vlc

Veitch Lister Consulting Pty Ltd

Prepared for

NSW TREASURY

Final Report

CGC Recurrent Transport Methodology

February, 2019

Review of CGC's Transport Assessment Methodology

Final report

Project No. 18-041

Date	Revision	Prepared By	Checked By	Approved By	Description
17 Jan 2019	V1.0	Stuart D	Daniel V	Daniel V	DRAFT REPORT
10 Feb 2019	V2.0	Stuart D	Daniel V	Daniel V	FINAL DRAFT REPORT
14 Feb 2019	V3.0	Stuart D	Daniel V	Daniel V	FINAL REPORT
28 Feb 2019	V4.0	Stuart D	Daniel V	Daniel V	REVISIONS

Contents

List of Figures

List of Tables

Executive Summary

The NSW Treasury commissioned Veitch Lister Consulting (VLC) to review the Commonwealth Grant Commission's (CGC's) methodology for recurrent transport expenditure. In this report, we consider whether there is evidence that non-policy factors affect net PT expenditure per capita in the five largest capital cities. Non-policy factors are outside of the control of state governments, at least in the short-term, and include—but are not limited to—density, congestion, and geography.

Compared to the average city in our sample, we estimate non-policy factors add 38% to Sydney's net PT expenditure per capita. This effect arises from three sources, specifically:

- *PT supply* per capita in Sydney is approximately 33% higher than average, due to higher employment density and increased congestion.
- *PT productivity* in Sydney is approximately 3.3% lower than average, which stems from lower bus/tram speeds, shorter bus/tram routes, and longer heavy rail routes.
- *PT revenue* per capita in Sydney is approximately 37% higher than average, because of higher density and increased congestion.

The total effect of non-policy factors on net PT expenditure per capita are summarised in the following table for each city, where costs and revenues are baselined to 100 and 25, respectively.

City	Costs	Revenues	Net Expenditure	Effect
Sydney	137.78	34.30	103.48	*+38%*
Melbourne	118.40	30.23	88.17	*+18%*
SE Qld	79.31	18.97	60.33	*-20%*
Perth	50.46	12.20	38.26	*-49%*
Adelaide	71.02	17.25	53.78	*-28%*
Baseline	*100.00*	*25.00*	*75.00*	*0%*

In short, we find non-policy factors increase Sydney's net PT expenditure per capita by approximately 38% compared to the average of our sample, driving an 87% wedge between net PT expenditure per capita in Sydney vis-à-vis Perth. This is equivalent to Sydney needing to spend around $1.17 and $2.70 per capita on public transport for every $1 that is spent in Melbourne and Perth respectively.

As costs greatly exceed revenues, it is misleading to compare the percentage effect of non-policy factors on costs vis-à-vis revenues. When measured in percentage terms, revenue (37%) responds more strongly to non-policy factors than costs (33%). Why, then, do we suggest non-policy factors increase Sydney's net PT expenditure per capita compared to the average? Putting productivity effects to one side for now, the reason is that *costs are approximately three times larger than revenues* in all five capital cities. The imbalance between costs and revenues means that, when measured in monetary terms, non-policy factors lead to higher net PT expenditure per capita. Put another way, while non-policy factors lead to a larger percentage increase in revenue than costs, the monetary value of the extra revenue is insufficient to overcome the monetary value of the additional costs. For this reason, we find non-policy factors increase net PT expenditure per capita.

Using an independent methodology, we confirm and extend the main findings of consultants engaged by the CGC. Consultants engaged by the CGC also considered the effects of non-policy factors on net PT expenditure per capita. Using a more top-down (macroeconometric) approach, these consultants identified the effects of non-policy factors based on variation *between* approximately 70 "significant urban areas", or SUAs. Their proposed model suggests net PT expenditure per capita increases with density, average travel-to-work distance, slope, and train / bus patronage. We use an independent methodology that draws on more detailed data, which provides greater statistical power and identifies effects based on variation both *between* and *within* cities. Our methodology allows us to extend and elaborate the findings of the consultants engaged by the CGC, providing more detail on the underlying causes of variations in net PT expenditure between jurisdictions. Arguably, our findings are also more policy-neutral with regards to transport mode, as we do not explain net PT expenditure in terms of bus and rail patronage but rather fundamental non-policy factors, such as density, congestion, and geography. The primary downside to our bottom-up (microeconometric) methodology is that we use several models to estimate various (partial) effects, which must then be pieced together. In doing so, some additional assumptions are required.

Our estimates are expected to represent a lower-bound for the total effect of non-policy factors on net PT expenditure per capita. The bottom-up (microeconometric) methodology we use is expected to underestimate the total effects of non-policy factors for two reasons:

- First, we assume labour unit cost rates are independent of city size. In practice, we expect unit cost rates for labour will increase with city size, which will tend to increase Sydney's PT costs relative to other capital cities. Labour costs are a large component of overall PT costs.
- Second, we do not consider productivity effects on the size of the vehicle fleet. The effect of non-policy factors in Sydney, such as slower speeds and increased route kilometres, is likely to lead to an increase in the number of vehicles required to deliver PT services.

For these two reasons, we expect the true effect of non-policy factors on net PT expenditure per capita is larger than that explained by our analysis. That is, we are underestimating the degree to which non-policy factors place Sydney at a relative disadvantage compared to other cities.

Horizontal fiscal equalisation (HFE) supports the provision of PT services, controlling for differences in policy choices. In our view, the CGC's policy of HFE seeks to enable states to deliver levels of PT supply per capita, consistent with demand. We find strong evidence PT supply, productivity, and revenue are affected by non-policy factors, such as density, congestion, and geography. Taken together, our results imply non-policy factors add 38% to Sydney's net PT expenditure per capita compared to the average of the five capital cities that we analyse. Our results imply, for example, the presence of an 87% wedge in net PT expenditure per capita between Sydney and Perth. And for the reasons noted above, we expect the differences in net PT expenditure between cities that we identify will likely underestimate the total effect of non-policy factors. In our view, our findings—when considered in the context of HFE—imply the CGC's chosen methodology for recurrent transport expenditure should produce differences in net PT expenditure per capita between cities that are similar to, if not larger than, those identified by our analysis.

1. Introduction

NSW Treasury commissioned Veitch Lister Consulting (VLC) to review the Commonwealth Grant Commission's (CGC's) methodology for recurrent transport expenditure. In this report, we consider whether there is evidence that non-policy factors affect net public transport (PT) expenditure per capita in Sydney vis-à-vis other large capital cities in Australia.

The following sections of this chapter are structured as follows:

- **Section 1.1** introduces policy principles that guide the Commission's activities, and which have informed our work, specifically the concept of horizontal fiscal equalisation.
- **Section 1.2** outlines our methodology in contrast to the work undertaken by consultants engaged by the CGC and presents theoretical foundations for our approach.
- **Section 1.3** describes the main types of econometric (empirical) models that we estimate and discusses the key metrics by which we evaluate their results.
- **Section 1.4** discusses the non-policy factors that are the focus of our analysis.

1.1 Policy Principles

The CGC's approach to funding state governments is designed to deliver on the general principle of "horizontal fiscal equalisation" (HFE), which is defined as follows: *"State governments should receive funding from the Commonwealth such that, if each made the same effort to raise revenue from its own sources and operated at the same level of efficiency, each would have the capacity to provide services at the same standards."*[1]

Three aspects of this definition of HFE are relevant to analysis of recurrent transport expenditure and have directly informed our methodology, specifically:

- Providing services "*at the same standards*" reflects the CGC's intention to fund levels of **PT supply** per capita that are consistent with underlying demand on a policy neutral basis;
- Operating "*at the same level of efficiency*" reflects the CGC's desire not to compensate states for differences in **PT productivity** arising from policy choices, such as ticketing systems; and
- Investing the same effort to "*raising revenue from its own sources*" reflects the CGC's desire not to compensate states for differences in **PT revenue** per capita arising from policy choices.

We appreciate the Commission's desire to equalise funding to deliver a consistent standard of supply while controlling for differences in productivity and revenue collection.

In practice, the situation is somewhat complex. PT supply, productivity, and revenue are all affected by a combination of *policy choices* and *non-policy factors*. As an example, PT productivity is affected by policy choices, such as stop-spacing and ticketing systems, while also being affected by non-policy factors, such as density, which are beyond the government's direct control.

In this context, our work seeks to disentangle the effects of policy choices and non-policy factors on net PT expenditure per capita. Building from the concept of HFE, we consider three channels through which non-policy factors may affect net PT expenditure: namely supply, productivity, and revenue. Our methodology is discussed in more detail in Section 1.2.

1.2 Methodology

1.2.1 Objectives and Data

Our methodology was developed with the following two objectives in mind: *(1) provide independent verification* of the non-policy factors identified by the consultants engaged by the CGC (c.f. Appendix

---

[1] Commonwealth Grants Commission, Report on State Revenue Sharing Relativities 2002 Update, p. 5

A) and *(2) consider additional evidence* for the role of other non-policy factors, such as the channels identified in the conceptual framework developed by Treasury NSW (c.f. Appendix B). Put simply, our methodology seeks to independently verify and extend the findings of consultants to the CGC. Our work is presented as a complement to, rather than a substitute for, these findings.

Unlike the CGC, we do not have access to data on net PT expenditure for "Significant Urban Areas" (SUA). Instead, we make use of more granular microdata, which confers two key advantages:

- **Statistical power.** Microdata provides us with a larger sample, which serves to increase the statistical significance of our findings and enables us to test a larger number of variables[2]; and
- **Stronger identification.** Microdata enables us to identify the effects of non-policy factors from variation that exists both *within* and *between* capital cities.[3]

The major limitation of our methodology is that we derive partial results from several econometric models, which must then be pieced together using additional assumptions. In contrast, the approach used by the CGC's consultants provides a simpler link between non-policy factors and net PT expenditure per capita. We suggest there is merit in both our bottom-up, microeconometric approach and the more top-down, macroeconometric approach used by the CGC's consultants.

1.2.2 Theoretical Foundations

Consultants engaged by the CGC estimated models that took net PT expenditure per capita as the dependent variable. In contrast, we treat net expenditure E as the result of two semi-independent (albeit linked) economic outcomes, namely gross costs, C, and fare revenue, R. That is,

$$E = C - R$$

To ease exposition, let us ignore differences by mode and city for the present time. We can then further decompose gross costs, C, into three major resource inputs:

- *Vehicle-hours* (h), which capture time-related costs, e.g. driver wages;
- *Vehicle-kilometres* (k), which capture distance-related costs, e.g. maintenance and fuel; and
- *Vehicles* (v), which capture vehicle-related costs, e.g. fleet and depots.[4]

To arrive at total costs, each resource input is multiplied by its unit cost (γ_j) and summed, formally

$$C = \gamma_1 h + \gamma_2 k + \gamma_3 v$$

Where γ_1, γ_2, and γ_3 denote unit cost rates for vehicle-hours, vehicle-kilometres, and vehicles, respectively, which we presume to be exogeneous, and h, k, and v are as previously defined.

Intuitively, resource inputs will increase with demand, D. Moreover, resource inputs will be affected by policy choices, such as the distance between stops, and non-policy factors, such as geography. We denote policy choices and non-policy factors by the vectors X and Y, respectively. Together, this implies our gross cost function can be expressed as follows:

$$C = \gamma_1 h(D, X, Y) + \gamma_2 k(D, X, Y) + \gamma_3 v(D, X, Y)$$

---

[2] Our PT productivity models, for example, use GTFS data on individual vehicle trips in five major capital cities and has sample sizes in excess of 100,000 observations. In contrast, the consultants engaged by the CGC were limited to a sample of approximately 70 SUAs.

[3] The findings of consultants to the CGC relies solely on variation between SUAs to identify the effects of non-policy factors. In some respects, identifying effects based on *within-city* variation is preferable to *between-city* variation, as the latter may be biased due to unobserved city-specific effects. In our case, we use both.

[4] Due to data limitations, we do not model the effects of non-policy factors on vehicle requirements. We note, however, that these effects are likely to be positively associated with vehicle-hours and vehicle-kilometres.

We also assume $R = r(D, X, Y)$. That is, revenue is a function of demand, D; policy choices, X; and non-policy factors Y, respectively. Substituting these expressions into net expenditure yields:

$$E = C - R = \gamma_1 h(D, X, Y) + \gamma_2 k(D, X, Y) + \gamma_3 v(D, X, Y) - r(D, X, Y)$$

Differentiating the above expression with respect to non-policy factor Y_j allows us to isolate the effect of the latter on net expenditure as follows:

$$\frac{\partial E}{\partial Y_j} = \left(\gamma_1 \frac{\partial h(\cdot)}{\partial Y_j} + \gamma_2 \frac{\partial k(\cdot)}{\partial Y_j} + \gamma_3 \frac{\partial v(\cdot)}{\partial Y_j}\right) - \frac{\partial r(\cdot)}{\partial Y_j}$$

This theoretical expression has a simple interpretation: The effect on net PT expenditure, E, of a small change in non-policy factor, Y_j, is the sum of its effects on costs (hours, distance, and vehicles) minus its effects on revenue. Our approach, then, estimates the effects of non-policy factors on costs and revenue separately before then combining these individual effects to estimate the total effect.

1.3 Econometric Models

Here we outline our three model specifications, discuss how we interpret results, and present model variants. The purpose of this section is to help the reader interpret results in Sections 2, 3, and 4.

1.3.1 Scope

As per the discussion of HFE in Section 1.1, we present three broad model specifications, specifically:

- **Supply (or cost) models,** which analyse the effects of non-policy factors on the quantity of PT services delivered in Australian capital cities, measured at the level of SA2s;
- **Productivity models**, which analyse the effects of non-policy factors on the efficiency with which PT services operate in Australian capital cities, measured at the route-level; and
- **Revenue models,** which analyse the effects of non-policy factors on revenue from PT services in Sydney, also measured at the level of SA2s.

Our supply and productivity models capture the effects of non-policy factors on costs, C, and use data sourced from Google Transit Feed Specifications (GTFS) and the Census for the five largest capital cities, as defined by ABS's Greater Capital City Statistical Areas (GCCSA). Our revenue model, in contrast, makes use of (confidential) Opal data for Sydney.

1.3.2 Interpretation

When interpreting the results of our models, we focus on the following three aspects:

- **Direction of parameters.** Whether the sign (positive or negative) of the estimated parameters align with our prior expectations.
- **Statistical significance.** The probability we can reject the hypothesis that model parameters equal zero (that is, non-policy factors are not associated with supply, productivity, or revenue).
- **Model fit.** The degree to which the model is a reasonable representation of the underlying data, considering overall explanatory power and the presence of extreme values.

The direction of parameters and, to a lesser extent, model fit is readily evaluated from regression output. Understanding the statistical significance of estimated parameters is, however, more complex because it requires assumptions on the distribution of the model's residuals. We discuss the issue of residuals, or standard errors, in more detail in Section 1.3.3 below.

1.3.3 Model Variants

In the following sections we generally present results for three model variants. Our first two variants estimate s.e. using the following assumptions:

- **Variant 1 (V1)** estimates "heteroskedasticity-robust" s.e. which allow for non-constant variance. In the presence of heteroskedasticity, s.e. will be incorrectly estimated; and
- **Variant 2 (V2)** estimates "cluster-robust" s.e., which also allows for correlation between groups in the data. In our models, we typically cluster s.e. by city.[5]

Model variants V1 and V2 are estimated using Ordinary Least Squares ("OLS"). In contrast, model variant (V3) makes use of Weighted Least Squares (WLS).

The reason we complement variants V1 and V2 with WLS is because *not all observations in our data are deemed to be equally important*. The GTFS data used in our productivity models, for example, groups identical PT services by trip-ID. Whereas some trip-IDs operate only once per week (for example, 10am on a Sunday), others operate seven days per week (for example, 10am every day). In this case, it seems reasonable to suggest the more frequently operated trip-IDs are more important.

For this reason, we estimate a model variant using WLS where individual observations are weighted based on their relative importance. Our choice of weights varies by model, specifically:

- For models where observations are defined by SA2s, such as the supply and revenue models presented in Sections 2 and 4, respectively, we weight observations by the total population and/or employment of each SA2, which gives more weight to more urbanised SA2s; and
- For models where observations are defined by trip-IDs, such as the productivity models presented in Section 3, we weight observations by the number of trips operated per week, which gives greater weight to trip-IDs that operate more frequently.

