

Multimodal pricing and optimal design of urban public transport: the interplay between traffic congestion and bus crowding

Alejandro Tirachini ^{1*}

David A. Hensher ²

John M. Rose ²

¹ Transport Engineering Division
Civil Engineering Department
Universidad de Chile, Santiago, Chile
Tel: +56 2 2978 4380
alejandro.tirachini@ing.uchile.cl

² Institute of Transport and Logistics Studies (ITLS)
The University of Sydney Business School,
The University of Sydney, NSW 2006, Australia
Tel: +61 2 9351 0169
david.hensher@sydney.edu.au
john.rose@sydney.edu.au

* Corresponding author

Citation: Tirachini, A., Hensher, D. A., Rose, J. M., 2014. Multimodal pricing and optimal design of public transport services: the interplay between traffic congestion and bus crowding. **Transportation Research Part B**, forthcoming.

Abstract

The interplay between congestion and crowding externalities in the design of urban bus systems is identified and analysed. A multimodal social welfare maximisation model with spatially disaggregated demand is developed, in which users choose between travelling by bus, car or walking in a transport corridor. Optimisation variables are bus fare, congestion toll, bus frequency, bus size, fare collection system, bus boarding policy and the number of seats inside buses. We find that optimal bus frequency results from a trade-off between the level of *congestion inside buses*, i.e., passengers' crowding, and the level of *congestion outside buses*, i.e., the effect of frequency on slowing down both buses and cars in mixed-traffic roads. A numerical application shows that optimal frequency is quite sensitive to the assumptions on crowding costs, impact of buses on traffic congestion, and overall congestion level. If crowding matters to users, buses should have as many seats as possible, up to a minimum area that must be left free of seats. If for any other reason planners decide to have buses with fewer seats than optimal (e.g., to increase bus capacity), frequency should be increased to compensate for the discomfort imposed on public transport users. Finally, the consideration of crowding externalities (on both seating and standing) imposes a sizeable increase in the optimal bus fare, and consequently, a reduction of the optimal bus subsidy.

Keywords: Bus design, congestion, crowding, fare, frequency, quality of service, walking

1. Introduction

When deciding whether or not to undertake a trip by public transport, travellers are influenced by a number of characteristics or attributes of the public transport mode, including accessibility, waiting time, travel time, price, reliability, comfort and safety. Demand is sensitive to the overall quality of service, which in turn depends on the design of the system; hence understanding the economic nature of urban public transport operations is crucial as a means to ensure the efficiency of a public transport network and, ultimately, the sustainability of the entire transport system. From a transport planner's perspective, the challenge associated with the design of public transport services lays in the myriad number of trade-offs that need to be considered at once, in order to establish an *optimal* service design. For example: increasing bus frequency reduces waiting time for users but increases the cost of operation; increasing the number of bus stops reduces users' access time but increases bus riding time for all-stop services; and investing in a quicker fare collection technology and dedicated road infrastructure for buses reduces bus travel time (and consequently may reduce operating cost) but increases capital cost. In this paper we introduce two trade-offs within the microeconomic modelling framework of optimal supply levels and pricing of an urban bus route, that are crucial to the level of service offered to public transport users: (i) the interaction between road congestion and passenger crowding externalities when setting supply levels of public transport, and (ii) the decision on the number of seats that public transport vehicles should have.

First, the influences of congestion and crowding on optimal bus service frequency have been treated independently in the literature. On the one hand, Jara-Díaz and Gschwender (2003) show that passenger crowding externalities push optimal public transport frequency up, with a total cost minimisation model without road congestion. On the other hand, Tirachini and Hensher (2011) find that the existence of bus congestion in the form of queuing delays at bus stops pushes optimal frequency down, with a total cost minimisation model that ignores crowding externalities. Moreover, buses may also slow cars down in shared roads. Therefore, there are possible counter-effects of congestion and crowding on optimal frequency that need to be addressed simultaneously. In this paper we propose a multimodal social welfare maximisation model that includes both passenger crowding and mixed-traffic congestion externalities, and find that optimal bus frequency is the result of crowding and congestion acting as colliding forces.

Second, going beyond microeconomic models that optimise public transport frequency and/or vehicle size to set supply levels (e.g., Mohring, 1972; Jansson, 1980; Oldfield and Bly, 1988; Chang and Schonfeld, 1991), we look at the internal design of vehicles by including the number of bus seats as a decision variable that influences both comfort and capacity. Different configurations of vehicles regarding number of seats and space for standees are relevant for the level of crowding and standing externalities in public transport (Whelan and Crockett, 2009). This is a key insight from the estimation of crowding and standing disutilities that has not been given attention in the literature on the design and optimisation of public transport systems. Microeconomic models that have included the level of crowding as an influence on the value of in-vehicle time savings do not distinguish between passengers sitting and standing (Jara-Díaz and Gschwender, 2003; Tirachini *et al.*, 2010); whereas Kraus (1991) applies a premium on

the value of travel time savings for passengers standing, but his research is concerned with the marginal cost and pricing of services considering the discomfort of standing, rather than with the design of vehicles. Thus, even though crowding and discomfort externalities have been analysed in the literature, previous studies always assumed a given internal design or layout of the vehicles involved, i.e., a given bus or train capacity. In short, it is assumed that size implies capacity. However, reducing the number of seats increases bus capacity by allowing more standees; thus a seat implies a trade-off between comfort and capacity that is allowed for in our model.

The analytical approach consists of a social welfare maximisation model with disaggregated origin and destination demand, and multiple travel alternatives. The model is applied to a single transport corridor in Sydney, Australia, which allows us to obtain detailed measures of crowding levels section by section. In contrast to other social welfare maximisation models (e.g., De Borger *et al.*, 1996; Proost and Van Dender, 2004; Wichensin *et al.*, 2007; Ahn, 2009; Parry and Small, 2009; Jansson, 2010; Basso *et al.*, 2011), our approach provides a more comprehensive modelling of the bus mode, including bus frequency, bus size, fare collection system, bus boarding policy, number of bus seats and fare level as decision variables. We also show that the inclusion of a non-motorised mode (walking) as an alternative to choosing bus and car for short trips may have a significant role when the transport system is optimised in highly congested scenarios¹. Results are discussed for several scenarios with different demand levels and modelling assumptions.

We show that optimal bus frequency is quite sensitive to the assumptions regarding crowding costs, impact of buses on traffic congestion and the overall congestion level. In particular, if the planner takes into account that crowding matters to users, our numerical application shows that bus frequency should increase (for a given bus size) with demand even under heavy congestion, however that might not be the case if the crowding externality is not accounted for, in which case an increase of total demand might be met by a decrease of both frequency and number of seats per bus, at the expense of crowding passengers inside buses and making more passengers stand while travelling. Likewise, we find that buses should be designed with as many passenger seats as possible (up to a minimum area that must be left free of seats for an aisle, next to the driver and doors, for a wheelchair and other possible uses). If for any other reason planners decide to have buses with fewer seats than optimal (e.g., to increase bus capacity), frequency should be increased to compensate for the discomfort imposed on public transport users.

The remainder of the paper is organised as follows. The theoretical model is developed in Section 2, including assumptions and definitions (Section 2.1), demand and crowding modelling (2.2), travel time and congestion (2.3), internal bus layout (2.4) and operator cost items (2.5); the section concludes with the formulation of the social welfare maximisation problem. Section 3 presents the numerical application of the model to Sydney and discussion of results in several scenarios. Conclusions are provided in Section 4.

¹ Excluding non-motorised alternatives, the properties of the bimodal car-bus competition for user equilibrium and/or system optimum solutions are analysed in a number of contributions, e.g., Ahn (2009), Li *et al.* (2012), and Gonzales and Daganzo (2012); 2013). The influence of an un-congestible non-motorised alternative on the optimal public transport fare is analytically studied by Tirachini and Hensher (2012) with a three-mode model.

2. Model set up

2.1 Assumptions and definitions

We consider a linear bi-directional road of length L and a single period of operation with directions denoted as 1 and 2. The road is divided into P zones denoted as $i \in \{1, \dots, P\}$, and the total demand Y^{ij} per origin-destination pair (i, j) is fixed. The distance between zone i and zone $i+1$ is denoted as L_i such that $L = \sum_{i=1}^{P-1} L_i$, as shown in

Figure 1. Users can choose to travel by car (a), bus (b) or to walk (e)². Then, if y_m^{ij} is the travel demand for mode m between zones i and j , it holds that:

$$Y^{ij} = \sum_m y_m^{ij} = y_a^{ij} + y_b^{ij} + y_e^{ij} \quad (1)$$

Let f_{a1}^i be the traffic flow in section i , between zone i and zone $i+1$ (direction 1) and f_{a2}^i be the traffic flow in section i , between zone $i+1$ and zone i (direction 2). The decision variables of the problem are denoted as follows:

f_b : bus frequency [bus/h]

s_b : bus length [m]

Δ : fare collection technology and boarding policy (one-door or all-door boarding)

n_{seat} : number of seats inside a bus

τ_a : car toll [\$/trip]

τ_b : bus fare [\$/trip]

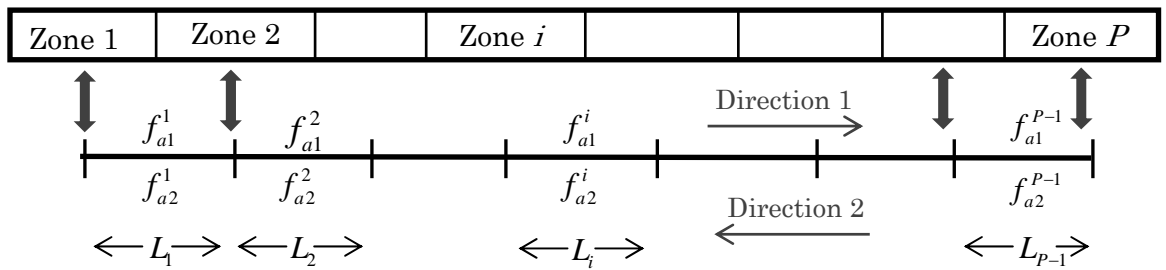


Figure 1: Transport corridor diagram

It is assumed that there is only one bus stop per zone³ and that the travel distance between zones is the same for the three modes. Bus frequency is assumed to be

² The model can be easily extended to more travel alternatives such as rail.

