes in occupancy, standing passengers, or headway (Cedar 2007).

Analysis of the relationship between volume (boardings) and total crowding costs across the sample data shows a positive correlation between the two indicators across Sydney’s heavy rail lines (Figure 4). However, increases in patronage do not fully explain the change in crowding cost per line, which is driven by the number of hours spent by passengers in crowded conditions. Boarding data, while more easily available, does not capture the length of time that passengers are exposed to overcrowding.

This shows the importance of understanding the underlying drivers of crowding-based discomfort, which become exponentially more uncomfortable to travellers as occupancy increases. It also demonstrates the importance of undertaking detailed cost-benefit analysis to determine whether interventions aimed at reducing crowding represent value for money, rather than targeting proxies such as occupancy levels or seating capacity.

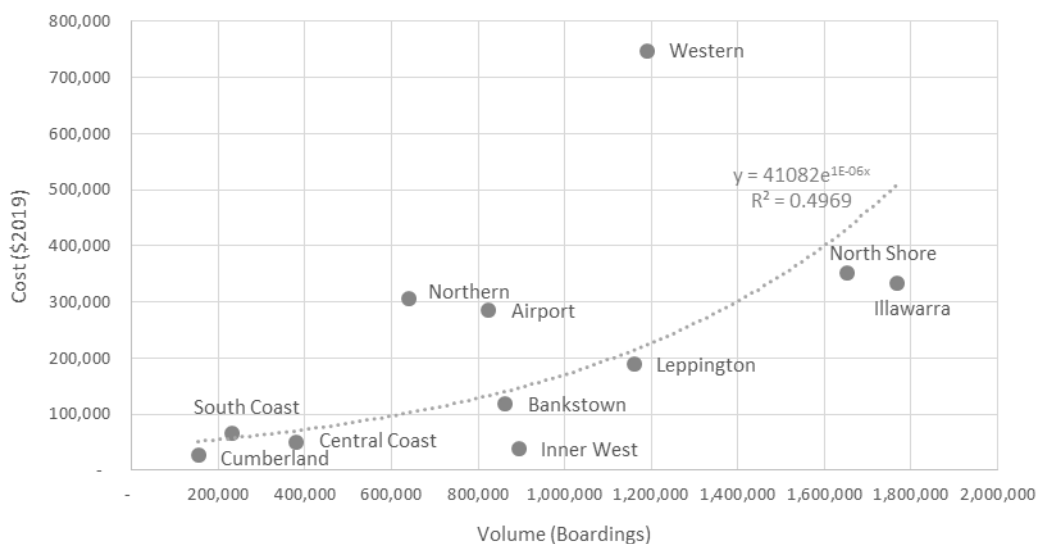


Figure 4: Relationship of trips to crowding cost, March 2019

The T1 Western Line is a significant cost outlier, driven by high levels of crowding on the Parramatta to Sydney CBD portion of the journey, which has both a high level of crowding and a long travel time. By contrast, analysis of the sample ROAM data found that most other services on the heavy rail network only reach significant levels of crowding much closer to the Sydney CBD, meaning that crowding costs are experienced for a shorter duration.

No clear relationship was found between volume or crowding costs and corresponding expansion factors (Figures 5 and 6). This may indicate that ‘peak spreading’ witnessed on busier lines is not from patronage diverted from the 1 hour peak to the peak shoulder, but instead a result of a broader increase in demand across the full peak and peak shoulder periods.

These results suggest that increases in crowding and patronage over time may not necessarily lead to a spread of demand from the peak to time periods outside of traditional peak periods,

so much as a broad increase across the full peak, and peak-shoulder periods. This could be the subject of further analysis and research in the future.

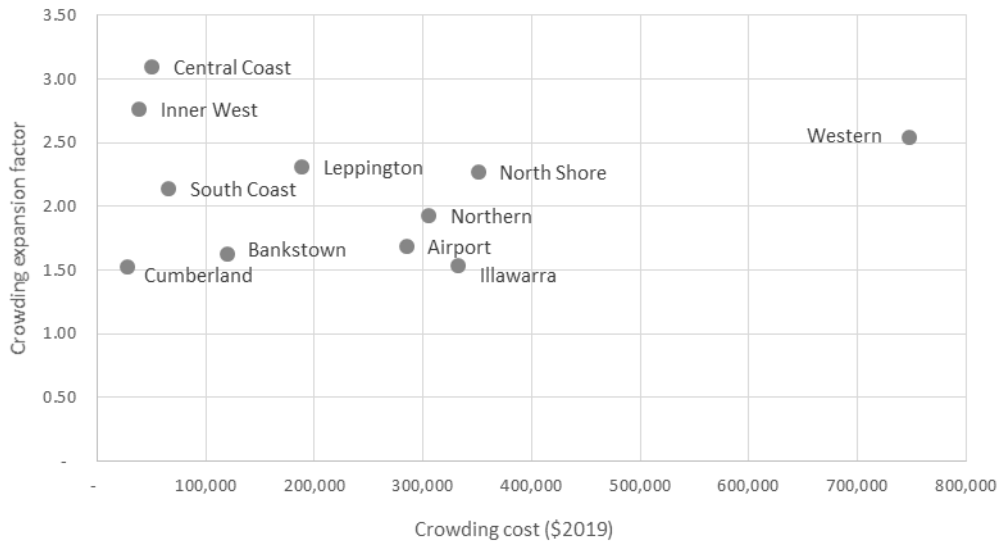


Figure 5: Relationship between crowding cost and crowding expansion factors, March 2019