For the supply and revenue models presented in Sections 2 and 4, model variant V2 (cluster-robust OLS) is preferred. We prefer model V2 over model V1 (robust OLS) because the latter does not account of clustering and typically overestimates the statistical significance of parameters (that is, underestimates standard errors). Similarly, we prefer model V2 over model V3 (cluster-robust WLS) because the weighting scheme in the latter tends to "double-count" the effects of density.

For Section 3, however, we prefer model variant V3 (cluster-robust WLS). We prefer model V3 over model V1 (robust OLS) for the same reason noted above, that is, V1 does not account for clustering and typically underestimates standard errors. In this case, however, we prefer model V3 over model V2 (cluster-robust OLS) because the weighting scheme used in the former, specifically trip frequency, adds new and useful information on the relative importance of individual trip-IDs.

Notwithstanding our preferences for models, we typically find consistent results across all three model specifications, which provides us with greater confidence in the robustness of our results.

1.4 Non-policy Factors

Our work seeks to identify the effect of non-policy factors on net PT expenditure per capita. In this section, we discuss non-policy factors and our choice of indicators. We are interested in non-policy factors that are largely beyond the control of state governments, at least in the short-run, and which lead to differences in net PT expenditure per capita.

1.4.1 Definition

Based on the work undertaken by consultants to the CGC, discussions with NSW Treasury, and our own experience, we identified the following non-policy factors are being of greatest relevance:

- **Transport outcomes,** such as road congestion and travel distances;
- **Economic geography**, such population/employment density and urban form; and

[5] As first noted by Moulton, econometric models where observations are based on geographic units will often have errors correlated across units. In most cases, failing to control for correlations between groups in the data will likely understate the s.e. (and, by extension, overstate the statistical significance of parameter estimates).

- **Physical geography**, such as barriers arising from water features and terrain.

There are no perfect measures for these non-policy factors. As with all empirical analyses, we rely on imperfect indicators that capture the most relevant, or salient, aspects of the factors in question. From a purely econometric stand-point, the presence of (unbiased) measurement error will tend to reduce the magnitude and significance of the effects we can identify.

1.4.2 Indicators

Our indicators are intended to capture the salient aspects of non-policy factors. The supply and revenue models in Sections 2 and 4, respectively, include the following two measures:

- **Density**, as measured by the number of residents or jobs within a certain area, and
- **Congestion**, as measured by daily delay hours incurred by vehicles.

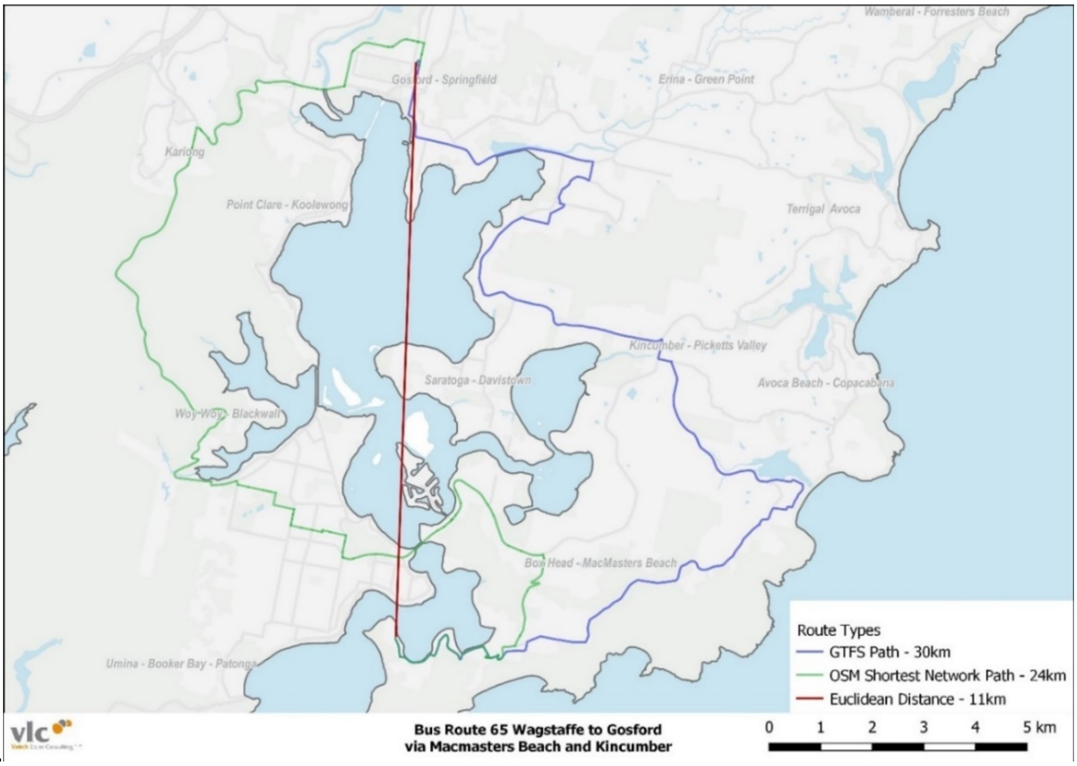
Both the supply and revenue models include the same non-policy factors.

The productivity models presented in Section 3 operate at the level of individual trip-IDs. In addition to density (or catchment) and congestion, we also test additional non-policy factors, including:

- **Geographical deviation**, which is designed to measure the effect of geographical barriers, such as harbours. We define the indicator, g_i, as (1) the shortest network distance minus (2) the Euclidean distance, as measured from the start to the end of the route; and
- **Absolute change in elevation**, which is an indicator of hilly terrain. We define the indicator, z_i, as the sum of the (absolute) changes in vertical elevation between stops along the route, where elevation data is sourced from SRTM (NASA).

Geographical deviation considers the shortest-distance when travelling via the *road network* between the start and end of a route, rather than the distance travelled by the PT route. To illustrate the difference between different measures, Figure 1 shows the route taken by bus route 65 in Sydney.

Figure 1: Geographical Deviation for a Bus Route in Sydney



The actual GTFS path (route length) will be affected by planning decisions that favour more circuitous PT routes than necessary, for example to increase population catchment. While warranted, such decisions are not attributable to geography. That is why we measure geographical deviation as the OSM shortest network path less the Euclidean distance.

2. Supply Model

Summary: We model the effect of non-policy factors on PT supply, or gross costs, in the five largest capital cities in Australia, as defined by ABS Greater Capital City Statistical Areas, or GCCSAs. We estimate PT supply for individual SA2s in terms of seat-kilometres. Results suggest non-policy factors, specifically density and congestion, have a positive effect on PT supply. Specifically, we estimate non-policy factors increase Sydney's net PT supply by 33% compared to the average.

2.1 Model

As noted previously, unlike the consultants engaged by the CGC, we do not have access to information on the gross costs of PT. Instead, we must approximate how costs vary within and between cities using readily-observed data.

We propose to approximate PT costs using *total seat-kilometres (seat-km)*. To estimate seat-km S_i in SA2 i, we multiply the number of vehicle kilometres k_i^m for each mode m with the seated vehicle capacity C^m of that mode. Formally:

$$S_i = \sum_m C^m k_i^m$$

Seat-kms satisfies two criteria: First, it is mode-neutral and, second, it is an accurate indicator of the costs of PT supply. In terms of the second criteria, we find an extremely high positive correlation (0.983) between seat-hrs and seat-kms at the SA2 level. This correlation, as well as our chosen model specification, which is discussed below, implies non-policy factors will have similar effects on seat-hrs as seat-km. While we do not have data on vehicle requirements, which is the third major cost driver noted in Section 1.2.2, we expect they will be determined largely by seat-kms and seat-hrs. For these reasons, seat-km provides a mode-neutral and accurate measure of the costs of PT supply.

That said, we note two limitations to the use of seat-km:

- First, as ferries do not operate on land, we exclude them from our measure.[6] We do not consider this to be a major issue given the small role played by ferries in most cities; and
- Second, to calculate seat-kms, we must make additional assumptions about the average vehicle capacity, C^m, for each PT mode that operates in each city.

We estimate vehicle capacities for each PT mode in each city from VLC's strategic transport models for 2016, as summarised in Table 1. These numbers denote approximate averages for each mode and city; actual capacity will vary depend on the rolling stock used to operate individual trips. Later in Section 2.4, we test the sensitivity of our results to these assumptions.

Table 1: Assumptions for Seated Public Transport Vehicle Capacities

City	Mode		
	Bus	Tram	Heavy rail
Sydney	52.5	239	1,165
Melbourne	50	115	875
SEQ	55	192	500
Perth	55	N/A	500
Adelaide	55	100	280

[6] The issue is that most ferry kilometres fall outside of the SA2 that benefit from the service. To resolve the issue, one could calculate ferry route kilometres along the entire route and then use a rule to assign these kilometres to the SA2s that are serviced by the ferry, i.e. where it stops. Rules could assign all kilometres to the SA2 where the service originates or assign proportionally based on the number of stops (usually only two).

With these capacities we can calculate our dependent variable S_i, and specify a basic model of PT supply. To start, we assume PT supply responds to density and congestion levels as follows:

$$S_i = F_i^q d_i^{\alpha_1} c_i^{\alpha}$$

Where d_i and c_i denotes density and congestion, respectively, and α_1 and α_2 denote parameters to be estimated. Taking logs yields an equation that is linear in parameters, formally:

$$\log S_i = \log F_i{}^q + \alpha_1 \log d_i + \alpha_2 \log c_i = f_i + \alpha_1 \log d_i + \alpha_2 \log c_i$$

Our priors are that PT supply increases with density and congestion, that is, $\alpha_1, \alpha_2 > 0$.

The constant $\log F_i^q = f_i$ is a "supply shifter", or fixed effect, which captures average differences in levels of PT supply between the SA2s in area q. In the models below, we define q to be SA3s. Including SA3 fixed effects f_i means the effects of density and congestion levels are identified from variation *between SA2s within an SA3*. In this way, we use SA3 fixed effects to control for unobserved determinants of PT supply, such as infrastructure, urban form, and policy choices. Put another way, we use fixed effects to capture differences in average levels of PT supply within SA3s that are not explained by the SA2-level explanatory variables in our model, namely density and congestion.[7]

Finally, one of the advantages of using a log-log model is the resulting estimates for α_1 and α_2 can be interpreted as "constant elasticities". These parameters provide a scale-invariant measure of the effects of explanatory variables that translates readily into relative percentage effects. This is important because, as noted above, our model uses seat-km as a proxy for total costs. Hence, we assume a percentage change in seat-kms translates into the same percentage change in costs.

2.2 Data

Our data was generated as follows:

- First, S_i is estimated by assigning route-kilometres in GTFS data to individual SA2s. For each mode and SA2, we then multiplied route kilometres by the vehicle seat capacities in Table 1;
- Second, we linked data on seat-km to Census data on the density of SA2s, such as population, employment, and area; and
- Third, we extracted data on SA2 vehicle delays from VLC's strategic transport models.[8]

Summary statistics for key variables are summarised in Table 2, where each row relates to individual capital cities and the final row presents the average for the sample.

Table 2: Summary Statistics – Averages for all SA2s

City	n	PT supply (S_i)	Population (p_i)	Employment (e_i)	Congestion (c_i)	Area A_i
SYD	308	10,025,876	15,662	7,172	1,763	31.74
MEL	300	6,060,838	14,731	6,743	1,588	26.72
SEQ	284	1,681,872	10,449	4,674	871	40.51
PER	163	2,442,383	11,904	5,118	800	29.99
ADL	101	836,054	12,654	5,449	1,266	24.65
Sample	*1,156*	*5,074,760*	*13,347*	*6,007*	*1,319*	*31.73*

[7] We are unsure why the consultants engaged by the CGC did not test models with fixed effects at the state level. This would ensure the effects they identify are not caused by average differences in net PT expenditure between states that happen to be correlated with the explanatory variables included in their model.
[8] Traffic congestion (specifically, total private vehicle delay hours for an average weekday in 2016) is extracted from VLC's Zenith models for each capital city at the SA1 level and then aggregated to SA2s.

Inspection of the data indicated ABS's definition of each of the five GCCSAs includes a small number of extremely large SA2s. These areas are typically associated with national parks and other natural areas that have extremely low population densities. These SA2s have a large effect on the average area of SA2s reported in Table 2, thereby distorting measures of density.

To ensure our analysis reflects PT supply in primarily urban areas, we apply a population density filter to remove these large, non-urban SA2s from our data. This raises the question of how to define "urban"? Various definitions of urban areas exist in the literature, many of which are relatively complex. The ABS, for example, use several criteria to define urban centres, including but not limited to "*SA1s are considered to be 'urban' if they … have a population density greater or equal to 100 persons per sq km.*"[9] Uchida and Nelson (2010) test the effects of population density thresholds of 150, 300, and 500 people per square kilometre for cities in the OECD.[10]

In our data, SA2s have already been designated by the ABS as part of a GCCSA. Moreover, we wish to maintain as large a sample as possible. All other things equal, we prefer a lower density threshold. As a starting point, we decide to follow the ABS's criteria for SA1s, that is, we apply a population density filter of $\tau = 100$ residents per square kilometre. This had the effect of excluding 98 non-urban SA2s. Summary statistics for the data with the density threshold applied are presented in Table 3. While most summary statistics barely change, the average area statistics are much reduced.

Table 3: Summary Statistics – Averages for Urban SA2s ($\tau = 100$)

City	n	PT supply (S_i)	Population (p_i)	Employment (e_i)	Congestion (c_i)	Area A_i
SYD	280	10,530,842	16,693	7,455	1,783	11.15
MEL	279	6,337,540	15,301	6,897	1,612	14.12
SEQ	267	1,744,157	10,495	4,724	890	13.72
PER	141	2,635,053	13,341	5,094	851	15.53
ADL	91	904,045	13,516	5,795	1,351	12.77
Sample	*1,058*	*5,327,325*	*14,042*	*6,161*	*1,351*	*13.30*

We estimate models using the data presented in Table 3. We also tested the sensitivity of our results to alternative density filters, finding no significant effect on our results.

2.3 Results

In this section we develop our baseline PT supply model. The first question we answer is how to define density. Several alternatives exist, which differ along the following two dimensions:

- **Measure**, specifically population (p_i) or employment (e_i); and
- **Method**, specifically average or weighted.[11]

Four possible density definitions can be constructed from these two dimensions, as summarised in Table 4 along with their respective formulae. The consultants engaged by the CGC used population-weighted density, which is the definition specified in the top-right cell of Table 4. As noted by the consultants themselves, weighted population density arguably has the advantage of being more

[9] www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1270.0.55.004~July%202016~Main%20Features~Design%20of%20UCL~8

[10] www.wider.unu.edu/sites/default/files/wp2010-29.pdf

[11] Weighted density W_j for each SA2 i is calculated as follows: (1) Calculate average density p_q / A_q or e_q / A_q for each ABS Destination Zone (DZN) $q \in i$; (2) Weight the resulting DZN densities by population p_q or employment e_q; (3) sum over all DZNs in SA2 i, and (4) Divide by the SA2 population p_i or employment e_i.