³ The location of bus stops is fixed in this model, which allows us to know the number of passengers that a bus carries in each segment of the route (between two consecutive zones). For a review of models that optimise the spacing of bus stops see Tirachini (2014).

continuous, whereas options on bus lengths are constrained by the size of commercial vehicles; four sizes are considered in the application of the model. These are mini (8 m. long, 1 or 2 doors), standard (12 m. long, 2 or 3 doors), rigid long (15 m. long, 3 or 4 doors) and articulated (18 m. long, 4 doors). The number of seats n_{seat} can be freely chosen subject to lower and upper bounds, the former is given by a minimum number of seats per bus that is exogenously decided in order to provide a minimum level of service, whereas the latter is determined by a minimum area on a bus that must be clear of seats (i.e., aisle, doors, space for a wheelchair, area next to the driver). Following Tirachini and Hensher (2011), we consider four alternative fare collection technologies: on-board payment with (i) cash, (ii) magnetic strip (with contact) and (iii) smartcard (contactless), plus (iv) off-board payment (on the bus stop). The bus boarding and alighting policy can be chosen as well, two alternatives are available to implement in buses with more than one door: (a) simultaneous boarding and alighting, in which boarding is allowed at the front door only while alighting takes place at the back(s) doors, and (b) sequential boarding and alighting, in which boarding is allowed at all doors giving priority to passengers alighting to unload first. In principle, we assume that cars and buses share the right-of-way and that bus stops do not directly affect cars, an assumption that is revised in Section 3.3.

2.2 Demand modelling and crowding

Previous research has shown the influence of crowding and standing on increasing the value of travel time savings (e.g., Maunsell and Macdonald, 2007; Whelan and Crockett, 2009; Hensher *et al.*, 2011; Wardman and Whelan, 2011). In this paper, mode choice models that include the proportion of available seats and the density of standees as attributes for buses are estimated. Data collected from a stated choice survey conducted in Sydney in 2009 is used to this end; the experimental design, study area, sample size and socioeconomic characteristics of respondents are described at length in Hensher *et al.* (2011), who estimate the crowding disutility as a function of the proportion of users seating (which affects the probability of getting a seat), and the total number of users standing, in order to estimate the willingness to pay to get a seat as a function of the number of people sitting and standing. In this paper, we use the density of standees per square metre, instead of the total number of standees, to represent the disutility of crowding, in order to have a common base for the application of the model with different internal bus layouts regarding allocation of space for seating and standing.

Let U_m^{ij} be the utility associated with travel by mode m in origin-destination (OD) pair (i, j) . In order to analyse differences in optimal bus service design due to alternative assumptions regarding user's valuations of seating, standing and crowding levels inside buses, we propose three different models that incorporate attributes representing the number of passengers seating and standing, interacting with travel time; these models will be compared with a specification that ignores any crowding or standing cost. The models, named M1 to M4, are described as follows:

- M1: Crowding is not explicitly considered as a source of disutility for users (eq. 2).
- M2: Only the density of standees [pax/m^2] imposes an extra discomfort cost (eq. 3).

- M3: The density of standees and the proportion of seats occupied are sources of disutility (eq. 4).
- M4: The density of standees and the proportion of seats occupied are squared in the utility function (eq. 5).

$$\text{Bus - M1: } U_b^{ij} = \alpha_b^{\text{M1}} + \beta_a^{\text{M1}} t_{ab}^i + \beta_h^{\text{M1}} h_b + \beta_{vb}^{\text{M1}} t_{vb}^{ij} + \beta_c^{\text{M1}} \tau_b \quad (2)$$

$$\text{Bus - M2: } U_b^{ij} = \alpha_b^{\text{M2}} + \beta_a^{\text{M2}} t_{ab}^i + \beta_h^{\text{M2}} h_b + \beta_{vb}^{\text{M2}} t_{vb}^{ij} + \beta_c^{\text{M2}} \tau_b + \beta_{den}^{\text{M2}} n_{den} t_{vb}^{ij} \quad (3)$$

$$\text{Bus - M3: } U_b^{ij} = \alpha_b^{\text{M3}} + \beta_a^{\text{M3}} t_{ab}^i + \beta_h^{\text{M3}} h_b + \beta_{vb}^{\text{M3}} t_{vb}^{ij} + \beta_c^{\text{M3}} \tau_b + \beta_{den}^{\text{M3}} n_{den} t_{vb}^{ij} + \beta_{seat} p_{seat} t_{vb}^{ij} \quad (4)$$

$$\text{Bus - M4: } U_b^{ij} = \alpha_b^{\text{M4}} + \beta_a^{\text{M4}} t_{ab}^i + \beta_h^{\text{M4}} h_b + \beta_{vb}^{\text{M4}} t_{vb}^{ij} + \beta_c^{\text{M4}} \tau_b + \beta_{den2} n_{den}^2 t_{vb}^{ij} + \beta_{seat2} p_{seat}^2 t_{vb}^{ij} \quad (5)$$

In (2) to (5), t_{ab}^i is the access time at zone i , h_b is the headway between two consecutive buses⁴, t_{vb}^{ij} is the in-vehicle time between zones i and j , τ_b is the bus fare, n_{den} is the density of standees per square metre, p_{seat} is the proportion of seats been used, α_b is a modal constant (which will be calibrated to predict an observed modal split) and β_k are the parameters associated to the different attributes. For each model M1 to M4, the utility of the alternative modes (car and walk) have the same specification:

$$\text{Car: } U_a^{ij} = \beta_{va}^{\text{M}} t_{va}^{ij} + \beta_c^{\text{M}} (c_r^{ij} + \tau_a) / o_r \quad (6)$$

$$\text{Walk: } U_e^{ij} = \alpha_e + \beta_{ve}^{\text{M}} t_{ve}^{ij} \quad (7)$$

where c_r^{ij} is the car running cost to travel between zones i and j , τ_a is the road charge (decision variable) and o_r is the average car occupancy rate (therefore U_a^{ij} is the average utility of car users). In expressions (2) to (6), bus fare and car toll are independent of the origin and destination of trips, an assumption that is consistent with an area charging scheme; extensions to distance-based fares and tolls are discussed in Section 4 (Conclusions).

Assuming a multinomial logit model for the estimation of demand, the number of trips by mode m in OD pair (i, j) is given by:

$$y_m^{ij} = Y^{ij} \frac{e^{U_m^{ij}}}{\sum_n e^{U_n^{ij}}} \quad \forall i, j \quad (8)$$

where Y^{ij} is the total demand between zones i and j . The estimation of parameters for models M1 to M4 is shown and discussed in Appendix A1. In this framework, the consumer surplus B is given by the logsum formula:

⁴ The assumed linearity in the headway attribute as a proxy of waiting time cost is more appropriate in high frequency services in which *ex-ante* schedules do not exist or are not relevant in the departure time decisions of users. A situation with low frequency in which passengers follow a schedule is better modelled with the inclusion of two or more headway parameters or with a non-linear formulation. The resulting frequencies in Table 3 are over 20 veh/h (headways of less than 3 minutes), which is consistent with the assumption of high frequency services.

$$B = \sum_{ij} \frac{Y^{ij}}{I_u} \ln \sum_m e^{U_m^{ij}} + B_0 \quad (9)$$

where I_u is the marginal utility of income⁵, equal to minus the cost parameter β_c^M estimated with the choice models⁶, and B_0 is a constant that has no effect on the solution of the problem, and therefore can be set to zero.

2.3 Travel time, congestion and bus stop delay

We assume that buses and cars share the right-of-way, which is subject to congestion. Furthermore, buses have to stop at bus stops to load and unload passengers. Bus stops are also subject to congestion in the form of queuing delays when the bus frequency is high and/or the dwell time is long. Taking direction 1 for illustration, we model travel time between zone i and zone $i+1$ by car (t_{va1}^i) and bus (t_{vb1}^i) as a function of traffic flow and bus frequency by using the well-known Bureau of Public Roads (BPR) formula:

$$t_{va1}^i(f_{a1}^i, f_b) = t_{a0}^i \left[1 + \alpha_0 \left(\frac{f_{a1}^i + \varphi(s_b) f_b}{K_r} \right)^{\alpha_1} \right] \quad (10)$$

$$t_{vb1}^i(f_{a1}^i, f_b) = t_{b0}^i \left[1 + \alpha_0 \left(\frac{f_{a1}^i + \varphi(s_b) f_b}{K_r} \right)^{\alpha_1} \right] + t_{s1}^i \quad (11)$$

where t_{a0}^i , t_{b0}^i , α_0 and α_1 are parameters (t_{a0}^i and t_{b0}^i are the free-flow travel times by car and bus, respectively), $\varphi \geq 1$ is the passenger car equivalency factor of a bus, which depends on the bus length s_b , and K_r is the capacity of the road. The travel time by bus includes the delay due to bus stops, t_{s1}^i , which consists of the acceleration and deceleration delay t_{ac1}^i , the average queuing time t_{q1}^i and the dwell time t_{d1}^i , i.e.,

$$t_{s1}^i = t_{ac1}^i + t_{q1}^i + t_{d1}^i \quad (12)$$

The delay in the process of accelerating and decelerating at bus stops is modelled by assuming uniform acceleration and deceleration; thus the extra stopping delay on top of the uniform travel time given by the running speed v_{b1}^i , is expressed as (13).

$$t_{ac1}^i = \frac{v_{b1}^i}{2} \left(\frac{1}{a_0} + \frac{1}{a_1} \right) \quad (13)$$

The queuing time t_{q1}^i is a measure of the external congestion caused by a bus stop, observed when a bus arrives at a stop and all berths are occupied. This delay is

⁵ The marginal utility of income is assumed constant, i.e., we ignore income effects on demand (Jara-Díaz and Videla, 1989; Jara-Díaz, 2007).