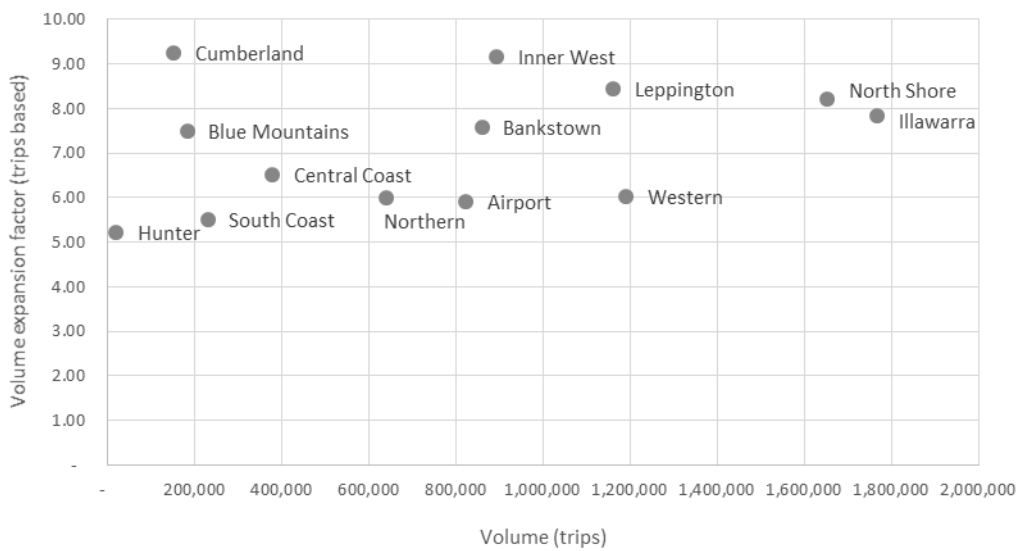


Figure 6: Relationship between trips and volume (trip based) expansion factors, March 2019

5. Suggestions for further work

This analysis was undertaken prior to the COVID-19 pandemic, and the associated impacts it has had on both patronage levels and crowding perceptions across Sydney. While the impact of the pandemic was relatively mild in Sydney by global standards, perceptions (and associated disbenefits) from crowding on public transport may have changed since the 2006 Douglas studies in response to additional health concerns. Future studies could focus on changes in the perceived cost of crowding experienced by passengers, perhaps through new willingness to pay studies.

In addition, analysis on ROAM data from early 2020 and 2021 could show the immediate and short-term impact of lockdowns, travel restrictions, and changed travel behaviour on crowding costs across the network.

Finally, this paper reports costs and expansion factors for the heavy rail network only. Sydney Metro's single-deck services were not included, as specific crowding cost functions for the Sydney Metro rolling stock (which have a different internal layout and seated capacity to the heavy rail fleet) have not yet been estimated. Once estimated, the analysis undertaken in this paper could be extended to include Sydney Metro services as well.

6. References

Ceder, A., 2007. Public Transit Planning and Operation Theory, modelling and practice. Civil and Environmental Faculty. *Transportation Research Institute, Technion–Israel Institute of Technology, Haifa*.

Douglas Economics, 2004, Value of rail travel time, a report to RailCorp.

Douglas, N. and Jones, M., 2016, November. Developing a suite of demand parameters for inner sydney public transport. In *Australasian Transport Research Forum 2016 Proceedings* (pp. 16-18).

Douglas, N. and Karpouzis, G., 2006, September. Estimating the passenger cost of train overcrowding. In *29th Australian Transport Research Forum* (pp. 1-8).

Douglas, N. and Karpouzis, G., 2005, October. Estimating the passenger cost of station crowding. In *28th Australasian Transport Research Forum, Sydney, Australia* (pp. 1-11).

Henn, L., Douglas, N. and Sloan, K., 2011. Surveying Sydney rail commuters' willingness to change travel time.

Svanberg, L. and Pyddoke, R., 2020. Policies for on-board crowding in public transportation: a literature review.

Tirachini, A., Hensher, D.A. and Rose, J.M., 2013. Crowding in public transport systems: effects on users, operation and implications for the estimation of demand. *Transportation research part A: policy and practice*, 53, pp.36-52.

Tirachini, A., Sun, L., Erath, A. and Chakirov, A., 2016. Valuation of sitting and standing in metro trains using revealed preferences. *Transport Policy*, 47, pp.94-104.

Transport for NSW, 2019. Transport for NSW Cost-Benefit Analysis Guide

Transport for NSW, 2020. Transport for NSW Economic Parameter Values

Waka Kotahi NZ Transport Agency, 2021, Monetised Benefits and Costs Manual

Wang, B. and Legaspi, J., 2012. Developing a train crowding economic costing model and estimating passenger crowding cost of Sydney CityRail network.