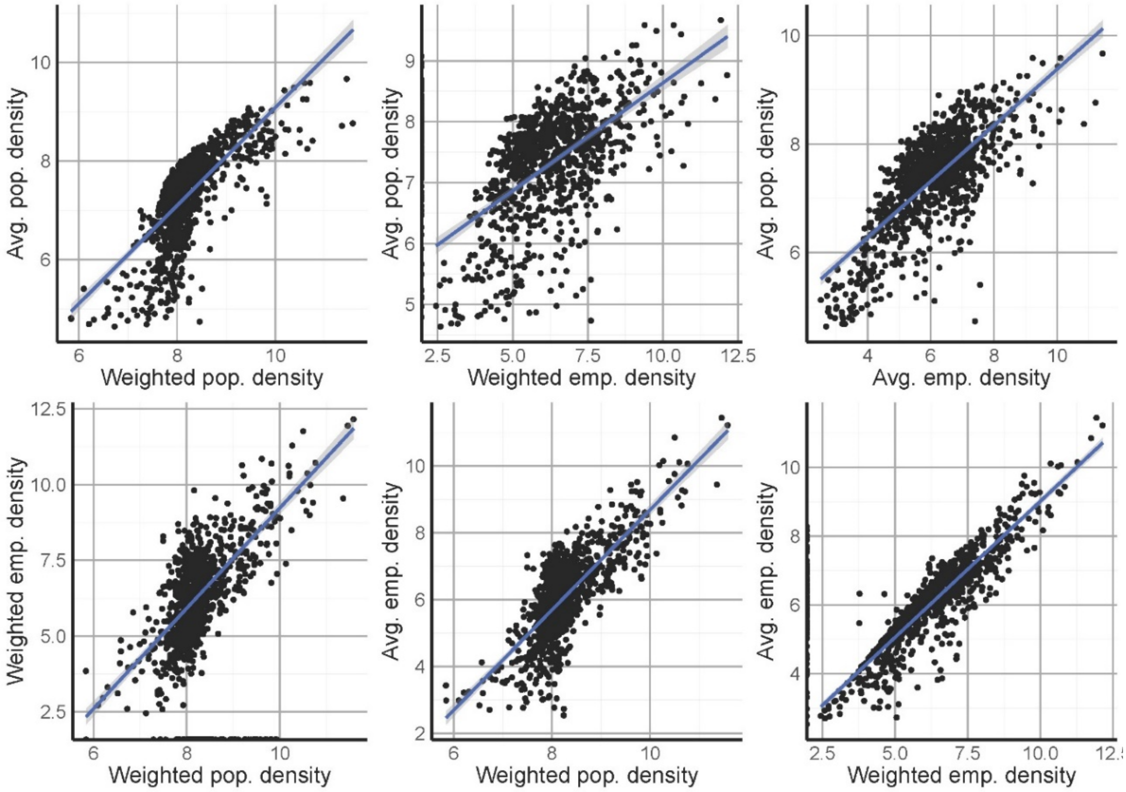
representative of the density that is "experienced" by the average resident. On the downside, weighted population density is more difficult to interpret than average density.[12]

Table 4: Alternative Density Measures – Comparing formulae

		Method	
		Average	**Weighted**
Measure	**Population**	$\frac{p_i}{A_i}$	$\frac{\sum_{q\in i} p_q^2/A_q}{p_i}$
	Employment	$\frac{e_i}{A_i}$	$\frac{\sum_{q\in i} e_q^2/A_q}{e_i}$

Figure 2 presents scatter plots for all six pair-wise combinations of the four alternative density definitions. As expected, all six combinations reveal a strong positive correlation. This suggests including multiple measures of density within a regression model may give rise to multi-collinearity, which in turn may thwart our ability to identify the effects of individual density measures.

Figure 2: Alternative Density Measures – Scatter plots



For weighted employment density, the ABS TableBuilder appears to report zeros for some Destination Zones (DZNs) that should clearly have employment. This problem is most severe in the suburbs in Sydney and Melbourne. Due to issues with the reliability of employment data, we do not pursue the use of weighted employment density. In terms of evaluating which of the other three density measures should be retained in our model, we apply the following three-step process:

1. We first estimate a model that includes the remaining three definitions of density (i.e. excluding weighted employment density due to data reliability issues), specifically (1) average population density, (2) average employment density, and (3) weighted population density.

[12] To see this, readers may want to try and take the derivative of weighted population density with respect to the Destination Zone (DZN) population p_q, noting that $p_i = f(p_q)$. The resulting expression is complex and non-linear. In contrast, the derivative of both average measures is simply $1/A_i$, which is constant.

2. If we find that an individual definition does not have the expected sign and/or is insignificant in model V2, then we remove this variable and re-estimate a restricted model
3. We repeat step 2 until all remaining density definitions in the model have the expected sign and are statistically significant.

The process lets the data speak while ultimately converging to the simplest model possible. Adopting this process saw us converge to a model that included only average employment density. Results for this model are summarised in Table 5.

Table 5: Regression Results – Supply Model

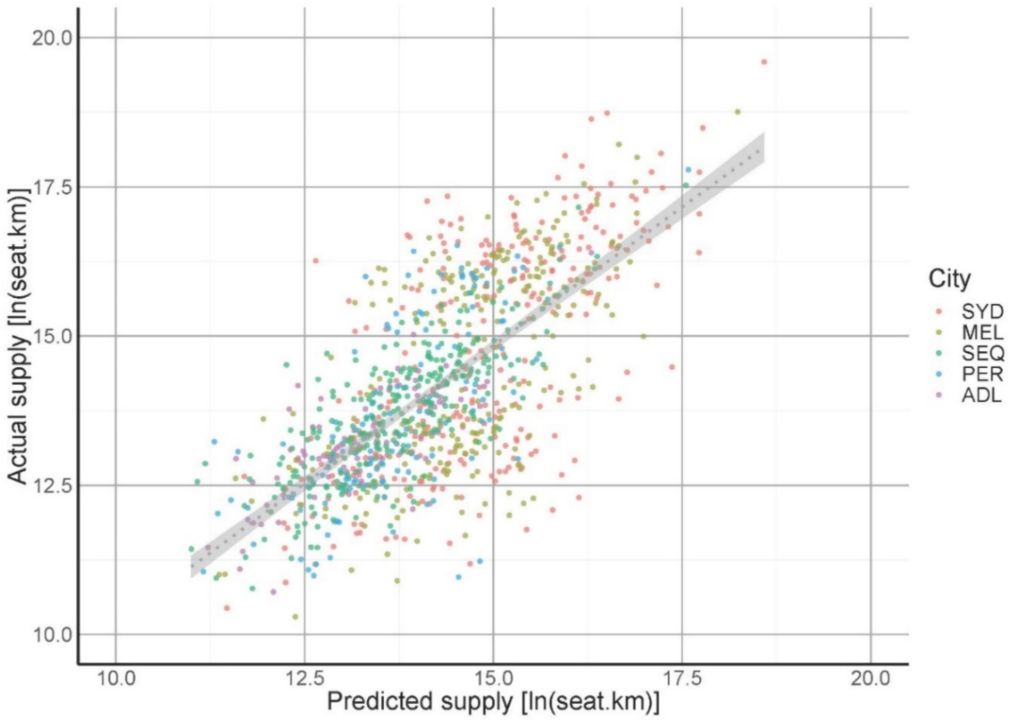
Parameters	SA3 Fixed Effects		
	V1: Robust	**V2: Cluster**	**V3: WLS**
log(emp. density)	0.55 (0.05)***	0.55 (0.06)***	0.66 (0.05)***
log(delays)	0.42 (0.05)***	0.42 (0.08)***	0.44 (0.06)***
R^2	0.53	0.53	0.63

Notes: All models include SA3 f.e. and have n = 1,058 obs.
Standard errors in brackets; $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

For reasons discussed in Section 1.3.3, we prefer model V2.[13] As expected, non-policy factors have a positive and statistically significant effect on PT supply. Specifically, we find the following elasticities: Employment density 0.55 – 0.66 and Car delays of 0.42 – 0.44.

In terms of model fit, the association between actual and predicted levels of PT supply is illustrated in Figure 3 for model V2 (preferred model), where points are coloured by city.

Figure 3: Supply Model – Model Fit (V2)



[13] We estimate models V1 and V2 using OLS and with robust and cluster-robust standard errors, respectively. Model V3 is estimated using WLS with cluster-robust standard errors. In model V3, our weights are the sum of population and employment in an SA2, which means more urbanised SA2s have more influence on parameters.

Promisingly, Figure 3 reveals a positive association between actual and predicted supply, with points clustered around the diagonal and no obvious extreme values. We find R-squared values of 0.53 and 0.63 for models V1/V2 and V3, respectively.

Finally, we note that while we find employment density is a stronger predictor of PT supply than population density at the SA2 level. This does not mean the latter is irrelevant. Indeed, the positive correlation that exists between the two measures of density (Figure 2; top right panel) implies higher employment density is often associated with higher population density. And as the size of the urban area increases, then we would expect this correlation to increase. While the consultants to the CGC focus on weighted population density at the SUA level, we expect this to be a reasonable proxy for employment density. This is one example of how we confirm and extend the findings of the consultants engaged by the CGC.

2.4 Robustness

We evaluated the robustness of our results in two broad ways:

- First, we tested alternative specifications, such as different vehicle capacities, alternative fixed effects, broader congestion measures, labour force variables, and average trip length; and
- Second, we estimated an instrumental variables (IV) version of our model to control for potential endogeneity of our explanatory variables.

In all cases, we find that our baseline results are largely unchanged.

2.4.1 Alternative Specifications

First, we tested the sensitivity of results to assumptions on seated vehicle capacities by estimating a model where all three PT modes have the same seated capacity in all five cities, specifically bus, 50; tram 125; and heavy Rail, 500. Regression results for this model are summarised in Table 6. We find our coefficient estimates are largely unchanged.

Table 6: Regression Results – Vehicle Capacity Sensitivity Test

Parameters	SA3 Fixed Effects			Vehicle Capacity Variation		
	V1: Robust	V2: Cluster	V3: WLS	V1: Robust	V2: Cluster	V3: WLS
log(emp. density)	0.55 (0.05)***	0.55 (0.06)***	0.66 (0.05)***	0.49 (0.04)***	0.49 (0.04)***	0.58 (0.03)***
log(delays)	0.42 (0.05)***	0.42 (0.08)***	0.44 (0.06)***	0.39 (0.04)***	0.39 (0.07)***	0.41 (0.05)***
R^2	0.53	0.53	0.63	0.55	0.55	0.66

Notes: *All models include SA3 f.e. and have n = 1,058 obs. Standard errors in brackets;* $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

We also tested replacing SA3 fixed effects with SA4 fixed effects. This did not significantly affect the coefficient estimates but did reduce the explanatory power of the model. For this reason, we prefer the baseline model above. Similarly, including variables for car delays at the SA3 or SA4 level were not found to be statistically significant and nor did they affect the other coefficients.

To capture aspects of the labour force, we formulated and tested the following three measures:

- Percentage of commutes by white collar workers;
- Percent of commutes to the city centre; and
- Percentage of commutes to employment centres.

These measures were intended to capture differences in workforce propensity to use PT, which in turn may explain PT supply. While all three variables entered the model with the expected positive sign, they were not statistically significant and hence are not considered to add to the baseline model.

Finally, we included two measures of travel demand: The average trip distance by motorised modes (car and PT) and the average in-vehicle distance by PT. While both variables entered the model with the expected positive sign, we found they were not statistically significant. This finding differs from the consultants engaged by the CGC, which reported average 'distance to work' as significant at the 5% level. The divergence in our findings is likely to reflect underlying empirical differences in our respective models, such as the *units of observations* (i.e. SUAs vis-à-vis SA2s), *model specification* (i.e. linear-log vis-à-vis log-log), *explanatory variables* (i.e. we use employment density and include car delays) or *assumptions on standard errors* (i.e. we use cluster-robust standard errors). We note that the decision to use cluster-robust standard errors may represent a more stringent standard for statistical significance than that used by the consultants to the CGC.

2.4.2 Instrumental Variables

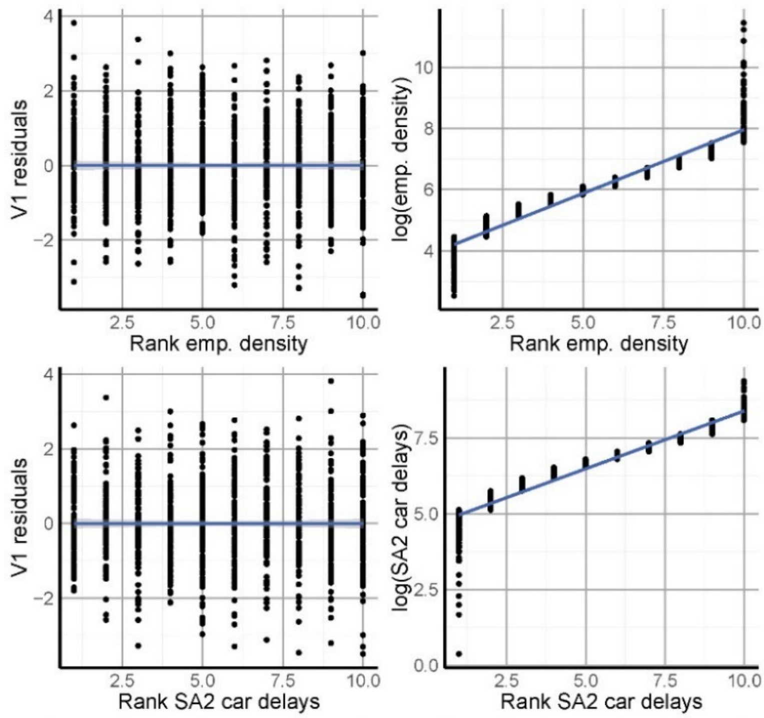
In the presence of endogenous regressors, our results may not capture causal effects. We addressed endogeneity using instrumented variables (IV). In technical terms, IV uses an *exogeneous variable* (the "instrument") to identify the exogeneous component of an endogenous regressor.

As well as being exogeneous, an instrument needs to satisfy the following two technical criteria: (1) *validity*, in the sense that it is independent of the model residuals; and (2) *relevance*, in the sense it is correlated with the endogenous regressor(s). For instruments, we follow the "rank instrument" method of Johnston et al. (1984)[14], as discussed in Kennedy (1992), which constructs instruments as follows:

- First, assign each observation to a quantile based on the endogenous regressors; and
- Second, the rank of the quantiles then defines the instrumental variable.

Several recent studies in the spatial economics literature use this process to generate instruments for endogenous variables.[15] We adapt the same process to generate instruments for our two potentially endogenous non-policy variables, namely employment density and car delays. The left and right panels of Figure 1 plot these instruments versus model residuals and the regressors, respectively.

Figure 4: Instrument validity (left panel) and relevance (right panel)



[14] Johnston, Jack, and John E. DiNardo. "Econometric Methods McGraw Hill." New York (1984)

[15] See, for example, *Fingleton, B. "Increasing returns: evidence from local wage rates in Great Britain." Oxford Economic Papers 55.4 (2003): 716-739.*

The zero correlation in the left panel provides informal evidence our instruments are *valid*, whereas the positive correlation in the right panel provides informal evidence our instruments are *relevant*. We also include an additional rank instrument for SA3 level congestion, which provides us with three instruments in total for our two variables (NB: The third instrument is useful for technical reasons, as it allows us to perform tests for over-identification and exogeneity, namely the Sargan-Hansen test).

Using these instruments, we re-estimate the SA3 Fixed Effects model, as presented in Table 7. We find coefficients that are similar in magnitude and significance across models and specifications. Interestingly, the coefficient for employment density in model V3 of the IV specification is much closer to those for models V1 and V2 (both in the baseline model and the IV version). This may suggest the weights used in model V3 increases the degree to which endogeneity biases our estimates.

Table 7: Regression Results – Instrumental Variables

Parameters	SA3 Fixed Effects			IV SA3 Fixed Effects		
	V1: Robust	**V2: Cluster**	**V3: WLS**	**V1: Robust**	**V2: Cluster**	**V3: WLS**
log(emp. density)	0.55 (0.05)***	0.55 (0.06)***	0.66 (0.05)***	0.56 (0.06)***	0.56 (0.09)***	0.57 (0.08)***
log(delays)	0.42 (0.05)***	0.42 (0.08)***	0.44 (0.06)***	0.41 (0.05)***	0.41 (0.06)***	0.43 (0.04)***
R^2	0.53	0.53	0.63	0.53	0.53	0.63

Notes: *All models include SA3 f.e. and have n = 1,058 obs. Standard errors in brackets; ***p < 0.001, **p < 0.01, *p < 0.05*

For model V2, formal tests of the instruments indicate we can:

- Reject the hypothesis our instruments are weak (i.e. our instruments are not irrelevant);
- Accept the hypothesis our variables are exogeneous (i.e. OLS is preferred to IV); and
- Accept the hypothesis our instruments are uncorrelated with model residuals (i.e. our instruments are valid, at least in a statistical sense).