⁶ Note that β_c^M is estimated with the choice of motorised modes only because walking is for free.

commonly present in high frequency services, but it may also occur in poorly controlled low frequency services where buses tend to bunch. Following Fernández *et al.* (2000) and Tirachini and Hensher (2011), we use the bus stop simulator *IRENE* (Gibson *et al.*, 1989; Fernández and Planzer, 2002) to estimate the queuing delay t_q [s/bus] as a function of the design of the bus stop (number and length of berths), bus length, bus frequency and average dwell time. Bus stops with one, two and three linear berths are considered. The regression model for the simulated data has the exponential form (14), as developed in Tirachini (2014) and explained in Appendix A2:

$$t_q = 0.001 \left[b_0 + b_{l1}s_b + (b_{d1} + b_{d2}Z_2 + b_{d3}Z_3)t_d \right] e^{0.001 f_b [b_f + b_{l2}s_b + (b_{d4} + b_{d5}Z_2 + b_{d6}Z_3)t_d]} \quad (14)$$

where s_b [m] is the bus length, t_d [s/bus] is the dwell time, f_b [veh/h] is the bus frequency and β_0 , β_{l1} , β_{l2} , β_{d1} , β_{d2} , β_{d3} , β_{d4} , β_{d5} , β_{d6} and β_f are estimated parameters; factors 0.001 are introduced for scaling of the parameters (see Appendix A2 for further details). Z_2 and Z_3 are dummy variables defined as follows:

$$Z_2 = \begin{cases} 1 & \text{if bus stop has two berths} \\ 0 & \text{otherwise} \end{cases}$$

$$Z_3 = \begin{cases} 1 & \text{if bus stop has three berths} \\ 0 & \text{otherwise} \end{cases}$$

Equation (14) is defined for bus stops with one, two and three berths; therefore it allows the number of berths to be a variable. In the numerical application of Section 3 we assume the existence of two berths per bus stop. The estimation of the dwell time per stops requires the cases with boarding allowed at all doors (*TnBn*) and at the front door only (*TnB1*) to be addressed separately, since in *TnBn* boarding and alighting is sequential at all doors, whereas in *TnB1* boarding at the front door occurs simultaneously with alighting at the rear doors. These two cases are summarised in expression (15)

$$t_d^i = \begin{cases} c_{oc} + p_b \beta_b \lambda^+ + p_a \beta_a \lambda^- & \text{if boarding at all doors (TnBn)} \\ c_{oc} + \max \{ \beta_b \lambda^+, p_a \beta_a \lambda^- \} & \text{if boarding at front door only (TnB1)} \end{cases} \quad (15)$$

where c_{oc} is the time to open and close doors, β_a and β_b are the average alighting and boarding times per passenger, λ^+ and λ^- are the number of passengers boarding and alighting a bus at the bus stop and factors p_a and p_b are the proportion of passengers boarding and alighting at the busiest door, which are given in Appendix A3 (Table A3.1). Equations (13), (14) and (15) conclude the derivation of the delay at bus stops (12).

2.4 The choice of bus size and internal layout

Using data from London, Jansson (1980) finds a linear relationship between bus running costs and bus size measured as the number of seats per bus, a relationship that has been used by Jansson and other authors to find the optimal size of buses in urban routes (e.g.,

Jara-Díaz and Gschwender, 2003; Tirachini and Hensher, 2011) under the implicit assumption that there is a unique relationship between bus size and capacity, measured as number of seats or total number of passengers that can be carried. However, the number of passengers that a bus can carry is not only given by the bus size, but also by the internal layout of space allocated to seating and standing, as a passenger sitting takes up more space than a passenger standing. A standard value for the area needed for a passenger sitting is 0.5 square metres (TRB, 2003), whereas, depending on crowding conditions, passengers standing may have a density of up to five or six passengers per square metre, and as such the minimum area required by a standee is approximately 0.17-0.20 square metres, i.e., less than half the space required for a person seated. Therefore, if the number of seats inside a bus can be varied, there is no one-to-one relationship between capacity and bus size, and the final capacity of a bus is the outcome of decisions on both the bus length and internal layout regarding seating and standing areas. In this context, capacity is not an absolute value, but rather a function of the number of seats and the maximum density of standees that is acceptable to have or provide, given by policy, demand and cultural constraints⁷.

Several physical constraints need to be considered when deciding the number of seats, including minimum space for aisles, doors in front of the bus (next to the driver) and for a wheelchair, that must be clear of seats. Let $A(s_b)$ be the total area available in a bus for seating and standing, which is a function of the bus length s_b . If P_s is the proportion of A allocated to seating, the areas for seating A_{seat} and standing A_{stand} can be formulated as:

$$A_{seat}(P_s, s_b) = P_s A(s_b) \quad (16)$$

$$A_{stand}(P_s, s_b) = [1 - P_s] A(s_b) \quad (17)$$

If a_{seat} is the area required by one bus seat (m²), then the number of seats n_{seat} per bus is

$$n_{seat} = \frac{A_{seat}}{a_{seat}} \quad (18)$$

For the estimation of in-vehicle time costs it is necessary to determine the proportion of seats being occupied p_{seat} and the density of standees n_{den} (if any) in each segment of a bus trip. Taking direction 1, if λ^{i+} and λ^{i-} are the number of passengers getting on and off a bus at stop i , the number of passengers q^i on board a bus between stops i and $i+1$ is calculated recursively:

$$q^0 = 0 \quad (19)$$

$$q^i = q^{i-1} + \lambda^{i+} - \lambda^{i-} \quad \forall i \in \{1, P-1\} \quad (20)$$

⁷ In some crowded public transport systems around the world (e.g., Moscow, Sao Paulo, Santiago de Chile, Tokyo), it is not unusual to operate at crush capacity, with 6 passengers standing per square metre in peak periods, however, such a high density of standees could not be acceptable in other regions.

Separating q^i among passengers seating q_{seat}^i and standing q_{stand}^i , we can obtain p_{seat}^i and n_{den}^i as follows:

$$p_{seat}^i = \frac{q_{seat}^i}{n_{seat}} = \frac{\min\{n_{seat}, q^i\}}{n_{seat}} \quad (21)$$

$$n_{den}^i = \frac{q_{stand}^i}{A_{stand}} = \frac{q^i - q_{seat}^i}{A_{stand}} \quad (22)$$

To calculate the area available for seating A_{seat} and standing A_{stand} , we need an estimation of the area occupied by seats, standees, doors and other elements. The U.S. Transportation Research Board recommends the following values (TRB, 2003):

Table 1: Area occupied by passengers sitting, standing and other objects

Situation	Projected area [m ²]
Standing	0.15-0.20
Standing with briefcase	0.25-0.30
Standing with daypack	0.30-0.35
Standing with suitcase	0.35-0.55
Transverse seating	0.50
Longitudinal seating	0.40
Wheelchair space	0.95
Rear door	0.80

We use Table 1 and the following assumptions in order to calculate seating and standing areas, feasible numbers of seats and total bus capacity:

(A1) Buses have transverse seating only, therefore 0.5 m² is the value used for the area occupied by passengers sitting.

(A2) The maximum density of standees d_{max} is around 6.7 pax/m², equivalent to an area of 0.15 m² per standee in Table A4.1. However, given the Sydney context d_{max} is set as 4 pax/m² in the Military Road application (Section 3).

(A3) Buses are 2.55 metre wide (regardless of length)

(A4) The front area must be left clear of passengers, for the driver and front door. This area is 1.5 metre long.

(A5) Next to each rear door there has to be a 0.8 m² area clear of standees. The number of doors per bus is denoted as n_{doors}

(A6) Buses must have a 0.95 m² area reserved for wheelchairs.

Using (A3) to (A6), the total area A (m²) available for seating and standing is:

$$A = A_{sit} + A_{stand} = 2.55(s_b - 1.5) - 0.8(n_{doors} - 1) - 0.95 \quad (23)$$

And the capacity of a bus (maximum number of passengers that can be accommodated) is:

$$K(s_b, P_s, n_{doors}) = \left[\frac{P_s}{a_{seat}} + (1 - P_s)d_{max} \right] A \quad (24)$$

Constraints

(C1) An aisle is provided in the centre of the bus, with a minimum width of 0.5 metre. This aisle does not necessarily have to cover the full length of the bus as the back row may have a seat in the middle (where the aisle ends). Therefore, assuming that 1.5 metre is left at the front and 0.7 metres is used for a seat at the back, the minimum area that has to be reserved for the aisle is $A_{stand}^{min} = 0.5(s_b - 2.2)$. Then, the number of seats is upper bounded by:

$$n_{seat} \leq n_{seat}^{max} = \frac{A - A_{stand}^{min}}{a_{seat}} \quad (25)$$

(C2) A minimum number of seats must be provided, i.e., the proportion P_s of A allocated to seating has a lower bound P_s^{min} , which is arbitrarily decided (e.g., $P_s^{min} = 0.3$ meaning that at least 30 percent of the available area must be reserved for passengers sitting). Therefore,

$$n_{seat} \geq n_{seat}^{min} = \frac{P_s^{min} A}{a_{seat}} \quad (26)$$

2.5 Bus operator cost and problem formulation

In this section, we formulate bus operator costs. Let operator cost be divided into four components:

c_1 : Station infrastructure cost [\$/station-h]

c_2 : Personnel costs (crew) and vehicle capital costs [\$/bus-h], and

c_3 : Running costs (fuel consumption, lubricants, tyres, maintenance, etc.) [\$/bus-km]

c_4 : Implementation cost related to the fare payment technology (e.g., software requirements) [\$/h]

The station cost c_1 (eq. 27), consists of two components: the station infrastructure cost which depends on the bus length i.e., $c_{10}(s_b)$, and the cost of fare vending machines and fare collection readers (if validation is undertaken at the station and not on bus), $c_{11}(\Delta)$, where the dependency on Δ denotes the fare payment method.