For model V3, the Sargan-Hansen test rejects the null hypothesis, suggesting there is evidence of a relationship between our instruments and model residuals. While undesirable, this provides additional evidence to prefer the estimates from model V2, which passes the same test and yet has coefficients that are similar in magnitude to those found for model V3.

Finally, to test the stability of results for IV model V3 to our choice of weights, we also estimate a weighted OLS model where weights are defined by SA2 population (rather than the sum of population and employment). Promisingly, in this model the coefficients for both employment density and car delays are close to models V1, V2, and V3. As such, we conclude the results for model V3 are not especially sensitive to the choice of weighting scheme.

2.5 Summary of Results

Our preferred model is the SA3 Fixed Effects model V2 first reported in Table 5. This model implies elasticities of +0.5543 and +0.4173 for employment density and car delays, respectively. We can use these elasticities, along with the summary statistics in Table 3, to estimate the percentage effect of these two non-policy factors on average PT supply in each city using the following formula:

$$\frac{S_A}{S_S} = \left(\frac{d_A}{d_S}\right)^{\alpha_1} \left(\frac{c_A}{c_S}\right)^{\alpha_2}$$

Where: (1) S_A and S_S denote PT supply (cost) in actual (A) and sample average (S) scenarios; (2) d_A and d_S denote average employment density in each scenario; (3) c_A and c_S denote congestion levels in each scenario; and (4) α_1 and α_2 denote elasticities for employment density and car delays. Table 8 summarises the results of this calculation for each city.

Table 8: Summary of PT Supply Model – Average Effect for SA3 Fixed Effects (model V2)

City	d_A	c_A	**% change**
Sydney	1,610	1,783	+33%
Melbourne	1,413	1,612	+19%
SEQ	1,013	890	-23%
Perth	523	851	-48%
Adelaide	661	1,351	-28%
Average	*1,181*	*1,351*	*0%*

We work through these calculations in more detail in Section 5.1.

3. Productivity Models

Summary: In this section we model the effect of non-policy factors on the *productivity* of PT services in each of the five largest capital cities. We measure productivity in terms of average speed and route length. For Sydney, we find non-policy factors affect PT productivity in the following ways: (1) Buses and trams operate slower (i.e. reduce productivity); (2) Bus/tram routes are shorter (i.e. increase productivity); and (3) Heavy rail routes are longer (i.e. reduce productivity). Compared to the average for our data, we estimate non-policy factors reduce Sydney's PT productivity by 3.2%.

3.1 Average Speed Bus and Train Speed

3.1.1 Models

The Basic model of average speed, s, of trip i is specified as follows:

$$s_i^B = \alpha_0 + \alpha_1 \ln(l_i) + \alpha_2 \ln(d_i) + \alpha_3^c D_i^c + \alpha_4^w D_i^w + \alpha_5^h D_i^h + \alpha_6^{cw} D_i^c D_i^w + \alpha_7^{ch} D_i^c D_i^h + \alpha_8^{wh} D_i^w D_i^h$$

Where:

- s_i^B denotes the average speed of trip i
- l_i denotes the route distance between the start and end of trip i [km]
- d_i denotes the average stop-spacing on trip i [km per stop]
- D_i^c denotes a capital city categorical variable, where Sydney is defined as the base category
- D_i^w denotes a weekday/weekend dummy variable, where weekday is the base category
- D_i^h denotes twenty-four hourly categorical variables, where 0400-0500 is the base category.
- α_0, α_1, α_2, α_3^c, α_4^w, α_5^h, α_6^{cw}, α_7^{ch}, and α_8^{wh} denote regression parameters to be estimated.

The categorical variables D_i^c, D_i^w, and D_i^h allow average speeds to vary by city, between weekdays / weekends, and by hour of the day. The three pairwise interaction terms are interpreted as follows:

- $D_i^c D_i^w$ allows the average speed in each city to vary between weekdays/weekends;
- $D_i^c D_i^h$, allows the average speed in each city to vary by hour of day; and
- $D_i^w D_i^h$ allows the average speed on weekdays/weekends to vary by hour of day.

Our model includes 5 x 24 x 2 = 240 categorical variables. While this is a large number, we suggest it is small when compared to the size of our sample (n = 282,225), which we discuss in more detail below. The the inclusion of these categorical variables will reduce the scope for non-policy factors to affect average bus and tram speed, largely because congestion effects—which we discuss in more detail below—are partly wrapped-up into our hourly dummies. In this way, our choice to include hourly categorical variables makes it harder for us to separately identify an effect for congestion.

In this way, the Basic model estimates average speed as a function of exogeneous variables that are directly observable from GTFS data. We expect average speed will:

- Increase with route length, l_i, because longer routes tend to operate (1) in peripheral areas with less congestion or (2) where they have greater priority over general traffic; and
- Increase with stop-spacing, d_i, because longer stop spacing allows vehicles to achieve a higher speed and are also more likely to be associated with express/pre-pay services.

We do not have strong priors on the city dummy variables, D_i^c, except perhaps that smaller cities and/or those with more extensive priority infrastructure will tend to enjoy higher average speeds. Similarly, we do not have strong priors on the weekend dummy variable, D_i^w, because of the potential for them to capture countervailing effects: While weekends experience lower demand and less congestion, they also have fewer express services and less priority. We expect the twenty-four hourly categorical variables D_i^h will be higher in off-peak periods and lower in peak periods.

The Basic model is the benchmark, or baseline, to which we compare the effects of non-policy factors. Specifically, we extend the Basic model by including two additional non-policy variables: Population catchment, p_i, and traffic congestion, c_i.[16] We measure p_i as the number of residents within 750m of stops, where we exclude areas of overlap between stops. The average speed of services in the Extended model, s_i^E, is estimated as Basic model plus these additional variables:

$$s_i^E = s_i^B + \alpha_9 p_i + \alpha_{10} \ln(c_i)$$

We expect average speed will decline with catchment, p_i, as catchment is associated with increased PT demand. Services that experience greater demand will, on average, be expected to run more slowly. Similarly, we expect car delays, c_i, will slow down bus and tram services. We also expect the effect of congestion to diminish at higher levels, due to proactive policies to mitigate the effects of congestion, such as priority infrastructure. For this reason, we take the log of congestion (NB: We found empirical support for this choice of functional form).

3.1.2 Data

We use GTFS data for the five largest capital cities, excluding all trip-IDs that (1) fell outside the boundaries of the relevant GCCSA defined by ABS and (2) were operated by heavy rail or ferry.[17]

We pool bus and tram (or light rail) trip-IDs as we expect both modes will be similarly affected by non-policy factors. That said, we include a dummy variable in both the Basic and Extended models to allow for potential differences in operating performance between bus and tram.[18]

We filtered out erroneous and/or unrepresentative values, specifically trip-IDs with:

- Average speeds in excess of 70 km/hr. Inspection suggests these trips often suffered from errors in the route geometries coded by the agencies/operators who prepare the data.
- Trip-lengths that are shorter than 5km or longer than 50km, or that have durations that are less than 20 minutes or more than 2-hours. Inspection of the data revealed that such trips are associated with relatively atypical routes, such as inner-city circulators, or shuttles.
- Two or fewer stops. Many of these trips appear to be associated with short-running and point-to-point "shuttle" services, such as those that may be provided for major sports events.
- Less than 100m (Euclidean distance) between start and end stops, which are associated with "loop" services that are slow / circuitous by design rather than due to non-policy factors.

These filters reduced the number of trip-IDs from n = 335,168 to n = 282,225, or 84% of the original data. We consider this to be a large and representative sample of PT services in the five capital cities.

Summary statistics for each capital city are presented in Table 9. We find Sydney operates the second highest number of bus and tram trips (services) after SEQ, which likely reflects the latter's extensive busway infrastructure (and associated high-frequency services) and limited heavy rail network. After SEQ, the number of trips declines with city size. For average speed, we see that Sydney is the slowest of all cities, while Perth is the fastest, approximately 1 km per hour faster than SEQ, which in turn is almost 1 km per hour faster than Melbourne. Although Sydney and Melbourne have similar average speeds, the former operates longer routes with larger distances between stops.

---

[16] In a similar vein to Section 2.1, we extract total vehicle delay hours from VLC's Zenith models for each capital city at the SA1 level and then assign it to trip-IDs based on the SA1s that they traverse. Population data was sourced from the ABS for 2016 at the meshblock level. Population catchment for each trip, p_i, is then the sum of the meshblock populations whose centroids lie within 750m of stops on the trip.

[17] GTFS is a standardised description of PT services. We use scheduled GTFS data, rather than real-time data, for two reasons. First, real-time data is not publicly available for all the cities we analyse. Second, scheduled data provides information on the *systematic effects* of non-policy factors, rather than *idiosyncratic variation* due to non-recurrent factors, such as weather, accidents, and events.

[18] Table 9 shows that the tram dummy is statistically significant in the Basic model but insignificant in the Extended model (V3). This implies there is no observable difference in speeds between the two modes.

Table 9: Summary Statistics for Productivity Model – Bus and Tram Average Speed by SA2

Market	Mode	n [trips]	Average speed (s_i) [km.hr]	Route length (l_i) [km]	Stop-spacing (d_i) [km/stop]	Pop. Catchment (p_i) [people]	Congestion (c_i) [veh.hr]
Syd.	Bus	186,620	22.388	17.653	0.528	75,435	1,938
	Tram	1,967	20.533	12.395	0.540	111,160	1,417
	Total	*188,587*	*22.369*	*17.598*	*0.528*	*75,808*	*1,933*
Mel.	Bus	149,966	23.458	16.690	0.417	51,336	1,099
	Tram	30,372	16.381	15.061	0.290	111,881	1,643
	Total	*180,338*	*22.266*	*16.416*	*0.396*	*61,533*	*1,190*
SEQ	Bus	318,493	24.331	17.774	0.735	46,710	1,515
	Tram	3,626	25.685	18.979	0.855	50,956	582
	Total	*322,119*	*24.346*	*17.788*	*0.736*	*46,758*	*1,504*
Perth	Bus	95,598	25.322	15.921	0.443	34,348	741
	Tram	N/A	N/A	N/A	N/A	N/A	N/A
	Total	*95,598*	*25.322*	*15.921*	*0.443*	*34,348*	*741*
Adl.	Bus	45,909	23.459	18.802	0.456	38,482	1,263
	Tram	3,152	17.513	14.740	0.534	43,619	1,207
	Total	*49,061*	*23.077*	*18.541*	*0.461*	*38,812*	*1,259*
Sample		*835,703*	*23.488*	*17.280*	*0.566*	*54,616*	*1,431*

In terms of non-policy factors, the average bus / tram trip in Sydney has higher catchment and congestion levels than is found in the other cities. For example, average population catchment is 23% higher in Sydney than in Melbourne, while congestion is 29% higher in Sydney than in SEQ.

3.1.3 Results

Regression results for the Basic and Extended productivity models are presented in Table 10 below.

Table 10: Regression Results for Productivity Model – Bus and Tram Average Speed

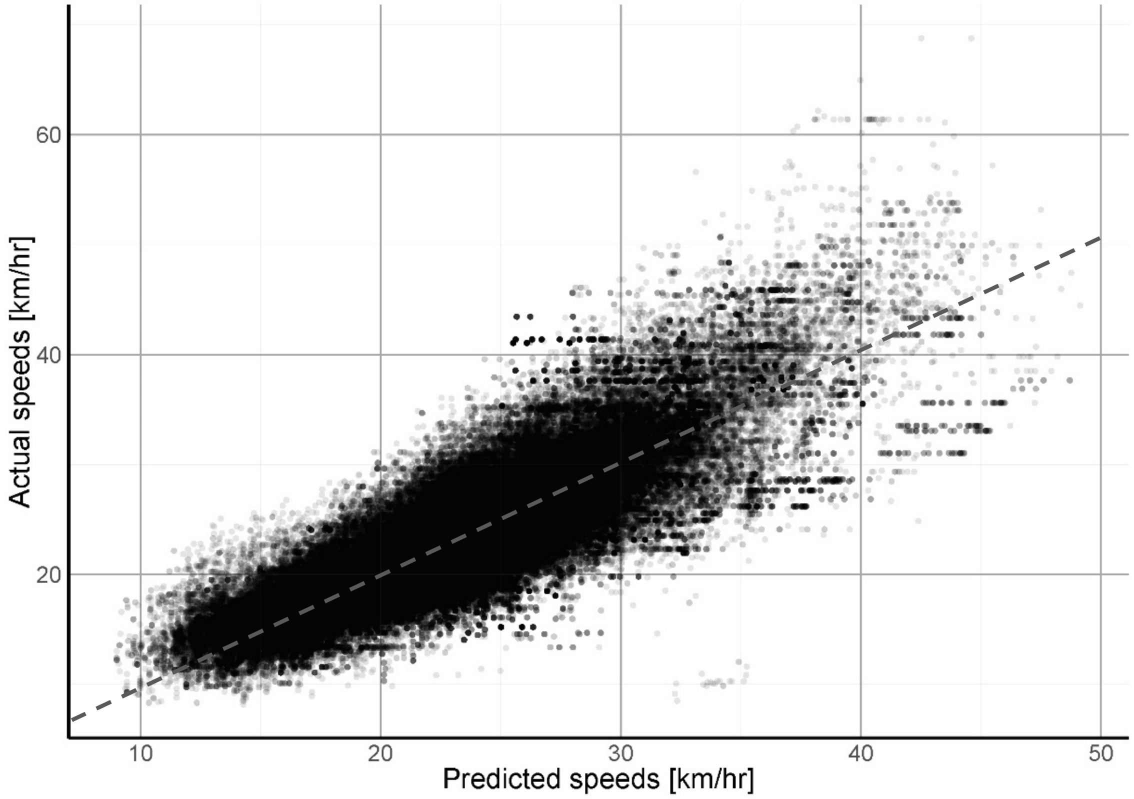
Parameters	Basic			Extended		
	V1: Robust	V2: Cluster	V3: Weighted	V1: Robust	V2: Cluster	V3: Weighted
Tram Dummy	-5.43 (0.02)***	-5.43 (0.53)***	-5.03 (0.58)***	-0.53 (0.03)***	-0.53 (0.77)	0.00 (0.80)
ln(length) (l_i) [km]	4.69 (0.02)***	4.69 (0.77)**	4.29 (0.63)**	8.75 (0.02)***	8.75 (0.43)***	8.39 (0.30)***
ln(stop-spacing) (d_i) [km/stop]	5.86 (0.02)***	5.86 (0.86)**	6.31 (0.63)***	3.92 (0.02)***	3.92 (0.78)**	4.44 (0.54)**
Population Catchment (p_i) [per 1,000 people]	-	-	-	-0.09 (0.00)***	-0.09 (0.01)**	-0.10 (0.02)**
ln(congestion) (c_i) [veh.hrs]	-	-	-	-0.98 (0.01)***	-0.98 (0.31)*	-1.11 (0.32)*
R^2	0.55	0.55	0.53	0.71	0.71	0.70
Notes to table	*n = 282,225. Standard errors in brackets; ***p < 0.001, **p < 0.01, *p < 0.05. All models include City x Weekend x Hour terms. Fixed effects are omitted but are available on request; inspection revealed a logical profile for hourly dummies.*					

In this case, our preferred model is model V3, which estimates clusters-robust standard errors using WLS. Our choice of weights is the number of times per week that an individual trip-ID operates. The latter provides useful additional information on the relative importance of individual observations.

We are primarily interested in coefficients for non-policy factors, which are reported in the bottom two rows of Table 10. These results find both population catchment and congestion have the expected negative sign and are statistically significant ($p < 0.05$ or smaller) in all the three Extended models that we test. These results provide evidence that higher population catchments and increased congestion leads to lower bus/tram speeds, which aligns with our priors.

To finish, we consider the overall explanatory power of the model. Figure 5 also illustrates predicted average speeds (horizontal axis) versus actual average speeds (vertical axis). Generally, we find a strong positive linear observation with clustering around the diagonal and few apparent extreme values. The Extended models have R-squared values of approximately 0.70.

Figure 5: Productivity Model – Bus and Tram Actual Speeds Extended Model Fit (model V3)



3.1.4 Robustness

We considered several alternative specifications of the Extended V3 Weighted model, specifically:

- **S1 city-specific catchments,** where we allow for the effect of population catchment on average speeds to vary by city, reflecting differences in underlying demand.
- **S2 interaction term**, where we include an interaction term between route length and stop spacing, reflecting our expectation that the effects of route length interact with stop-spacing.
- **S3 500m catchments**, where we use a 500m population catchment around stops as opposed to the 750m used in previous specifications.
- **S4 road infrastructure**, where we include additional variables to capture the presence of bus lanes and road type for each trip-ID as a percentage of total route distance.

Regression results for these sensitivity tests are summarised below in Table 11, where we report only those coefficients that relate to non-policy factors and our alternative model specifications.[19] Considering S1 first, we find several of the city-specific catchment variables are statistically significant at the 0.1% or 1% levels. The negative effect of catchment on average speeds in SEQ and Perth is estimated to be approximately twice as large as in Sydney. Including city-specific catchment effects improves the overall explanatory power of the model, while also reducing the size of the fixed effects

[19] In all four alternative model specifications, the dummy variable for tram services remains statistically insignificant. This supports the decision to pool data for bus and tram services.

estimated for each city. We view the latter as a positive development, because the city fixed effects capture residual unexplained differences in speeds between cities.

Turning now to model S2, we find that the interaction term between length and stop-spacing is statistically significant at the 5% level. The inclusion of the interaction term, however, has relatively negligible effects on the other model coefficients. Model S3 is the same as S2, except that we now use a 500m population catchment rather than the 750m that is used in the other models. While the coefficients for population catchment are scaled up to account for the reduced size of the catchment, our main findings with regards to catchment and congestion are largely unchanged.

Model S4 includes information on road infrastructure, specifically the percentage of the route that is classed as busway, motorway, or main roads (NB: The final category of road type, namely residential streets, is omitted due to collinearity with the other variables).[20] While all three road types are found to be statistically significant, the estimated coefficients for our non-policy factors is unaffected.

Table 11: Productivity Model Sensitivity Tests – Bus and Tram Average Speed

Parameters	V3: Weighted	S1: City-specific	S2: Interaction	S3: 500m catch.	S4: Road infra.
Population Catchment (p_i) [per 1,000 people]	-0.10 (0.02)**	-0.09 (0.00)***	-0.08 (0.00)***	-0.13 (0.01)***	-0.12 (0.01)***
ln(congestion) (c_i) [veh.hrs]	-1.11 (0.32)*	-0.80 (0.40)	-0.82 (0.39)	-0.99 (0.30)*	-1.44 (0.21)**
Melbourne catchment	-	-0.00 (0.00)	-0.00 (0.00)	-0.01 (0.01)	-0.02 (0.01)
SEQ catchment	-	-0.11 (0.01)***	-0.12 (0.00)***	-0.20 (0.01)***	-0.20 (0.02)***
Perth catchment	-	-0.11 (0.02)**	-0.11 (0.01)***	-0.19 (0.02)**	-0.19 (0.03)**
Adelaide catchment	-	-0.09 (0.01)**	-0.08 (0.01)**	-0.14 (0.02)**	-0.13 (0.02)**
ln(length)*ln(stop-spacing)	-	-	1.37 (0.44)*	1.61 (0.35)**	4.84 (0.52)***
Busway [%]					5.73 (0.99)**
Motorway [%]					0.68 (0.15)**
Main road [%]					1.51 (0.30)**
R^2	0.70	0.73	0.73	0.73	0.75
Notes to table	*n = 282,225. Standard errors in brackets; ***p < 0.001, **p < 0.01, *p < 0.05. All models include City x Weekend x Hour terms*				

Notwithstanding the apparent improvement in explanatory power, the estimated coefficients for catchment and congestion are not significantly different from model V3 in any of these sensitivity tests. By using model V3 in later sections, we are effectively calculating a smaller productivity effect than that predicted by model S4. That is, we are erring on the side of underestimating effects,

3.2 Route Length

In this section we develop productivity models of route length. We first present Basic and Extended models used for both modes, after which are mode-specific sub-sections.

3.2.1 Models

We follow a similar modelling process to that used previously in Section 3.1. That is, we first specify a Basic model that contains a set of general exogeneous controls, after which we specify an Extended model that includes our non-policy factors of interest. For route distance, the Basic model is given by:

---

[20] The coefficient denotes the speed associated with shifting from 0% to 100% for each road type. For example, if a route is moved from 0% to 100% busway, we predict a 5.73 km/hr (24%) increase in average speed.

$$l_i^B = \alpha_0 + \alpha_1 \bar{l}_i + \alpha_2 \bar{d}_i + \alpha_3^c D_i^c + \alpha_4^w D_i^w + \alpha_5^h D_i^h + \alpha_6^{cw} D_i^c D_i^w + \alpha_7^{ch} D_i^c D_i^h + \alpha_8^{wh} D_i^w D_i^h$$

Where l_i^B denotes the route distance between the start and end of trip i [km]; $\bar{l}_i$ denotes the Euclidean (or "crow flies") distance between the start and end of the route; and $\bar{d}_i$ denotes the number of stops. As per Section 3.1.1, D_i^c, D_i^w, and D_i^h denote categorical variables for city, weekday / weekends, and time of day, respectively. We include all pairwise interaction effects between categorical variables. Coefficients α_0, α_1, α_2, α_3^c, α_4^w, α_5^h, α_6^{cw}, α_7^{ch}, and α_8^{wh} denote regression parameters to be estimated.

We extend the Basic model with four non-policy variables. In addition to catchment *(p_i)* and the log of congestion (c_i), which were previously discussed in Section 3.1.1, we include *geographical deviation*, g_i, and *vertical elevation,* z_i (NB: These non-policy factors are defined in detail in Section 1.4). The Extended model, l_i^E, is thus equal to the Basic model plus these additional variables:

$$l_i^E = l_i^B + \alpha_9 p_i + \alpha_{10} \log c_i + \alpha_{11} g_i + \alpha_{12} z_i$$

In terms of priors, for the Basic model we expect l_i^B will increase with Euclidean distance, $\bar{l}_i$, and the number of stops, $\bar{d}_i$. In the Extended model, we expect l_i^E declines with population catchment, p_i, for two reasons: First, in denser areas routes do not need to travel as far to reach people, second, in denser areas PT planners design shorter routes to maintain reliability. We do not have strong priors on congestion, c_i, although we expect l_i^E increases with barriers, g_i; and elevation, z_i.

3.2.2 Bus and Tram

3.2.2.1 Data

For the route length model, we use the same data source (including filters) as previously described in Section 3.1.2. Summary statistics are presented below in Table 12.

Table 12: Summary Statistics for Productivity Model – Bus and Tram Route Length

Market	Mode	n [trips]	Length (l_i) [km]	Eu. Dist ($\bar{l}_i$) [km]	Stops ($\bar{d}_i$)	Catchment (p_i) [people]	Congestion (c_i) [veh.hr]	Deviation (g_i) [km]	Elevation (z_i) [km]
Syd.	Bus	186,620	17.65	10.32	43.93	75.00	1,938	2.44	0.29
	Tram	1,967	12.40	6.89	22.96	111.00	1,417	1.44	0.12
	Total	188,587	17.60	10.29	43.72	75.38	1,933	2.43	0.29
Mel.	Bus	149,966	16.69	10.05	43.32	51.00	1,099	1.707	0.22
	Tram	30,372	15.06	12.03	51.65	112.00	1,643	1.609	0.27
	Total	180,338	16.42	10.07	43.41	51.64	1,105	1.71	0.22
SEQ	Bus	318,493	17.77	10.83	29.64	47.00	1,515	2.417	0.24
	Tram	3,626	18.98	14.79	22.22	51.00	582	4.206	0.13
	Total	322,119	17.79	10.87	29.56	47.04	1,505	2.44	0.23
Perth	Bus	95,598	15.92	9.94	38.78	34.00	741	2.125	0.21
Adl.	Bus	45,909	18.80	12.13	45.74	38.00	1,263	2.018	0.26
	Tram	3,152	14.74	9.82	27.58	44.00	1,207	2.385	0.14
	Total	49,061	18.54	12.10	45.55	38.06	1,262	2.02	0.25
Sample		835,703	17.28	10.59	37.96	54.51	1,432	2.22	0.24

We find average route length, l_i; Euclidean distance, $\bar{l}_i$, and number of stops, $\bar{d}_i$ are similar in Sydney to other cities. Comparing the number of stops across cities, we find that SEQ is the outlier with fewer stops per trip, on average. Again, this likely reflects the effects of SEQ's extensive busways.

In terms of our non-policy factors, summary statistics for catchment and congestion are identical to Table 9 and are not discussed further. For our new non-policy factors, namely geographical deviation, g_i, and vertical elevation, z_i, we find Sydney has the second-largest (after SEQ) and largest values, respectively. The implication is that bus and tram routes in Sydney tend to face greater geographical barriers and larger changes in vertical elevation than the average route.

3.2.2.2 Results

Regression results for the Basic and Extended models are present in Table 13. In the Basic model, we find the expected positive relationship between average route length and our control variables for Euclidean distance and number of stops.

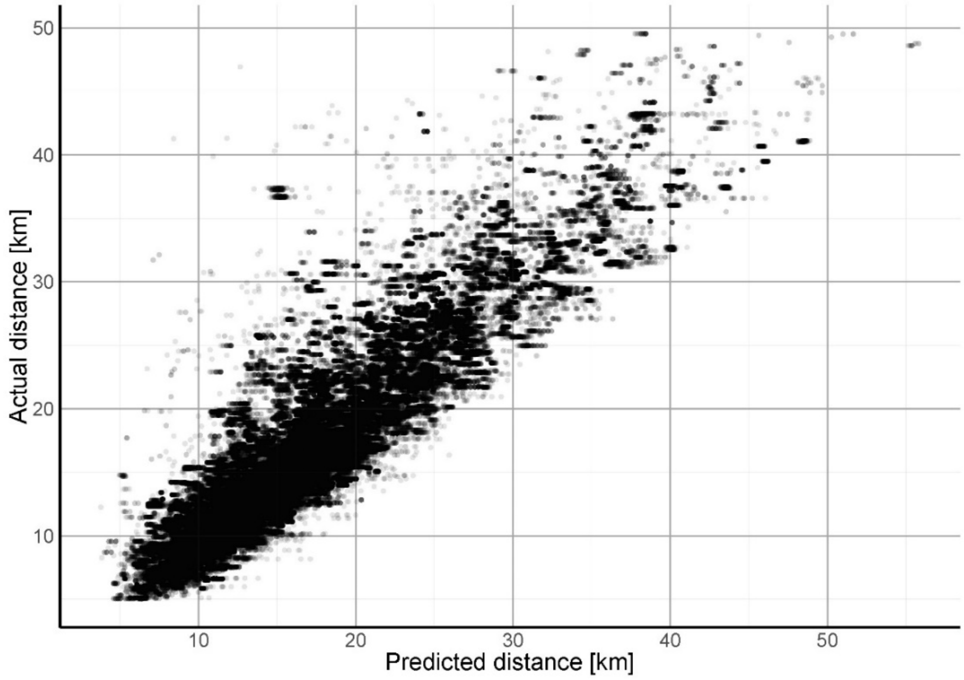
When we introduce our four non-policy factors, we find both geographical barriers and vertical elevation have the expected positive effect on route length ($p < 1\%$), whereas population catchment has the expected negative effect ($p < 1\%$). Congestion is significant ($p < 0.1\%$) in model V1 but is insignificant in models V2 and V3.

Table 13: Regression Results for Productivity Model – Bus and Tram Route Length

Parameter	Basic			Extended		
	V1: Robust	V2: Cluster	V3: Weighted	V1: Robust	V2: Cluster	V3: Weighted
Tram Dummy	-3.45 (0.02)***	-3.45 (1.09)*	-3.53 (1.03)*	-1.62 (0.03)***	-1.62 (0.73)	-1.69 (0.67)
Euclidean distance	0.99 (0.00)***	0.99 (0.08)***	1.00 (0.07)***	0.92 (0.00)***	0.92 (0.08)***	0.94 (0.07)***
Number of stops	0.13 (0.00)***	0.13 (0.01)***	0.13 (0.00)***	0.14 (0.00)***	0.14 (0.02)***	0.14 (0.01)***
Population Catchment (p_i) [per 1,000 people]	-	-	-	-0.04 (0.00)***	-0.04 (0.01)**	-0.04 (0.01)**
ln(congestion) (c_i) [veh.hrs]	-	-	-	0.41 (0.01)***	0.41 (0.18)	0.32 (0.16)
Deviation (g_i) [km per route]	-	-	-	0.39 (0.01)***	0.39 (0.08)**	0.34 (0.06)**
Elevation (z_i) [Δ height per route]	-	-	-	4.98 (0.08)***	4.98 (1.15)*	5.59 (1.21)**
R^2	0.81	0.81	0.81	0.83	0.83	0.83
Notes to table	*n = 282,225. Standard errors in brackets; ***p < 0.001, **p < 0.01, *p < 0.05. All models include City x Weekend x Hour terms*					

Both the Basic and Extended models have good explanatory power, with R-squared values of 0.81 and 0.83, respectively. Predicted versus actual values are illustrated in Figure 6. This shows a strong positive association with most values clustered around the 45-degree line. We observe more variation on the upside, which suggests that errors are heteroskedastic, supporting the use of robust s.e.

Figure 6: Productivity Model – Bus and Tram Route Length Extended Model Fit



3.2.2.3 Robustness

We considered the following alternative model specifications: (1) *S1 city-specific catchments***,** where the effect of population catchment varies by city, (2) *S2 interaction term*, between route length and number of stops, (3) *S3 500m catchments*, where we use a 500m population catchment as opposed to 750m; and (4) *S4 road infrastructure*, where we include variables for length of busways and road type. Regression results for these alternative model specifications are summarised in Table 14.

Table 14: Productivity Model Sensitivity Tests – Bus and Tram Route Length

Parameters	Extended V3: Weighted	S1: City-specific	S2: Interaction	S3: 500m catch.	S4: Road infra.
Population Catchment (p_i) [per 1,000 people]	-0.04 (0.01)**	-0.03 (0.00)***	-0.03 (0.00)***	-0.05 (0.00)***	-0.04 (0.00)***
ln(congestion) (t_i) [veh.hrs]	0.32 (0.16)	0.37 (0.15)	0.22 (0.12)	0.18 (0.13)	0.34 (0.11)*
Deviation (g_i) [km per route]	0.34 (0.06)**	0.35 (0.06)**	0.38 (0.05)**	0.38 (0.05)**	0.38 (0.02)***
Elevation (z_i) [Δ height per route]	5.59 (1.21)**	5.08 (1.14)*	4.92 (1.46)*	4.89 (1.33)*	4.97 (1.21)*
Melbourne catchment	-	-0.03 (0.01)*	-0.02 (0.00)*	-0.02 (0.01)	-0.03 (0.01)*
SEQ catchment	-	-0.00 (0.01)	-0.01 (0.00)**	-0.02 (0.00)*	-0.03 (0.01)
Perth catchment	-	-0.05 (0.01)**	-0.05 (0.00)***	-0.07 (0.01)***	-0.06 (0.01)**
Adelaide catchment	-	-0.04 (0.01)*	-0.02 (0.01)*	-0.02 (0.01)	-0.02 (0.01)
Eu_dist*stops	-	-	-0.01 (0.00)***	-0.01 (0.00)***	-0.01 (0.00)***
Busway [%]	-	-	-	-	-2.55 (1.73)
Motorway [%]	-	-	-	-	1.871 (1.96)
Main Road [%]	-	-	-	-	-3.14 (0.48)**
R^2	0.83	0.84	0.85	0.85	0.86
Notes to table	*n = 282,225. Standard errors in brackets; ***p < 0.001, **p < 0.01, *p < 0.05. All models include City x Weekend x Hour terms*				

Including additional explanatory variables marginally improves the model's explanatory power, which increases to 0.86 in model S4. Nonetheless, our main results are robust to all four sensitivity tests. That is, our non-policy factors, namely population catchment, geographical deviation, and vertical elevation, all retain their sign, approximate magnitude, and statistical significance. We prefer Extended model V3 because it is simple and generates comparable coefficients to models S1 – S4.