$$c_1(s_b, \Delta) = c_{10}(s_b) + c_{11}(\Delta) \quad (27)$$

Second, the cost per bus-hour c_2 also has two elements: the personnel cost (wages) and the capital cost of a vehicle, which includes the cost of the fare collection readers (validation devices) installed in buses. If $c_{20}(s_b)$ is the cost associated to bus size and $c_{21}(\Delta, s_b)$ is the cost of the fare collection readers (which in turn depends on bus size s_b as well, because for on-board payment methods, fare collection devices are installed at each boarding door, see Appendix A4), the total cost per bus-hour c_3 is simply expressed as:

$$c_2(s_b, \Delta) = c_{20}(s_b) + c_{21}(\Delta, s_b) \quad (28)$$

The third component of operator cost is the running cost per vehicle-kilometre $c_3(s_b)$, which includes fuel consumption, lubricants, tyres, maintenance, etc., and depends on the size of the bus (the capital cost of garages, which depends on bus and fleet size, is ignored). Finally, $c_4(\Delta)$ accounts for the cost of software and implementation of the alternative fare collection technologies and boarding and alighting policies. The estimation of parameters for the operator cost components is given in the Appendix A4. With this, the total operator cost C_o can be defined as:

$$C_o(s_b, \Delta, F) = c_1(s_b, \Delta)S + c_2(s_b, \Delta)F + c_3(s_b)VF + c_4(\Delta) \quad (29)$$

where S is the number of bus stops, F is the fleet size and V is the operating speed (commercial speed including running and stops). The fleet size requirement is given by $F = f_b T_c$, in which T_c is the cycle or round-trip time (given by the summation of bus travel time (11) at all sections and both directions, plus a scheduled slack time at termini if required). Rewriting T_c as $2L/V$, we obtain that the third term in (18) does not depend on the operating speed and passenger demand. Therefore, the final expression for bus operator cost is given by (30).

$$C_o(f_b, s_b, \Delta) = c_1(s_b, \Delta)S + c_2(s_b, \Delta)f_b T_c(f_b, s_b, \Delta) + 2c_3(s_b)Lf_b + c_4(\Delta) \quad (30)$$

Importantly, we are assuming that the number of seats inside a bus (and consequently, the number of passengers) has no effect on the bus capital cost, which is only determined by the bus size and arrangements regarding fare collection readers (i.e., the cost of seats if assumed negligible relative to the cost of the bus). After obtaining an expression for the operator cost (30), we can formulate the social welfare maximisation problem as follows:

$$\text{Max} \quad SW = \sum_{ij} \frac{y^{ij}}{I_u} \ln \sum_m e^{U_m^{ij}} + \sum_{ij} y_a^{ij} \tau_a + \sum_{ij} y_b^{ij} \tau_b - C_o \quad (31)$$

Subject to

$$\max_i \{y_{b1}^i, y_{b2}^i\} \leq \kappa f_b K (s_b, n_{seat}) \quad (32a)$$

$$n_{seat}^{\min} \leq n_{seat} \leq n_{seat}^{\max} \quad (32b)$$

$$f_b^{\min} \leq f_b \leq f_b^{\max} \quad (32c)$$

$$s_b \in \{s_{b1}, \dots, s_{b4}\} \quad (32d)$$

$$\Delta \in \{\Delta_1, \dots, \Delta_6\} \quad (32e)$$

$$y_m^{ij} = Y^{ij} \frac{e^{U_m^{ij}}}{\sum_n e^{U_n^{ij}}} \quad \forall i, j \quad (32f)$$

Inequality (32a) is a capacity constraint that ensures that the bus transport capacity ($\kappa f_b K$) is large enough to accommodate the maximum bus load (y_{b1}^i and y_{b2}^i are the bus demands on section i in directions 1 and 2, respectively; see Figure 1); κ is design factor introduced to have spare capacity to absorb random variations in demand (for example, $\kappa = 0.9$ assumes a system in which capacity is 90 percent of the maximum average demand). (32b) states that the number of seats is constrained by minimum and maximum values (equations 25 and 26). Frequencies are also constrained by a minimum policy frequency f_{min} (set to have a minimum level of service, if desired) and the maximum feasible frequency f_{max} as given in expression (32c). Expressions (32d) and (32e) establish that bus size s_b and the boarding and alighting policy and fare collection technology Δ are taken from available choices. Finally, a set of equilibrium constraints is necessary because modal choice depends on travel times, which in turn depend on modal choice; inducing a fixed-point problem that is solved by iterating between modal choice and travel times until convergence is reached, using equations (32f).

Contrary to simpler models in which it is possible to analytically find the optimal level of variables such that bus frequency, size and fare (e.g., Mohring, 1972; Jansson, 1980; Chang and Schonfeld, 1991; Tirachini *et al.*, 2010), the constrained optimisation (31)-(32) requires a numerical approach. The problem is solved using the Sequential Quadratic Programming method implemented in the constrained optimisation toolbox of Matlab v7.12.0. The solution procedure applied considers bus frequency as a continuous variable, while the number of seats, car toll and bus fare are discrete (fare and toll are constrained to be a multiple of 5 cents).

3. Application

3.1 Physical setting and input parameters

In order to illustrate the effects of explicitly accounting for the crowding discomfort in the design of buses and the pricing structure of urban transport systems, we apply the social welfare maximisation model with demand and supply data from a specific transport corridor, Military Road in North Sydney, Australia, shown in Figure 2. The section modelled comprises 3.44 km of road which is divided in 12 zones (therefore the average zone length is 286 metres). The origin-destination matrix for car and bus trips is

obtained from a traffic simulation study undertaken in this corridor by the Roads and Traffic Authority (RTA)⁸. In order to add walking trips to the matrix we use Sydney's Household Travel Survey (TDC, 2010) to obtain the city's modal split by trip distance; 66.7 percent of trips shorter than one kilometre are made on foot, a figure that drops to 24.7 percent for trips between 1 and 2 km, and 5.7 percent for trips between 2 and 5 km (considering car, bus and walk only). Then we amplify each cell (bus+car trips) by the respective percentage of walking trips according to the distance between origin and destination. The matrix obtained with this procedure is presented in Figure 3, with a total of 19,231 trips in the morning peak (7.30 to 8.30am), from which 54.3 percent are from east to west, towards the CBD (Direction 2 in Figure 1).



Figure 2: Test corridor, Military Road

⁸ This corridor is chosen because of the availability of origin-destination demand data at the level of small zones. The estimation of taste parameters for utility functions (2) to (5) is done with data collected in adjacent area in Sydney (the CBD and the North West); we assume that the estimated parameters are also applicable to the Military Road area.

O/D	1	2	3	4	5	6	7	8	9	10	11	12
1	0	856	1324	54	23	8	74	99	419	71	16	1405
2	165	0	192	15	4	1	20	19	68	14	3	326
3	829	93	0	0	0	0	0	0	0	0	0	0
4	50	12	0	0	0	0	1	3	13	1	0	91
5	146	0	0	0	0	0	0	0	1	0	0	11
6	235	9	3	0	0	0	0	3	9	0	0	17
7	87	13	4	0	0	0	0	12	48	12	0	187
8	18	1	0	0	0	0	0	0	3	9	0	8
9	396	22	5	1	1	3	24	9	0	27	3	763
10	7	0	0	0	0	0	0	0	0	0	12	1511
11	119	11	1	0	0	0	12	0	3	123	0	1027
12	1780	277	54	21	16	27	151	65	207	3763	1685	0

Figure 3: Origin-Destination matrix

The road has two lanes per direction, BPR functions (10) and (11) are assumed to represent travel times with commonly used parameter values $\alpha_0 = 0.15$ and $\alpha_1 = 4$, and a capacity $K_r = 2000 \text{ veh/h}$ obtained by assuming a 60 percent for effective green time ratio at signalised intersections. Speed at free flow is 50 km/h. With these assumptions, the average car speed is 26.3 km/h in direction 1 (outbound) and 21.5 km/h in the direction 2 (inbound), similar to the measured average speed of 22 km/h on this road (RTA, 2011, which only reports average speed in the inbound direction in the morning peak). The bus equivalency factors $\varphi(s_b)$ are 1.65 for small buses (8 m), 2.19 for standard buses (12 m), 2.60 for rigid long buses (15 m) and 3.00 for articulated buses (18 m).

Users can choose between travelling by car, bus or to walk; other alternatives like switching time period or changing origin and/or destination are not considered. The car operating cost is 14 cents/km (fuel consumption) and the average car occupancy 1.45 pax/car (TDC, 2010), which we assume remains unchanged after pricing reforms (the sensitivity of car occupancy to raising tolls is ignored). Walking speed is assumed to be 4 km/h.

The parameter estimates for the utility functions (2) to (7) is presented in Table 2. The estimation of parameters is explained and discussed in the Appendix A1, including goodness-of-fit tests and t-ratios. The constraints for the minimum and maximum number of seats per bus are explained in Section 2.4.

Table 2: Parameter values

Attribute	M1	M2	M3	M4
Access time β_a	-0.016	-0.017	-0.017	-0.017
Headway β_h	-0.009	-0.010	-0.010	-0.010
In-vehicle time bus (t_{vb}) β_{vb}	-0.019	-0.013	-0.004	-0.006
In-vehicle time car (t_{va}) β_{va}	-0.016	-0.018	-0.018	-0.018
Travel time walk β_{ve}	-0.035	-0.035	-0.035	-0.035
Cost β_c	-0.062	-0.064	-0.064	-0.064
Modal constant bus α_b	-2.080	-2.112	-2.129	-2.134
Modal constant walk α_e	-0.092	-0.099	-0.099	-0.100
$t_{vb} \times \text{den stand } \beta_{den}$		-0.004	-0.003	
$t_{vb} \times \text{prop seat } \beta_{seat}$			-0.013	
$t_{vb} \times (\text{den stand})^2 \beta_{den2}$				-0.0005
$t_{vb} \times (\text{prop seat})^2 \beta_{seat2}$				-0.013

Note: Time in minutes, cost in \$ (AUD).

3.2 Base results

Results with the current OD matrix (Figure 3) for demand models M1 to M4 are shown in Table 3. First, the solution regarding bus size, frequency, fare, toll and number of seats is similar for M1 and M2, and for M3 and M4. In the case of M1 (no crowding or standing externality internalised) it is optimal to operate with mini buses (8 m. long) at a frequency of 21.7 veh/h and to charge a fare of 10 cents, whereas in M2 (with standing disutility) the optimal solution has a slightly greater frequency of 23.7 veh/h. The similarity of results is because at this level of bus demand almost all passengers are sitting, as shown by the maximum occupancy rate (over number of seats), which is 1.08 for M1 and 0.98 for M2 (tenth row in Table 3), therefore, due to the absence of standees, both models have similar optimal outputs. A different result is obtained if we assume that the proportion of bus riders sitting is also a source of disutility, either in a linear (M3) or quadratic (M4) form; in these cases the optimal solution comprises bigger (12 m) and more frequent buses (between 25.0 and 26.1 veh/h). Importantly, the external marginal cost of crowding reflects an increase in the optimal fare of 30 cents per passenger; as in M1 and M2 the marginal cost of carrying an extra passenger is only given by the extra boarding and alighting time, whereas for M3 and M4 the optimal fare also accounts for the discomfort caused by a passenger that reduces the number of free seats on a bus. This strong increase in optimal bus fare points to the large effect of including a *crowded seating* disutility when setting optimal public transport prices.