3.2.3 Heavy Rail

3.2.3.1 Data

We use the same data as described in Section 3.1.2, except we now focus on heavy rail services rather than bus and tram services. Summary statistics are presented below in Table 15.

Table 15: Summary Statistics for Productivity Model – Heavy Rail Route Length

Market	n [trips]	Length (l_i) [km]	Eu. Dist ($\bar{l}_i$) [km]	Stops ($\bar{d}_i$)	Catchment (p_i) [people]	Congestion (t_i) [veh.hr]	Deviation (g_i) [km]	Elevation (z_i) [km]
Syd.	47,096	39.48	29.18	19.38	138,000	2,819	4.59	0.25
Mel.	57,240	31.12	25.55	18.75	93,000	2,312	2.60	0.19
SEQ	4,748	58.07	43.70	26.91	82,000	2,408	7.80	0.20
Perth	15,627	22.92	19.83	14.53	35,000	1,090	2.45	0.08
Adl.	3,572	27.53	22.75	17.61	35,000	1,061	2.63	0.17
Sample	*128,283*	*34.09*	*26.78*	*18.73*	*100,433*	*2,318*	*3.50*	*0.20*

Sydney operates the most kilometres (trips x length) by heavy rail. In terms of average route length, we find that Sydney is slightly longer than the sample average and has slightly more stops. As for non-policy factors, Sydney's population catchment, congestion levels, and elevation are higher than in other cities, whereas for deviation Sydney has the second largest after SEQ. While Sydney and SEQ appear to be outliers by this measure, most of the difference is explained by the length of rail trips in these cities, which are longer than average. Indeed, if we divide the deviation by the average route length to calculate deviation per kilometre travelled, then we find that the ratios for all five cities are much closer together. We return to this issue in Section 3.2.3.3.

3.2.3.2 Results

Regression results for our Basic and Extended heavy rail route models are presented in Table 16.

Table 16: Regression Results for Productivity Model – Heavy Rail Route Length

Parameter	Basic			Extended		
	V1: Robust	V2: Cluster	V3: Weighted	V1: Robust	V2: Cluster	V3: Weighted
Euclidean distance	1.04 (0.00)***	1.04 (0.03)***	1.03 (0.03)***	0.89 (0.00)***	0.89 (0.03)***	0.90 (0.02)***
Number of stops	0.37 (0.00)***	0.37 (0.09)*	0.35 (0.10)*	0.17 (0.00)***	0.17 (0.09)	0.17 (0.11)
Population Catchment (p_i) [per 1,000 people]	-	-	-	0.02 (0.00)***	0.02 (0.01)	0.01 (0.01)
ln(congestion) (t_i) [veh.hrs]	-	-	-	0.71 (0.02)***	0.71 (0.81)	0.41 (0.75)
Deviation (g_i) [km per route]	-	-	-	1.03 (0.01)***	1.03 (0.18)**	1.00 (0.13)**
Elevation (z_i) [Δ height per route]	-	-	-	12.15 (0.13)***	12.15 (2.00)**	12.80 (1.53)**
R^2	0.95	0.95	0.94	0.97	0.97	0.97
Notes to table	*n = 70,745. Standard errors in brackets; ***p < 0.001, **p < 0.01, *p < 0.05.*					

Considering the Basic model first, we find that route length tends to increase with Euclidean distance and the number of stops, as expected.

In the Extended model, we add our four non-policy factors. For our preferred models V2 and V3, coefficients for deviation and elevation are positive and statistically significant ($p < 0.1\%$), whereas those for catchment and congestion are not. This aligns with our prior expectations.[21]

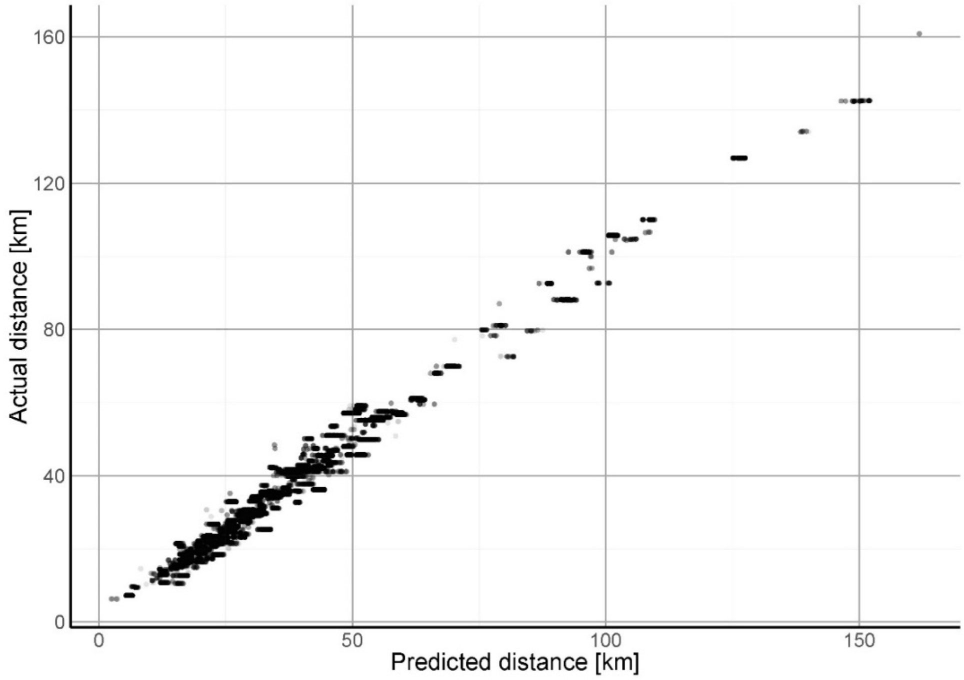
We find that deviation and elevation seem to have a greater effect on the route length for heavy rail services compared to buses and trams. Indeed, if we compare the coefficients on these two variables in Table 16 to their counterparts in Table 13, we find that the effect of deviation and elevation to be 2-3 times greater for heavy rail than for bus/tram. This is an interesting finding, which suggests that heavy rail is more sensitive to underlying geography and topography.

Some may question the latter finding on the grounds that rail infrastructure tends to overcome geography, for example using tunnels. While this is may be true for modern railways, we suggest it is less applicable to older rail infrastructure, which was constructed at a time when geographic barriers were less easily overcome. In SEQ, this is seen by comparing the relatively indirect (old) rail network with the relatively direct (new) busways, for example. For this reason, we expect older railways, such as those that exist in most Australian capital cities, are more sensitive to geography than buses.

In terms of explanatory power, we find both the Basic and Extended heavy rail route length V3 models have high R-squared values of 0.94 and 0.97, respectively. Model fit is illustrated in Figure 7. This reveals an excellent alignment between the predicted and actual values, with no extreme values.

[21] Levels of road congestion are unlikely to influence rail route length because the length of rail routes is determined by the existence of infrastructure more so than reliability.

Figure 7: Productivity Model – Heavy Rail Route Length Extended Model Fit



3.2.3.3 Robustness

We test the following alternative models:

- **S1 Constrained,** where we remove the (insignificant) catchment and congestion variables.
- **S2 Control interaction,** in which interact Euclidean distance and the number of stops.
- **S3 Data subset,** in which we re-estimate the model without SEQ and Sydney.

To motivate model S3, we refer to the summary statistics presented in Table 15. Here, we see heavy rail routes in SEQ and, to a lesser extent, Sydney have high values for deviation and elevation. To understand whether these cities drive our results, we remove them from the data and re-estimate model S3 on the remaining sub-set of data. Results for all three models are presented in Table 17.

Table 17: Productivity Model Sensitivity Tests – Heavy Rail Route Length

Parameters	V3: Weighted	S1: Constrained	S2: Interaction	S3: Data subset
Population Catchment (p_i) [per 1,000 people]	0.01 (0.01)	-	-	-
ln(congestion) (t_i) [veh.hrs]	0.41 (0.75)	-	-	-
Deviation (g_i) [km per route]	1.00 (0.13)**	0.96 (0.11)**	1.00 (0.14)**	0.94 (0.15)*
Elevation (z_i) [Δ height per route]	12.80 (1.53)**	13.50 (1.26)***	12.70 (1.84)**	14.47 (3.55)
Eu_dist*stops	-	-	-0.01 (0.00)**	-0.00 (0.00)
R^2	0.97	0.97	0.97	0.98
Notes to table	*n = 70,745, except S3 where n = 36,951. Standard errors in brackets; ***p < 0.001, **p < 0.01, *p < 0.05.*			

The estimated coefficients for the two statistically significant non-policy factors, namely deviation and elevation, are largely unchanged in each of these three alternative models. The largest change in the coefficients is associated with model S3, where the effect of deviation and elevation both vary marginally. That said, these changes are not statistically significant.

4. Revenue Model

Summary: We model the effect of non-policy factors on PT *revenue* in Sydney. Within the Sydney GCCSA, we use Opal ticketing data to estimate PT revenue for individual SA2s. We find employment density and car delays have a positive effect on revenue. Compared to average, we estimate non-policy factors increase Sydney's net PT revenue per capita by 37%.

4.1 Model

Our dependent variable is total fare revenue for individual SA2s, R_i. We model R_i using a similar log-log model to that used in Section 2.1:

$$\log R_i = f_i^R + \beta_1 \log d_i + \beta_2 \log c_i$$

Where:

- R_i denotes fare revenue by SA2 i;
- d_i denotes employment density in SA2 i;
- c_i denotes total daily vehicle delay hours that are incurred in SA2 i; and
- β_1 and β_2 denote parameters to be estimated.

Our prior expectations are that revenue increases with employment density and car delays.

4.2 Data

Summary statistics for our revenue data are summarised in Table 18 below.

Table 18: Summary Statistics for Revenue Model – Averages by Urban SA2

City	n	PT revenue (R_i)	Population (p_i)	Employment (e_i)	Congestion (c_i)	Area A_i
SYD	280	10,530,842	16,693	7,455	1,783	11.15
Sample	*1,058*	*5,327,325*	*14,042*	*6,161*	*1,351*	*13.30*

While we do not have revenue data for other capital cities, we can apply the elasticities from our model to estimate the effect of non-policy factors on revenue in other cities. This assumes PT revenue in other cities responds to non-policy factors in a manner that is similar Sydney.

4.3 Results

Regression results for the revenue model are summarised in Table 19. For models V2 and V3 we cluster standard errors by SA4.

Table 19: Regression Results – Revenue Model

Parameters	SA3 Fixed Effects		
	V1: Robust	**V2: Cluster**	**V3: WLS**
log(emp. density)	0.64 (0.10)***	0.64 (0.10)***	0.79 (0.12)***
log(delays)	0.43 (0.08)***	0.43 (0.06)***	0.36 (0.12)**
R^2	0.59	0.59	0.72

Notes: All models include SA3 f.e. and have n = 280 obs.
Standard errors in brackets; $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Specifically, our results suggest the following elasticities of PT revenue with respect to:

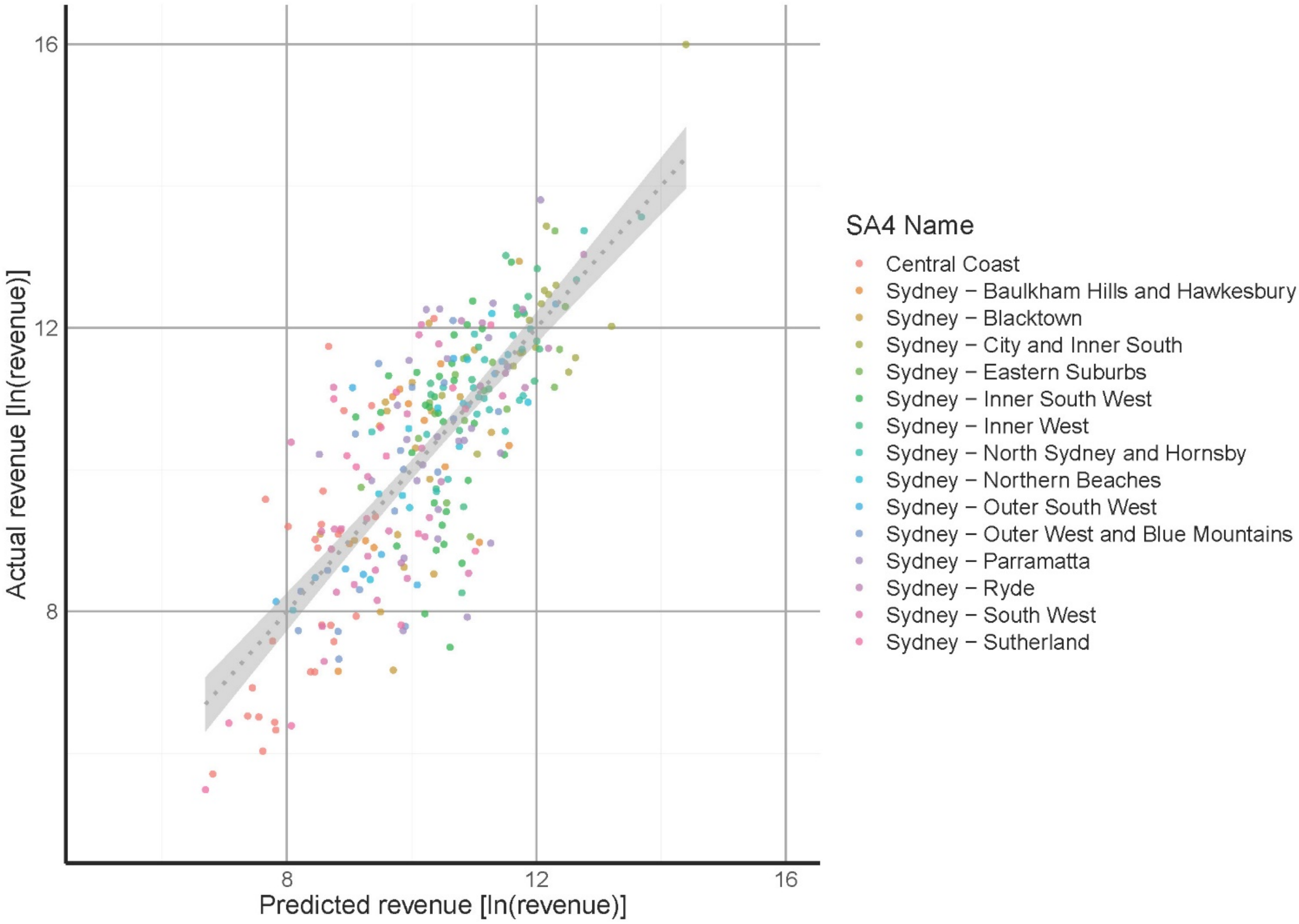
- Employment 0.64 – 0.79,

- Vehicle delays of 0.43 – 0.36.

In model V3, the estimated elasticities for car delays is significant at the 1% level, whereas all other coefficients in all other models are significant at the 0.1% level.

In terms of model fit, we find R-squared values of 0.59 and 0.72 for models V1 / V2 and V3, respectively. For model V2, actual and predicted PT revenues are illustrated in Figure 8.

Figure 8: Revenue Model – Model Fit (V2)



Promisingly, Figure 8 illustrates a strong positive association with points clustered around the diagonal and no extreme values.

4.4 Robustness

In the following sections we test the sensitivity of our revenue model to alternative specifications and estimate an instrumental variables version.

4.4.1 Alternative Specifications

First, we included population density as an additional argument. Regression results indicated this explanatory variable was statistically insignificant.

Second, we estimated a model using SA4 Fixed Effects rather than SA3 Fixed Effects. Regression results for this model are summarised in Table 20. Differences between the coefficient estimates for either model specification are not statistically significant. Estimates for the SA4 Fixed Effects model are marginally more precise, although the model has lower overall explanatory power. Given the relative similarities between the sets of results, we choose to prefer the SA3 Fixed Effects model.

Table 20: Regression Results – SA4 Fixed Effects

Parameters	SA3 Fixed Effects			SA4 Fixed Effects		
	V1: Robust	**V2: Cluster**	**V3: WLS**	**V1: Robust**	**V2: Cluster**	**V3: WLS**
log(emp. density)	0.64 (0.10)***	0.64 (0.10)***	0.79 (0.12)***	0.53 (0.08)***	0.53 (0.08)***	0.71 (0.10)***
log(delays)	0.43 (0.08)***	0.43 (0.06)***	0.36 (0.12)**	0.43 (0.07)***	0.43 (0.07)***	0.35 (0.11)**
R^2	0.59	0.59	0.72	0.53	0.53	0.67

Notes: *All models include SA3 or SA4 f.e. and have n = 280 obs. Standard errors in brackets; ***p < 0.001, **p < 0.01, *p < 0.05*

To capture salient aspects of the labour force, we then tested three measures related to the percentage of commutes:

- By white collar workers;
- To the city centre; and
- To employment centres.

As discussed previously, these measures were intended to capture workforce propensity to use PT, which in turn may explain PT revenue. When all three variables enter the revenue model, only the first two listed above were found to be statistically significant and have the expected positive sign. For this reason, we dropped the percentage of commutes and re-estimated the model with only the first two variables. Regression results for this restricted model are summarised in Table 21.

Table 21: Regression Results – Labour Force

Parameters	SA3 Fixed Effects			Labour Force		
	V1: Robust	**V2: Cluster**	**V3: WLS**	**V1: Robust**	**V2: Cluster**	**V3: WLS**
log(emp. density)	0.64 (0.10)***	0.64 (0.10)***	0.79 (0.12)***	0.46 (0.10)***	0.46 (0.12)***	0.50 (0.12)***
log(delays)	0.43 (0.08)***	0.43 (0.06)***	0.36 (0.12)**	0.42 (0.08)***	0.42 (0.08)***	0.30 (0.10)**
% white collar	-	-	-	3.25 (1.13)**	3.25 (1.47)*	2.91 (1.19)*
% city centre	-	-	-	8.57 (3.13)**	8.57 (4.09)*	9.35 (2.05)***
R^2	0.59	0.59	0.72	0.62	0.62	0.76

Notes: *All models include SA3 or SA4 f.e. and have n = 280 obs. Standard errors in brackets; ***p < 0.001, **p < 0.01, *p < 0.05*

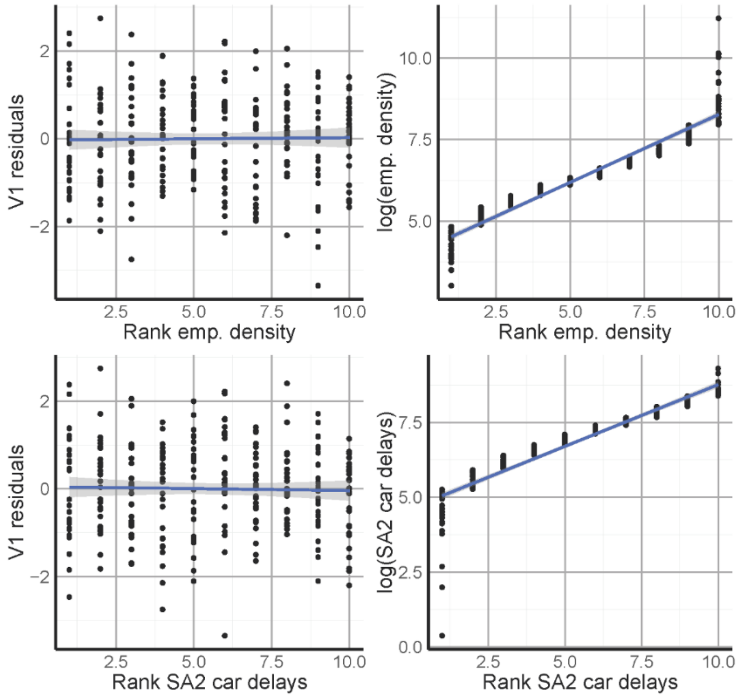
Interestingly, the coefficients for employment density tends to decline when the labour force variables are included in the model, whereas the coefficient for car delays is essentially unchanged. The overall explanatory power of the model increases slightly to 0.62 – 0.76. We re-consider the Labour Force model again in Section 4.4.2.

Finally, we included two measures of travel demand: The average trip distance by car and PT. Neither variable enters the model with the expected positive sign or are they statistically significant.

4.4.2 Instrumental Variables

As noted earlier in Section 2.4.2, due to endogeneity our estimated coefficients may not describe causal effects. We again address the issue of endogeneity using instrumented variables (IV), where we generate instruments the same "rank instrument" process for our two potentially endogenous non-policy variables, namely employment density and car delays. The left and right panels of Figure 9 plot our instruments versus model residuals and the regressors, respectively. The zero correlation in the left panel provides informal evidence our instruments are valid, whereas the positive correlation in the right panel provides informal evidence our instruments are relevant.

Figure 9: Instrument validity (left panel) and relevance (right panel)



We also include an additional rank instrument for SA3 level congestion, which provides us with three instruments in total for our two variables (NB: The third instrument is useful for technical reasons, as it allows us to perform tests for over-identification and exogeneity, specifically the Sargan-Hansen test).

Using these instruments, we re-estimate our SA3 Fixed Effects model, as illustrated in Table 22. The estimated coefficients are similar in magnitude and significance across all specifications and models.

Table 22: Regression Results – Instrumental Variables Baseline Model

Parameters	SA3 Fixed Effects			IV SA3 Fixed Effects		
	V1: Robust	V2: Cluster	V3: WLS	V1: Robust	V2: Cluster	V3: WLS
log(emp. density)	0.64 (0.10)***	0.64 (0.10)***	0.79 (0.12)***	0.65 (0.11)***	0.65 (0.12)***	0.68 (0.12)***
log(delays)	0.43 (0.08)***	0.43 (0.06)***	0.36 (0.12)**	0.37 (0.09)***	0.37 (0.09)***	0.32 (0.15)*
R^2	0.59	0.59	0.72	0.59	0.59	0.71

Notes: All models include SA3 or SA4 f.e. and have n = 280 obs. Standard errors in brackets; ${}^{***}p < 0.001$, ${}^{**}p < 0.01$, ${}^{*}p < 0.05$

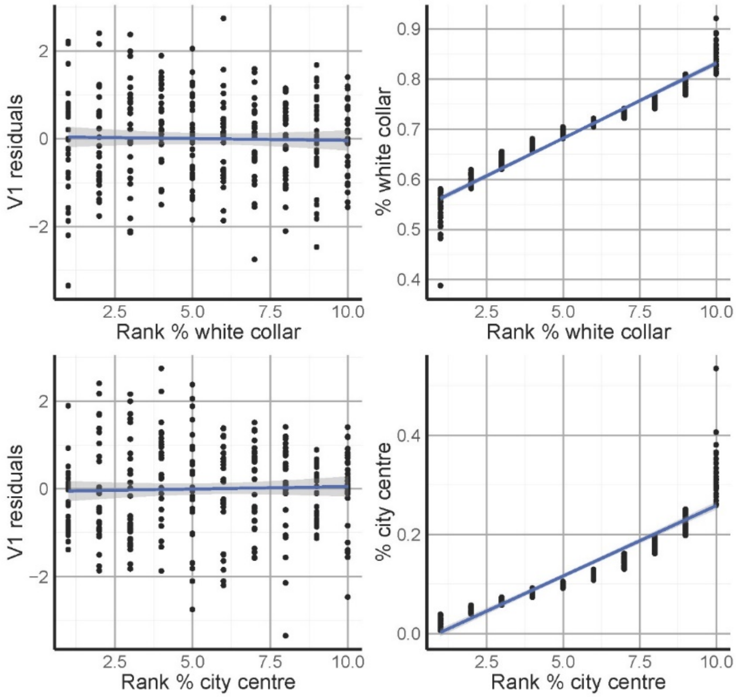
Again, the main effect of the IV model is to reduce the coefficient for employment density in model V3 from 0.79 to 0.68, which brings the latter closer to the coefficients estimated for models V1 and V2 (both in the baseline model and the IV version).

For model V2, formal tests of the instruments indicate that we can:

- Reject the hypothesis our instruments are weak (i.e. our instruments are not irrelevant);
- Accept the hypothesis our variables are exogeneous (i.e. OLS is preferred to IV); and
- Accept the hypothesis our instruments are uncorrelated with model residuals (i.e. our instruments are valid, at least in a statistical sense).

We also estimate an IV version of our Labour Force model treating all non-policy variables as endogenous. This means the Labour Force model includes the instruments illustrated in Figure 9, as well as the additional instrument illustrated in Figure 10.

Figure 10: Instrument validity (left panel) and relevance (right panel)



Results are summarised in Table 23. We find the coefficients for employment density and car delays are largely unchanged in terms of magnitude and significance. The only exception is the car delays coefficient in the IV model V3, which becomes insignificant. In terms of the Labour Force variables, we find that the % white collar is not significant in any of the IV models we test. In contrast, the % city centre variable increases in statistical significance in model V2. Formal tests reject the hypotheses that our instruments are weak and that OLS is unbiased. We also find some evidence of endogeneity.

Table 23: Regression Results – Instrumental Variables Labour Force Model

Parameters	Labour Force			Instrumental Variables		
	V1: Robust	V2: Cluster	V3: WLS	V1: Robust	V2: Cluster	V3: WLS
log(emp. density)	0.46 (0.10)***	0.46 (0.12)***	0.50 (0.12)***	0.49 (0.11)***	0.49 (0.13)***	0.57 (0.13)***
log(delays)	0.42 (0.08)***	0.42 (0.08)***	0.30 (0.10)**	0.34 (0.09)***	0.34 (0.10)***	0.16 (0.11)
% white collar	3.25 (1.13)**	3.25 (1.47)*	2.91 (1.19)*	2.27 (1.25)	2.27 (1.67)	1.80 (1.38)
% city centre	8.57 (3.13)**	8.57 (4.09)*	9.35 (2.05)***	14.88 (5.22)**	14.88 (5.35)**	15.49 (3.85)***
R^2	0.62	0.62	0.76	0.61	0.61	0.74

Notes: *All models include SA3 or SA4 f.e. and have n = 280 obs. Standard errors in brackets;* $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Given the insignificance of the % white collar variable in our IV Labour Force model, we estimate a restricted Labour Force model that omits this variable. Restricting the Labour Force model to only the percentage commutes to the city centre has negligible effects on the resulting coefficients.

In summary, our preferred estimates are drawn from the SA3 Fixed Effects model in Table 19, which implies elasticities for employment density and car delays of 0.64 and 0.43. We prefer the SA3 Fixed Effects V3 model as it is stable and, unlike the Labour Force model, does not rely on ad-hoc definitions of occupational status/or and the geographical extent of the city centre. By extension, the SA3 Fixed Effects model has, in our view, clearer causal implications. That said, we note the confidence intervals for the estimated elasticities of employment density and car delays from the SA3 Fixed Effects model typically overlap with those of the Labour Force model. This provides added confidence that our baseline model is generating relatively robust estimates.

4.5 Summary of Results

Our preferred model is the SA3 Fixed Effects model V2 first reported in Table 19. This model implies elasticities of +0.6396 and +0.4259 for employment density and car delays, respectively. We can use these elasticities, along with the summary statistics in Table 3, to estimate the percentage effect of these two non-policy factors on average PT supply in each city using the following formula:

$$\frac{R_A}{R_S} = \left(\frac{d_A}{d_S}\right)^{\beta_1} \left(\frac{c_A}{c_S}\right)^{\beta_2}$$

Where: (1) S_A and S_S denote PT supply (cost) in actual (A) and sample average (S) scenarios; (2) d_A and d_S denote average employment density in each scenario; (3) c_A and c_S denote congestion levels in each scenario; and (4) β_1 and β_2 denote elasticities for employment density and car delays. Table 24 summarises the results of this calculation for each city.

Table 24: Summary of PT Supply Model – Average Effect for SA3 Fixed Effects (model V2)

City	d_A	c_A	% change
Sydney	1,610	1,783	+37%
Melbourne	1,413	1,612	+21%
SEQ	1,013	890	-24%
Perth	523	851	-51%
Adelaide	661	1,351	-31%
Average (d_S and c_A)	*1,181*	*1,351*	*0%*

We work through these calculations and their implications in more detail in Section 5.1.

5. Implications and Extensions

5.1 Implications

In this section, we work through the fiscal implications of our findings. To summarise, the effect of non-policy factors on net PT expenditure per capita is captured using three types of models:

- *Supply*, as measured in seat-km
- *Productivity*, as measured in vehicle-hours (speed) and vehicle-kilometres (route-kilometres)
- *Revenue*, as measured in monetary terms.

We examine the implications of these models for net PT expenditure per capita in each of the five capital cities for which we have data. To begin with, we ignore the more complex productivity effects and instead focus our analysis solely on understanding the implications of the supply and revenue models, which consider the same non-policy factors, specifically employment density and car delays.

Whereas the revenue model is estimated in monetary terms, the PT supply is estimated in seat-km. For the purposes of our analysis, we assume seat-km exhibit a 1:1 relationship with costs. That is, a 1% increase in seat-km leads to a 1% increase in costs. As well as being simple, this approach has the advantage of being policy-neutral, in the sense it is unaffected by the costs of different modes.[22] We also expect that it closely approximates reality in the five cities that we analyse, which all have established PT networks that are close to, if not at, capacity during peak times.

Conceptually, our analysis then proceeds by comparing two scenarios: **actual outcomes** vis-à-vis the **sample average.** In the sample average scenario, the level for non-policy factors is defined by the average for our sample. That means all cities face the same non-policy factors, in terms of employment density and car delays. By extension, in the sample average scenario all cities will have the same baseline net PT expenditure per capita.

We then calculate the effects of shifting each city from the sample average to their actual outcomes. Differences in net PT expenditure per capita between the actual and average scenarios define the estimated effect of non-policy factors. The effects of non-policy factors on PT supply, or costs, per capita are estimated for each city in the actual (A) and sample (S) scenarios as follows:

$$\frac{S_A}{S_S} = \left(\frac{d_A}{d_S}\right)^{\alpha_1} \left(\frac{c_A}{c_S}\right)^{\alpha_2}$$

Where:

- S_A and S_S denote the supply, or cost;
- d_A and d_S denote average employment density;
- c_A and c_S denote congestion levels; and
- α_1 and α_2 denote elasticities for employment density and car delays from our supply model. Using results for model V2 in Table 5, we have $\alpha_4 = 0.5543$ and $\alpha_5 = 0.4173$.

Similarly, the fiscal implications of the revenue model for net PT expenditure are calculated as:

$$\frac{R_A}{R_S} = \left(\frac{d_A}{d_S}\right)^{\beta_1} \left(\frac{c_A}{c_S}\right)^{\beta_2}$$

[22] Alternatively, one could calculate the mode weighted-average cost per seat kilometre for each city. This would, however, be sensitive to the mix of modes in each city and seems less likely to be policy-neutral.

Where: R_A and R_S denote revenue; β_1 and β_2 denote estimated elasticities for employment density and car delays from our revenue model. Using results for model V2 in Table 19, we have $\beta_1 = 0.6396$ and $\beta_2 = 0.4295$. All other variables are as defined above.

Data for each city in each of the two scenarios are summarised in the first two columns of Table 25. Using this data and the above formulae, we can then estimate the effects of non-policy factors on PT costs and revenue and, by extension, the change in net PT expenditure per capita. For all cities, we set the average index for costs and revenues to 100 and 25, respectively (NB: The choice of index has no effect, instead what matters is the relative size of costs to revenue).

Table 25: Net PT Expenditure per capita – Effects of Non-policy Factors on Costs and Revenue

City	d_i	c_i	Costs	Revenues	N.E.	Effect
Sydney	1,610	1,790	133.31	34.30	99.01	+32%
Melbourne	1,413	1,624	118.90	30.23	88.67	+18%
SE Qld	1,013	881	77.16	18.97	58.19	-22%
Perth	523	854	52.50	12.20	40.30	-46%
Adelaide	661	1,356	72.49	17.25	55.24	-26%
Baseline	1,181	1,359	*100.00*	*25.00*	*75.00*	*0%*

The effect of non-policy factors is then calculated by applying (1) the cost and revenue equations presented above to (2) the data in Table 25. We find non-policy factors add 32% to Sydney's net PT expenditure per capita compared with the average within the sample, while the same factors reduce Perth's net PT expenditure per capita by approximately 46%.