Next, regarding the optimal number of seats, in all cases the optimal result is having the maximum number of seats that is technically possible (24 for 8 m.-long buses, 39 for 12 m.-long buses), constrained by the minimum area required free of seats⁹. Out of the

⁹ Note that bus utility in M1 is indifferent to the number of seats inside buses, therefore as long as the capacity constraint is not binding, any number of seats would produce the same level of social welfare. In

available area for seating and standing, 80 percent is allocated to seating and 20 percent to standing. The greater frequency and bus size of models M3 and M4 considerably reduces the average occupancy rate (as a function of the number of seats) from over 50 percent in M1 and M2, to 30 percent in M3 and M4 (the supply of seats per hour is almost doubled from 521 in M1 to 1,017 in M3). The key role of the number of seats as a decision variable is unveiled in Section 3.3, when an increase in road congestion challenges the optimality of having the maximum number of seats if the planner ignores the external cost of crowding.

Table 3: Base case results

Optimal value	M1	M2	M3	M4
Bus length [m]	8	8	12	12
Frequency [veh/h]	21.7	23.7	26.1	25.0
Fare [\$]	0.1	0.1	0.4	0.4
Toll [\$]	2.0	2.0	2.0	2.0
Number of seats	24	24	39	39
Bus capacity [pax/bus]	36	36	58	58
Seating area/total bus area	0.58	0.58	0.63	0.63
Seating area/ (seating plus standing area)	0.80	0.80	0.80	0.80
Average occupancy rate (over number of seats)	0.57	0.52	0.30	0.31
Max. occupancy rate (over number of seats)	1.08	0.98	0.56	0.58
Max. occupancy rate (over total capacity)	0.62	0.56	0.32	0.33
Seat capacity bus route (seats/h)	521	569	1,017	975
Total capacity bus route (pax/h)	782	854	1,512	1,450
Fare collection technology	Off-board	Mag. strip	Mag. strip	Mag. strip
Boarding regime	All doors	All doors	All doors	All doors
Social welfare [\$]	129,544	122,984	122,897	122,801
Consumer surplus [\$]	114,290	107,721	107,454	107,319
Bus operator profit [\$]	-671	-645	-467	-436
Toll revenue [\$]	15,925	15,908	15,909	15,917
Subsidy/bus operator cost	0.83	0.83	0.46	0.44
Fleet size [buses]	11	12	13	13
Modal split bus	7.1%	7.0%	7.0%	7.1%
Modal split car	60.0%	59.9%	60.0%	60.0%
Modal split walk	32.9%	33.1%	33.0%	32.9%

Table 3 the capacity constraint is inactive for buses with the maximum number of seats n_{seat}^{max} , therefore n_{seat}^{max} is arbitrarily chosen for M1.

The outputs regarding number of doors, bus boarding policy and fare collection technique are described as follows. First, in all cases it is optimal to have the maximum number of doors given by the bus size, i.e., 2 doors for 8 metre buses and 3 doors for 12 metre buses, as the more doors are in place the shorter are the boarding and alighting times¹⁰. Second, sequential boarding and alighting at all doors (*TnBn* system) is more efficient than operating with boarding at the front door only (*TnB1*). Third, the optimal fare collection technology is off-board with M1 and on-board with a magnetic strip with M2 to M4.

The consideration of a *crowded seating* disutility has a strong effect on the financial state of the public transport provider and the subsidy required to run the system: in M1 and M2 with an optimal fare of 10 cents it is required a subsidy that needs to cover 83 percent of the operator cost, whereas if the optimal fare of M3 is charged (40 cents) the required subsidy is halved¹¹. In all cases the toll revenue is more than enough to cover the bus operator deficit (ignoring toll collection costs).

Since the model includes disaggregated origin-destination (OD) demand information, we are able to compute marginal external costs of crowding and congestion per OD pair. The marginal external costs are computed by deriving travel time costs with respect to modal demand in order to account for the effect of an extra traveller resulting in increasing travel time on the road (i.e., the congestion externality captured by BPR functions 10 and 11) and in increasing the proportion of seats occupied or the density of standees (crowding externality captured in utility functions, equations 3, 4 and 5). In all cases, external time costs [min/trip] are divided by the marginal utility of cost (β_c) to have the correct [\$/trip] measure. For illustration, marginal external cost curves for trips from five origins are depicted in Figure 4 for model M3, in which, for example, “crowding 1-12” shows the marginal external crowding cost (MECrC, between 10 and 63 cents of dollar) of bus trips starting in origin 1 with destination in all zones between 2 and 12, whereas “crowding 12-1” shows the MECrC for trips with origin in zone 12 and destinations between 1 and 11 (between 16 and 74 cents in the dollar). The two MECrC curves grow approximately linearly as a trip gets longer, and the value is slightly larger in Direction 2 (from 12 to 1) given the larger demand in this direction (54.3 percent of total). Using the distance in kilometres between stops as explanatory variable, the following relationship between MECrC and trip length L_{km} can be estimated:

$$MECrC_{1-12} = 0.02 + 0.16L_{km} \quad (R^2 = 0.98) \quad (33a)$$

$$MECrC_{12-1} = 0.07 + 0.18L_{km} \quad (R^2 = 0.96) \quad (33b)$$

Equations 33(a) and (b) imply that for trips starting in sections 1 and 12, MECrC grows by between 16 and 18 cents of dollar for every kilometre that a person is inside a bus. On the other hand, the marginal external congestion cost (MECoC) is larger than MECrC

¹⁰ This result ignores that the time to open and close doors may increase with the number of doors, because drivers may spend more time to check that all doors are clear of passengers if more doors are provided in a bus.

¹¹ The current operation of Sydney buses has a minimum fare of \$2.10 for a single ticket, which in our model would produce profits, however the current system has to be subsidised. This divergence is explained by a number of elements, including the likely existence of a large amount of fixed costs that is not considered in this application, and that we are only modelling the morning peak period (in which demand is the highest) in an area of high demand density.

as shown in Figure 4. The relatively larger level of the curve “Congestion 12-1” is derived from the large travel demand from section 12 to sections 11 and 10, whilst in comparison, a person starting his or her trip in section 11 (curve “crowding 11-1”) has a much lower MECoC.

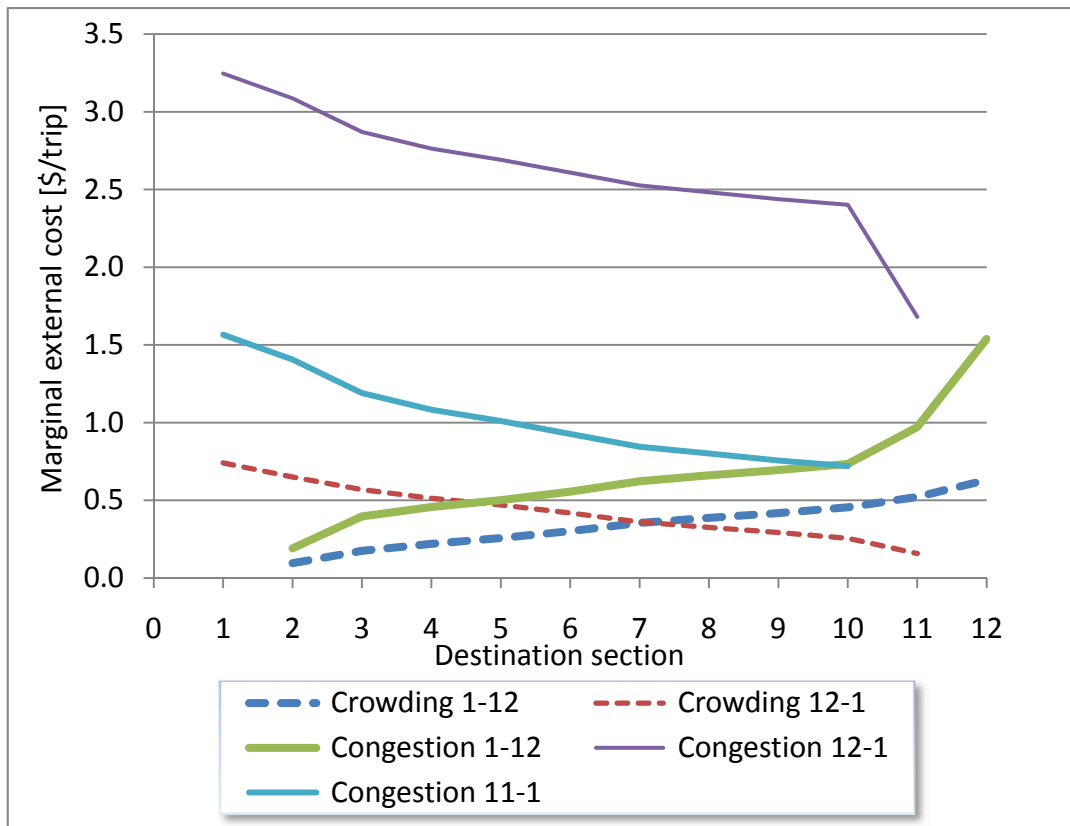


Figure 4: Marginal external crowding and congestion cost

In the next sections, we analyse how the bus service and pricing levels (fare and toll) should be adapted when faced with an increase in transport demand (e.g., through a future urban densification around the corridor), assuming that it is not possible to increase road capacity ($K_r = 2000 \text{ veh/h}$). The idea is to analyse the evolution of key design variables when the system is stressed and severe congestion arises. The trips by origin and destination of Figure 3 are uniformly scaled in five steps, up to a total demand of 28,850 trips/h (50 percent higher than the current number of trips). The main results for modal split, bus service design, pricing, crowding and congestion are discussed. Model M4 is not shown because it produces similar results to those of M3.

3.3 Optimal bus frequency: The trade-off between congestion and crowding

The evolution of the optimal bus frequency is presented in Figure 5. It is evident that regardless of the demand model considered, frequency does not vary monotonically with demand, in particular optimal frequency can decrease as demand grows, although the reasons for this result are not the same across the models. Focusing on M2 (standing disutility) first, we observe that frequency is increasing up to 32 veh/h for 1,900 pax/h, but drops to 22 veh/h for 2,100 pax/h; this is because up at 1,900 pax/h the optimal bus is mini (8 metres) whereas at 2,100 pax/h it becomes optimal to operate with standard 12 metre buses with a higher capacity. Similarly for M3, a discrete increase in bus length (from 12 to 15 metres) also explains the drop in frequency from 26 to 24 veh/h with

1,550 pax/h. However, if bus size remains unaltered, frequency is always an increasing function of demand if we assume that crowding and standing disutilities matter (case of M2 up to 1,900 pax/h and M3 beyond 1,550 pax/h) which is in line with all total cost minimisation models that optimise bus frequency either assuming a fixed bus size, or that bus size can be freely adjusted to meet demand once frequency has been optimised (e.g., Mohring, 1972; Jansson, 1980; Jara-Díaz and Gschwender, 2003).

What happens with the optimal frequency in the model that is insensitive to crowding as a source of disutility (M1) is even more noteworthy. In this case the optimal bus size does not change across the whole demand range (mini buses) and in spite of that, frequency slightly decreases from 21.7 to 20.8 veh/h as demand increases from 1,370 to 2,250 pax/h¹². This is because of congestion on the road: as total demand grows so does the number of people that use the congestible road facility (the actual speed drop for cars and buses is shown in Figure 6 for both directions, D1 and D2), and given that bus frequency adds to traffic congestion, the model tries to reduce the number of buses on the street at the expense of increasing crowding levels inside buses, which in M1 is welfare improving because crowding comes at no comfort loss. This is a clear sign of the relevance of including crowding in the optimisation of transport corridors that have cross-congestion between buses and cars.

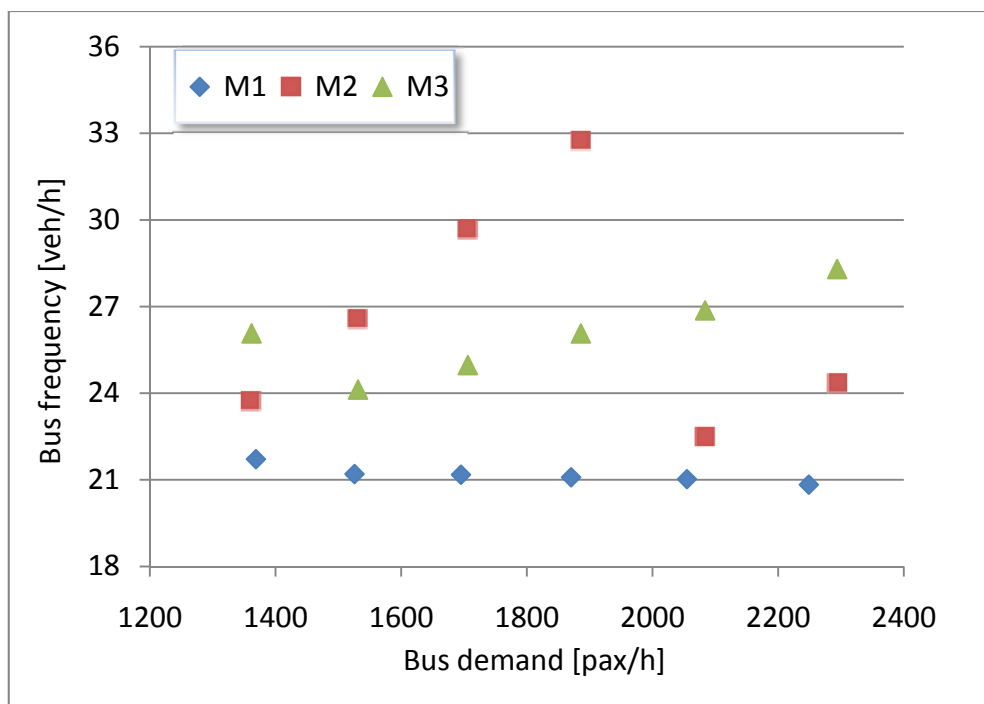


Figure 5: Optimal frequency

¹² This frequency reduction is not necessarily in opposition of traditional bus optimisation models that predict bus frequency to increase with demand, such as Mohring (1972). In Mohring's "square root rule", frequency decreases with the cycle time, which in this case is increasing with demand because of road congestion.

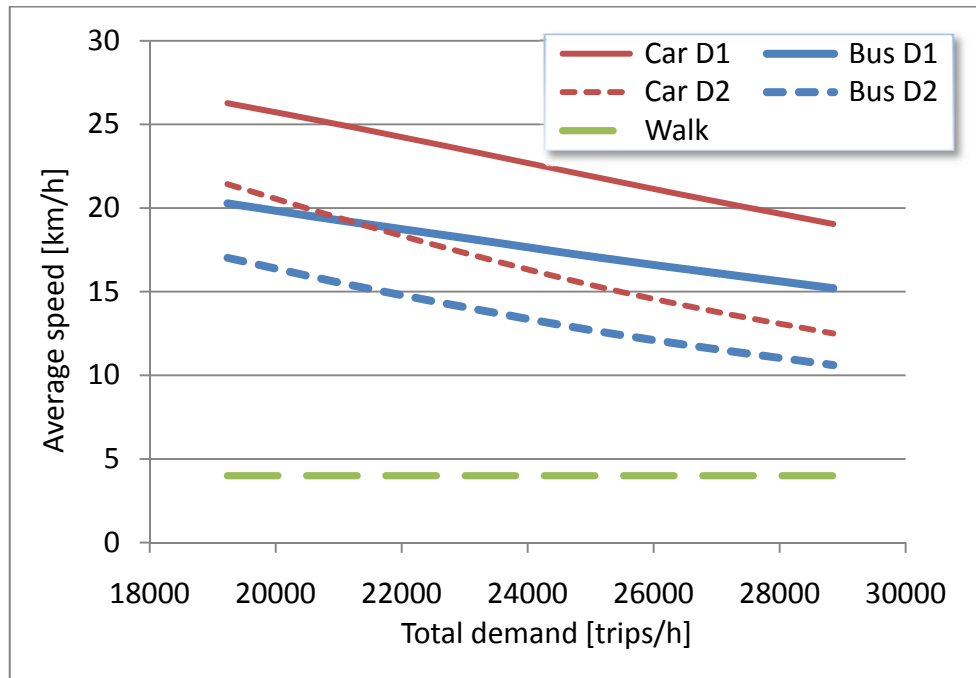


Figure 6: Average speed M1

The examination of optimal frequencies does not provide a full picture of the transport supply being provided by the bus operator because different optimal bus sizes are chosen in Figure 5. The total seat supply (frequency times number of seats per bus) and seat plus stand supply (frequency times bus capacity) are shown in Figures 7 and 8. It is clear that the optimal capacity that a planner would choose is quite sensitive to the characterisation of the crowding and standing disutilities. M1 is insensitive to the number of seats chosen as long as the bus capacity constraint (32a) is not binding; therefore we have shown the maximum number of seats per bus such that (32a) is not active, which passes from 24 seats per bus when demand is 1,370 pax/h, to 11 seats per bus when demand is 2,250 pax/h, as reflected in Figure 7 with a total seat supply decreasing for M1. In other words, when confronted with an increase in demand, part of the (optimal) increase in supply is provided simply by reducing the number of seats in order to increase the number of passengers that can be accommodated in a bus, at no crowding cost in M1¹³. A completely different outcome is obtained if crowding and standing matter, in which case the number of seats is kept at the maximum possible given constraint (32b) and total seat capacity is increasing for the whole demand range on M2 and M3, as shown in Figure 7.

The fact that planners or bus operators would choose to reduce the number of seats per bus if crowding and standing disutilities are not explicitly accounted for (M1) does not mean that the total transport capacity (seat plus stand) is decreasing; as Figure 8 shows that with M1, total capacity is actually increasing, due to the increase in bus capacity coupled with a slightly decreasing (almost flat) bus frequency (Figure 5).

¹³ Note that a model in which the number of seats cannot be adjusted would force the frequency and/or bus size to increase if the capacity constraint is binding, which comes at a cost for the operator.