Now we consider the somewhat more complex effect of non-policy factors on PT productivity. Using the results presented in Section 3, we calculate the effect of non-policy factors on bus/tram and heavy rail productivity in each of our five cities by calculating the percentage change in average performance for bus/tram (speed and distance) and heavy rail route (distance). Estimated average productivity effects for each mode are summarised in Table 26.

Table 26: Bus/Tram and Heavy Rail Productivity – Effect of Non-policy Factors

City	Bus and Trams					Heavy rail		
	Speed (hours)		Distance		Cost Effect [%]	Distance		Cost Effect [%]
	KM/hr	%	KM	%		KM	%	
SYD	-2.45	-10.4%	-0.39	-2.2%	4.38%	1.73	5.1%	2.54%
MEL	-0.49	-2.1%	-0.25	-1.5%	0.74%	-1.03	-3.0%	-1.51%
SEQ	0.73	3.1%	0.33	1.9%	-1.00%	4.30	12.6%	6.31%
PER	2.76	11.7%	0.41	2.4%	-4.78%	-2.59	-7.6%	-3.79%
ADL	1.72	7.3%	0.61	3.5%	-2.55%	-1.25	-3.7%	-1.84%

To put these effects on a monetary basis, we must make some assumptions. For buses and trams, we assume vehicle-hours and vehicle kilometres represent 50% and 30% of vehicle operating costs, respectively (NB: Implying 20% is attributable to vehicle fleet, which we do not consider in our analysis and that is likely to make our estimates relatively conservative, in the sense that we underestimate productivity effects).[23] Under these assumptions, non-policy factors are predicted to increase Sydney's bus and tram operating costs by 10.4% x 50% - 2.2% x 30% ≈ 4.38%. Similarly, for heavy rail we assume vehicle-kilometres represent 50% of total operating costs, such that non-policy factors are estimated to increase operating costs by 5.1% x 50% ≈ 2.54%.

[23] These figures are loosely derived from unit cost rates provided in Transport for NSW "Principles and Guidelines for Economic Appraisal of Transport Investment and Initiatives" (source).

We can then combine these productivity effects by assuming bus / tram and heavy rail operating costs represent 50% and 45% of total net PT expenditure, respectively.[24] Sydney's total percentage productivity loss attributable to non-policy factors, L_N, can then be calculated as follows:

$$L_N = 45\% \times 4.38\% + 50\% \times 2.54\% = 3.24\%$$

We calculate productivity effects for each city using the same assumptions on relative operating cost splits between hours vis-à-vis kilometres and bus/tram vis-à-vis heavy rail. Results are summarised in Table 27, where the "productivity factor" (PF) represents the estimated net effect of non-policy factors on PT productivity. A PF smaller than one implies non-policy factors decrease PT productivity, and vice versa for a number greater than one. Our analysis suggests non-policy factors lead to PF in Sydney and Perth that are 3.2% and 4.1% lower and higher than average, respectively.[25]

Table 27: Calculating Productivity Factors

City	Fiscal effect			PF
	Bus/Tram	Rail	Total cost	
SYD	1.97%	1.27%	3.24%	0.968
MEL	0.33%	-0.75%	-0.42%	1.004
SEQ	-0.45%	3.15%	2.70%	0.973
PER	-2.15%	-1.90%	-4.05%	1.040
ADL	-1.15%	-0.92%	-2.07%	1.021

Table 28 incorporates these PF into our analysis of net PT expenditure per capita. Here, the PF are applied to costs while revenue is left unchanged from that in Table 25. We include relative net PT expenditure per capita from Table 25 for ease of comparison, noting it excludes productivity effects.

The final ("+ PF") column in Table 28 contains the estimated total effect of non-policy factors on net PT expenditure per capita. Specifically, we find Sydney's net PT expenditure per capita is 38% higher than average once productivity effects are accounted for. In contrast, non-policy factors *reduce* Perth net PT expenditure per capita from 46% to 49%. Taken together, these two results imply non-policy factors cause net PT expenditure per capita in Sydney to be 87% higher than in Perth. Non-policy factors also cause Sydney's net PT expenditure per capita to be 20% higher than Melbourne.

Table 28: Net PT Expenditure per capita – Adding Productivity Effects

City	Costs		Revenues	Net Expenditure		Effect	
	No PF	+PF		No PF	+PF	No PF	+PF
Sydney	133.31	137.78	34.30	99.01	103.48	+32%	*+38%*
Melbourne	118.90	118.40	30.23	88.67	88.17	+18%	*+18%*
SE Qld	77.16	79.31	18.97	58.19	60.33	-22%	*-20%*
Perth	52.50	50.46	12.20	40.30	38.26	-46%	*-49%*
Adelaide	72.49	71.02	17.25	55.24	53.78	-26%	*-28%*
Baseline	*100.00*		*25.00*	*75.00*		*0%*	

[24] The 5% balance is for ferries.

[25] We tested the sensitivity of PFs to changes in assumptions on resource costs (specifically the percentage of costs associated with vehicle hours vis-à-vis kilometres for bus / tram and heavy rail) and operating costs (specifically the percentage of total budget allocated to bus / tram and heavy rail). This sensitivity analysis calculated the PF different assumptions that led to high and low scenarios. Sydney's PF varied from 2.6% to 3.8% in our low and high scenario, respectively, which represents a range of ≈±20% compared to the 3.2% used in Table 27. The direction (or sign) and order of the PF remained the same in both scenarios for all cities to that in Table 27. For these reasons, we consider the estimated PFs to be robust to underlying assumptions. We also tested a scenario where we assumed productivity effects on vehicle fleet were proportional the combined effects on vehicle-hours and vehicle-kilometres. This reduced the PF for Sydney from 0.968 to 0.950.

The final column in Table 28 represent our best estimate of the effects of non-policy factors on net PT expenditure per capita in the five capital cities that we analyse.

To finish, we address the concern of double-counting. That is, do we risk "double-counting" the effects of non-policy factors on *vehicle-kilometres* (as predicted by our productivity models) with the effects on *seat-km* (as predicted by our supply model). We suggest this risk is low because our productivity and supply models include different non-policy factors. Specifically, our productivity models for vehicle-kilometres, only find that geographical barriers and vertical elevation are statistically significant. Neither of these two non-policy variables feature in our supply models, which consider the effects of employment density and car delays. Given the different non-policy factors included in our productivity and supply models, we see less risk of double-counting. Indeed, any residual indirect effects, such as interactions between non-policy factors, seem as likely to run in the opposite direction, which would cause us to underestimate the effects of non-policy factors.

5.2 Direction of Error

The fiscal implications presented in the previous section are our best estimate of the effects of non-policy factors on net PT expenditure. That said, in our view these numbers are a lower-bound estimate of the effect of non-policy factors on net PT expenditure per capita. More specifically, we are likely to underestimate the effects of non-policy factors for at least two reasons:

- First, we assume labour unit cost rates are independent of city size. In practice, we expect unit cost rates for labour to increase with city size, which will tend to increase Sydney's PT costs relative to other capital cities. Labour costs are a large component of overall PT costs.
- Second, we do not consider productivity effects on the size of the vehicle fleet. The productivity effect of non-policy factors, such as slower speeds and increased route kilometres, is likely to increase the number of vehicles required to deliver PT services.

To give a sense of the magnitude of the error for the second point, if we assume costs of vehicles scale proportionally with the costs of hours and kilometres, then Sydney's productivity factor declines from 0.968 to 0.950, that is, a decline of just under 2%. This in turn would cause Sydney's net PT expenditure premium to increase from the 38% reported in Table 28 to 41%. While not insignificant, we suggest these errors are relatively small in the wider scheme of things.

For these two reasons, we expect the total effect of non-policy factors on net PT expenditure per capita is larger than that explained by our analysis. That is, we are underestimating the degree to which non-policy factors place Sydney at a relative disadvantage compared to other cities.

5.3 Extensions

Our work could be extended in several ways, such as:

- *Verify the accuracy of assumptions used to estimate productivity effects.* One of the downsides of our bottom-up, microeconometric approach is the need for additional assumptions to piece together our results, especially with regards to PT productivity. While we test the sensitivity of our analysis to changes in these assumptions (c.f. footnote 26; pg. 35), further work could revisit these assumptions using detailed data from multiple jurisdictions.
- *Develop a revenue model that includes other capital cities.* The confidential nature of PT ticketing data means that we were unable to estimate our revenue model for cities other than Sydney. While we have confidence in the general direction and magnitude of the effects that we identify, there is merit in extending our revenue model to other cities. As the revenue model uses aggregate revenue data by SA2s, it should be possible for state governments to share data while preserving the confidentiality of the underlying travel patterns.
- *Develop a monetary measure of supply at the SA2 level.* In formulating our supply model, we developed relatively innovative techniques for assigning PT supply (kilometres and hours) to SA2s, which in turn could be converted to seat-kms and seat-hrs. Further work could seek to

monetise these supply measures by applying unit cost rates for each mode and jurisdiction. If this data was then linked to data on fare revenue at the SA2 level (as per the comment above), then it would be possible to model net PT expenditure at the SA2 level.

- *Incorporate ferries into our supply-side model.* Ferries were excluded from our supply-side model because their vehicle-kms largely fall outside of the SA2 that they service. Including them would require calculating their seat-kms and assigning them to the SA2s where they stop, rather than travel through. While we do not expect including ferries will have much of an effect on our overall results, given the small role they play in most cities, this issue seems technically surmountable and should be the subject of further work.

We note that the second and third points above would present the opportunity to integrate our revenue and supply models into one model of net expenditure. In doing so, our methodology could be conceptually aligned with that used by the consultants engaged by the CGC. At the same time, because our methodology makes use of SA2s, our integrated model would provide considerably more detail than SUAs, with associated benefits in terms of identification and statistical power.

In the future, we suggest the CGC consider working at the level of SA2s rather than SUAs. In such an approach, the GTFS feed for each network could be assigned to SA2s using a consistent methodology, as we have done here. To monetise the costs of supply, the CGC would then need states to supply only two high-level pieces of information, specifically:

(1) *Approximate PT unit cost rates*, e.g. per vehicle-hour and vehicle-kilometre; and
(2) *Total contract costs* for each mode within jurisdictions.

In this way, the CGC's relies less on information supplied and processed by states are reduced, ensuring greater consistency in the data between states. This seems preferable to a situation where each state assigns costs and revenues to areas, potentially using different methodologies.

Appendices

Appendix A – The Commission's Work

Consultants engaged by the CGC model net PT expenditure per capita for approximately 70 statistical urban areas ("SUAs") in Australia. By seeking to explain the variation in net expenses across a large sample of urban areas, this work represents a useful advance over prior modelling work undertaken for the CGC (NB: This prior work relied on fewer than 10 cities). Based on their analysis, the consultants engaged by the CGC ultimately recommend a model of the following form:

$$E_i = \beta_0 + \beta_1 dense_i + \beta_2 dist_i + \beta_3 slope_i + \beta_4 \ln\left(pax_i^{train}\right) + \beta_5 \ln\left(pax_i^{bus}\right)$$

Where: (1) E_i denotes net public transport expenses per capita ("net expenses") for SUA i; (2) $dense_i$ denotes the population-weighted density; (3) $dist_i$ denotes the average travel-to-work distance from the census; (4) $slope_i$ denotes mean land slope; (5) pax_i^{train} and pax_i^{bus} denote annual train and bus patronage, respectively; and $\beta_j, j \in \{0, \dots, 5\}$ denotes parameters to be estimated.

The consultants estimate the above model using data for 70 SUAs in Australia. Regression results are summarised below.

Table 1: Estimated parameters for preferred model (Source: Jacobs, 2018)

	Coefficient estimate	Standard error	95% confidence interval	
Intercept	-154.5637**	46.8811	-248.2194	-60.90792
$dense_i$	0.0715307***	0.0200746	0.0314271	0.1116343
$dist_i$	3.411582*	1.647887	0.1195494	6.703616
$slope_i$	6.963933	4.882911	-2.790803	16.71867
$\ln(pax_{i,train})$	18.07401***	4.036532	10.01011	26.13791
$\ln(pax_{i,bus})$	6.719857	6.659917	-6.584856	20.02457

*** p > |t| <=0.1%
** p > |t| <=1%
* p > |t| <=5%
^ p > |t| <=10%

Results suggest net PT expenditure per capita increase linearly with density, average travel-to-work distance, and slope and logarithmically with train and bus patronage. Coefficients for density, distance, and train patronage are statistically significant at the 95% level, while those for slope and bus patronage are not. On the surface, this model addresses several conceptual issues identified by NSW Treasury in earlier submissions to the CGC (NB: These are discussed in Appendix B).

In our view, the methodology adopted by the Consultants can be characterised as a top-down, "macroeconometric" approach to modelling net PT expenditure per capita. By macroeconometric, we mean that the model (1) treats net expenses as a single function, rather a composite function of gross costs and fare revenue and (2) aggregates data to the level of SUAs, smoothing out within SUA variation (for example at the route or area level). The macroeconometric approach used by the Consultants is useful for capturing the total effects of explanatory variables on net expenses.

On the other hand, the macroeconometric approach suffers from three weaknesses, which we seek to address in our own work:

- First, the sample of 70 SUAs remains relatively small, reducing the statistical power of multivariate regressions;
- Second, by focusing on variation between rather than within SUAs, the precise channels are not necessarily identified explicitly and/or precisely; and
- Third, by including rail and bus patronage separately, the model is not necessarily policy neutral, in the sense that it may be sensitive to the cost structures for different modes.

In general, we believe the work undertaken by consultants engaged by the Commission represents a useful advance on earlier work, which we seek to confirm and extend with our analysis.

Appendix B – Treasury NSW's Conceptual Framework

Compared with the current revenue allocation formula previously used by the CGC, the work undertaken by Consultants for the CGC aligns more with the conceptual framework previously advanced by Treasury NSW in their CGC submissions, as illustrated below.

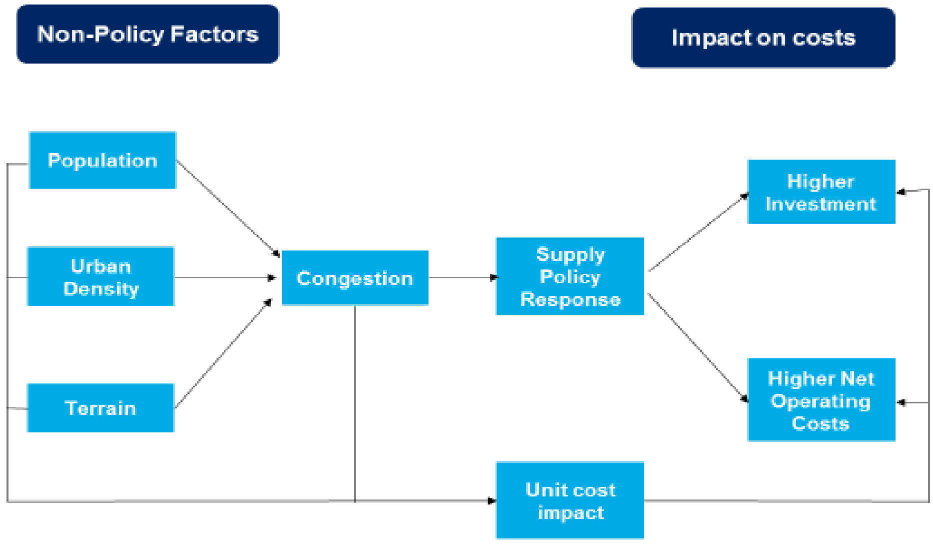


Figure 1: Conceptual model – Impact of non-policy factors on net expenses (Source: NSW Treasury, 2018).

This conceptual model highlights how three "Non-Policy Factors", specifically population, urban density, and terrain, may combine to affect PT operating costs, operating via the twin channels of supply-side policy responses and unit costs.

Of the non-policy factors identified in Treasury's conceptual model, the model recommended by the Consultants engaged by the Commission directly addresses two, namely urban density ("density") and terrain ("slope"), which they find increase net expenses. The third factor, namely population, is not directly addressed in the recommended model, but is likely to be indirectly captured via train and bus patronage variables, both of which tend to increase with population.