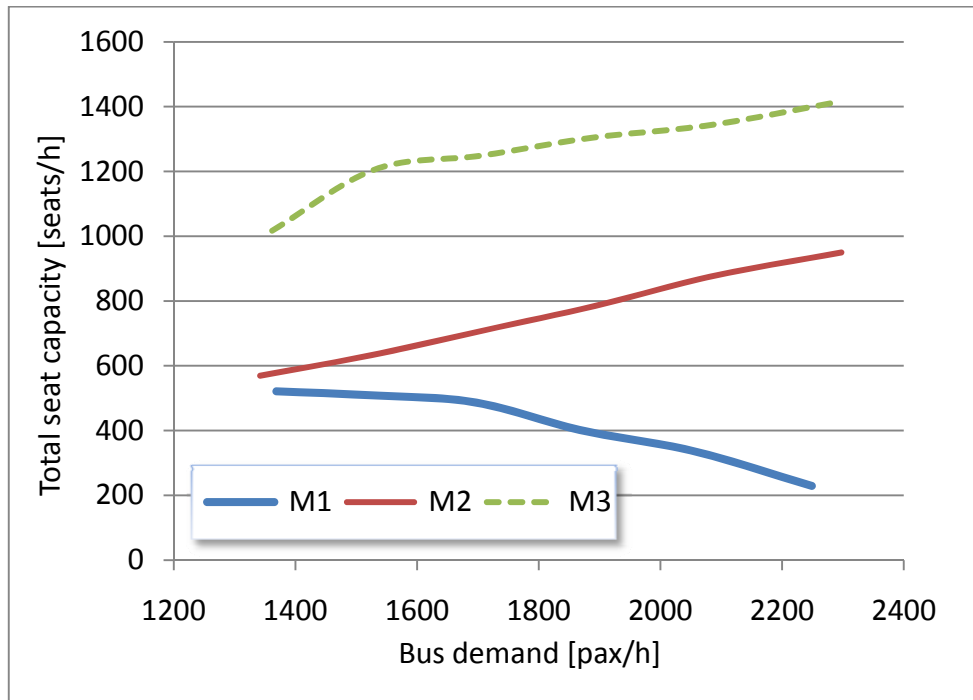


Figure 7: Optimal seat supply

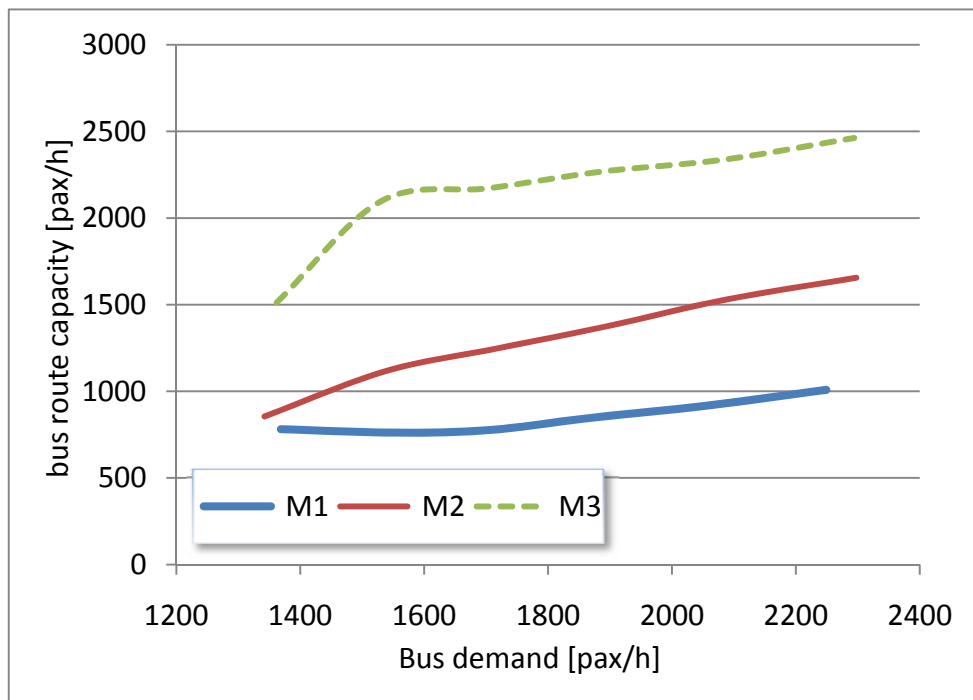


Figure 8: Optimal seating+standing supply

In order to show that M1's optimal frequency drop (Figure 5) is actually due to the congestion interaction with cars, an alternative scenario in which there is an exclusive bus lane is modelled, and therefore, bus frequency does not influence car travel time, which is only given by traffic flow (cars remain in two lanes). As depicted in Figure 9, when buses do not affect cars, optimal frequency has an increasing tendency along the demand range.

Finally, it is worth mentioning that in all scenarios, bus frequencies are low enough not to cause any queuing delay at bus stops (equation 14), which are assumed to have two berths each.

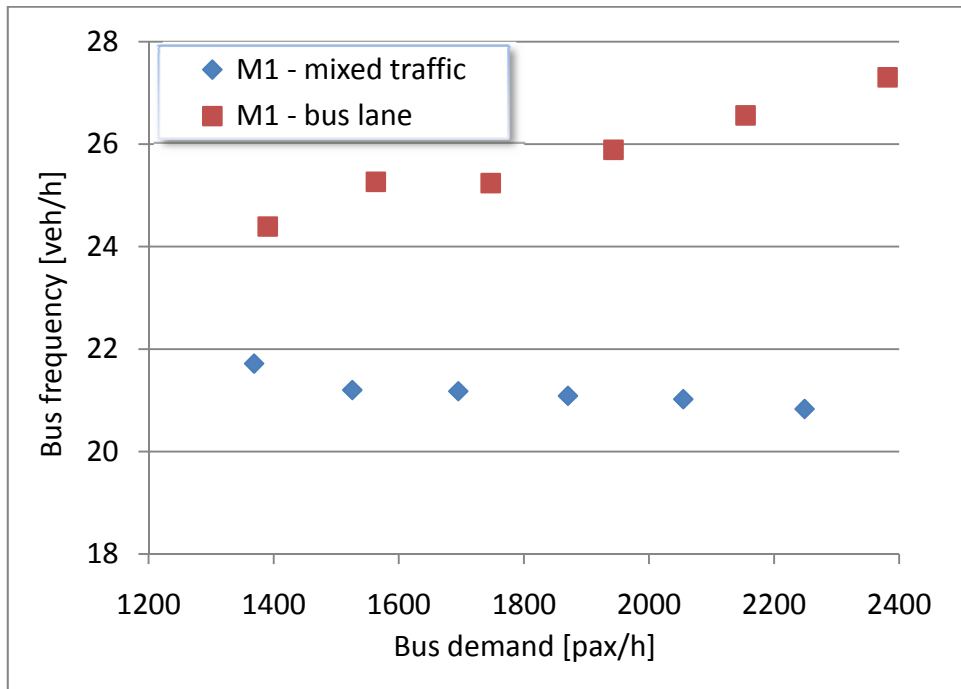


Figure 9: Optimal bus frequency on shared and dedicated right-of-way

3.4 Optimal pricing and modal split

The change in optimal toll and bus fare is analysed in this section. The optimal toll is largely insensitive to the specification of crowding in the bus utility functions, therefore only one value is presented in Figure 10, which shows the increase in the optimal toll as total demand (and consequently road congestion) grows. On the other hand, the optimal bus fare slightly increases for M2 and M3, up to 50 cents per ticket, whereas in M1 the fare is maintained at 10 cents.

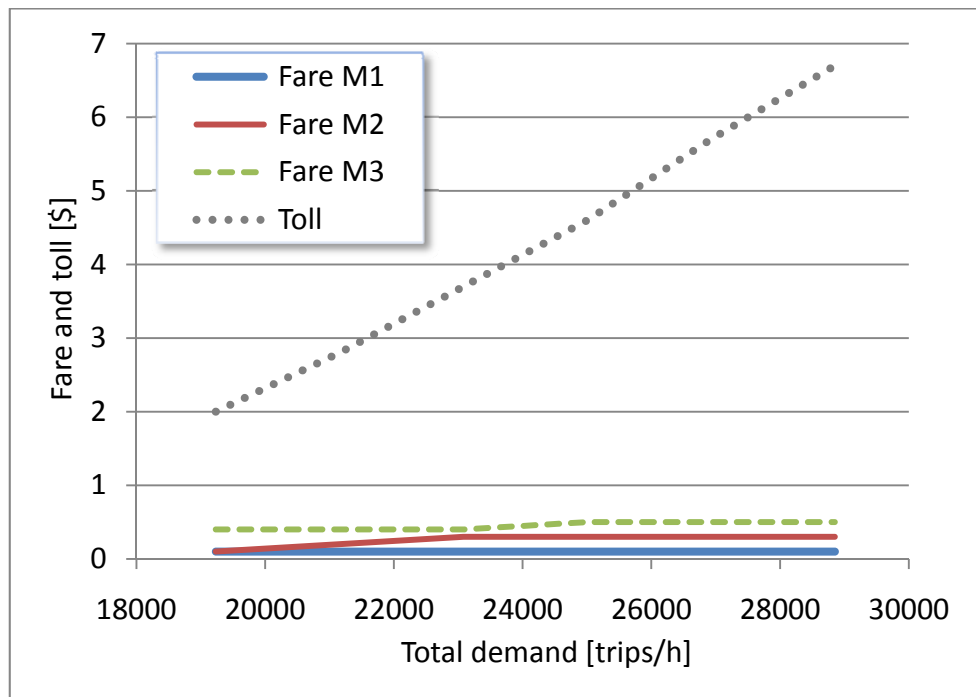


Figure 10: Optimal toll and bus fare

Next, modal shares are analysed. The predicted modal splits are almost identical under the four models, and that. The resulting modal split with optimised bus design and pricing structure is almost identical under all demand models (Table 3), thus only M2 is shown for illustration in Figure 11. Compared to the observed modal split (62.5 percent car, 31.6 percent walk and 5.9 percent), more people decide to walk (33 percent) and ride a bus (7 percent), reducing the car modal split to 60 percent. The relatively modest increase in bus modal share (from 5.9 to 7 percent) after optimisation is due to the large and negative value of the bus modal constant (around -2.1 in Table 2), which has been calibrated to predict observed modal shares in the base situation (very unbalanced in favour of the car mode).

The worsening of road congestion¹⁴ as total demand grows encourages walking. This results in a car modal share dropping from 60 to 54 percent, the (assumed uncongestable) alternative of walking increases from 33 to 38 percent, and the bus choice grows from 7 to 8 percent of all trips (due to increased frequency and price difference between toll and fare, see Figure 10). Therefore, if transport demand grows in the future and road capacity is held constant, the model predicts walking to become more relevant as a travel alternative, which in this example is supported by the fact that trips are relatively short (the corridor is 3.4 km long). In fact, Figure 12 displays the modal split per trip length for trips starting in Zone 1; it is clear that there is a loss of competitiveness of walking as an alternative to motorised modes as trip length increases.

¹⁴ Shown in Figure 5 for M1; the result for M2 is similar.

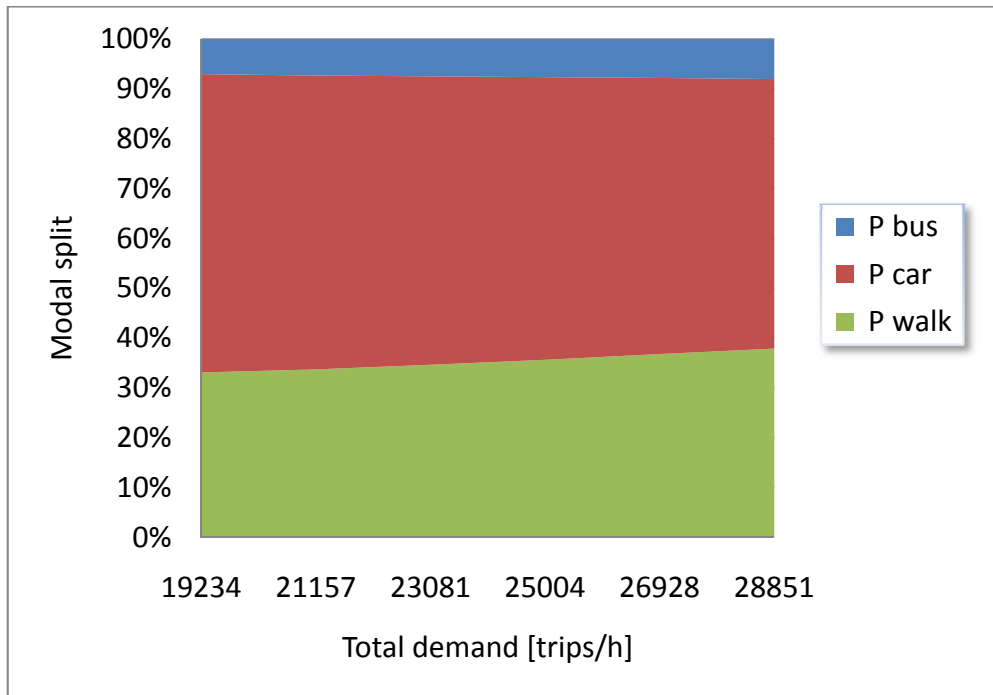


Figure 11: Modal split M2

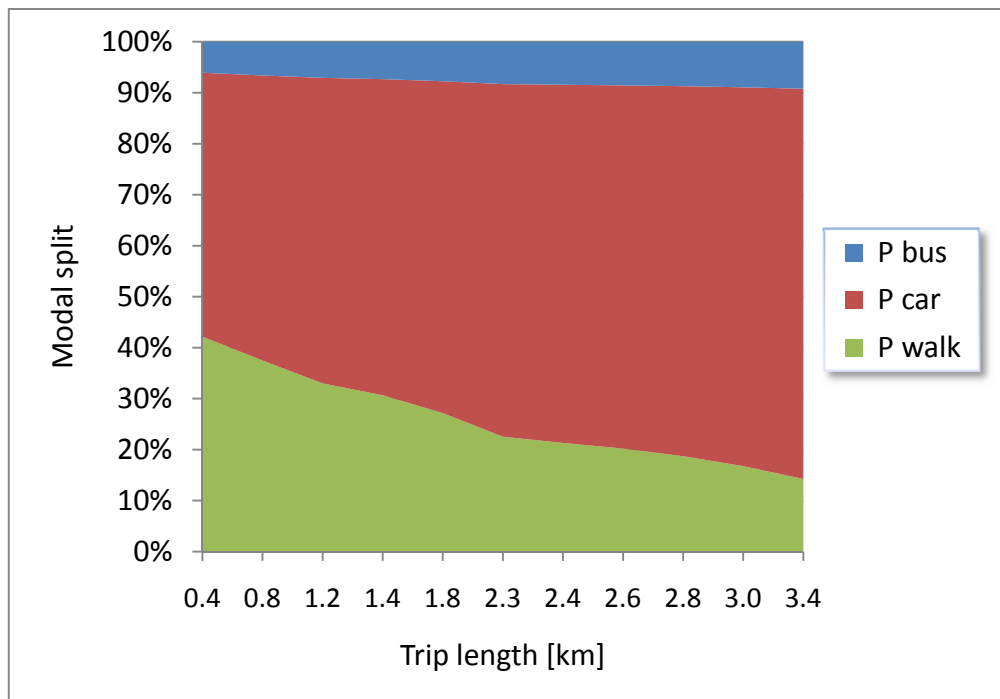


Figure 12: Modal split per trip length, M2 base case (total demand=19,234)

3.5 The case with increased bus-induced congestion

In the previous scenario it was assumed that passenger car equivalency factor for buses $\varphi(s_b)$ is solely given by bus size, from 1.65 for mini buses (8 metres) to 3 for articulated buses (18 metres). However, some authors such as Parry and Small (2009) assume that, in mixed traffic, buses should be given a greater weight in the congestion functions (10) and (11), given that their stops to load and unload passengers have an effect on the capacity of lanes and impose delays on other modes including cars (Koshy and Arasan, 25

2005; Zhao *et al.*, 2007). We find that when doubling the passenger car equivalency factor (to between 3.3 and 6) optimal bus frequency is reduced, and that the impact is stronger if no crowding externalities are explicitly modelled (M1, Figure 13) than when the crowding disutility is accounted for, in which case the crowding externality dominates over the congestion externality (M2, Figure 14). In Figure 13, the increase in frequency for a bus demand beyond 1,800 pax/h is because the minimum number of seats n_{seat}^{min} has been reached, the capacity constraint (26) is binding and therefore the operator has no option but increasing the bus frequency to meet demand.

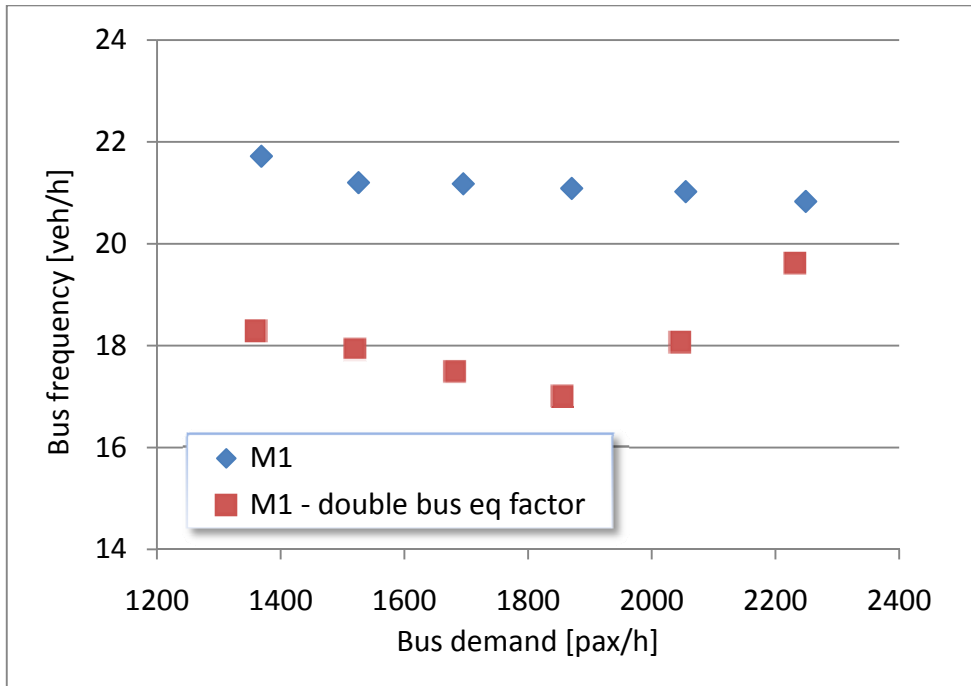


Figure 13: Optimal bus frequency M1, double equivalency factor for buses

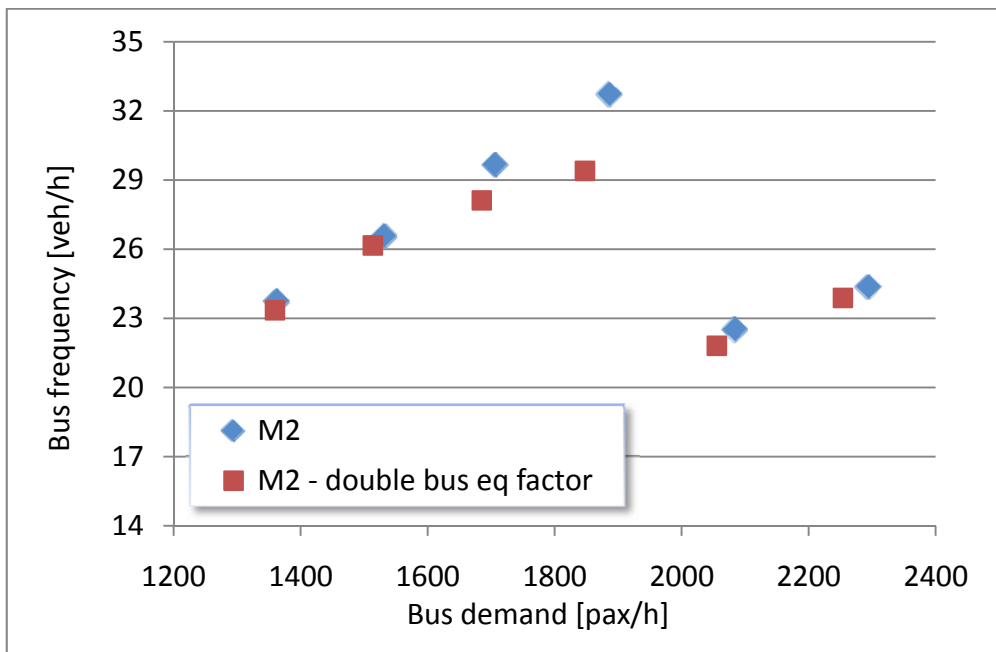


Figure 14: Optimal bus frequency M2, double equivalency factor for buses

3.6 The relationship between the number of seats and optimal frequency

In this section, we study the sensitivity of the optimal bus frequency to alternative bus layouts regarding number of seats. As previously discussed, in all scenarios in which the crowding externality is considered (M2, M3 and M4), the optimal bus design comprises having as many seats as possible, given an optimal bus size, in order to reduce the crowding effects of seating and reduce the number of standees. In this context, we study what happens if the number of seats is exogenously chosen to be lower than the maximum (and therefore the bus capacity is increased); Figure 15 shows that for both M2 (mini buses, 8 m.) and M3 (standard buses, 12 m.) frequency should be increased as a response to the users' discomfort of having fewer seats. In other words, the number of seats inside a bus does have an effect on the optimal design of a public transport system if the planner acknowledges that users dislike crowding.

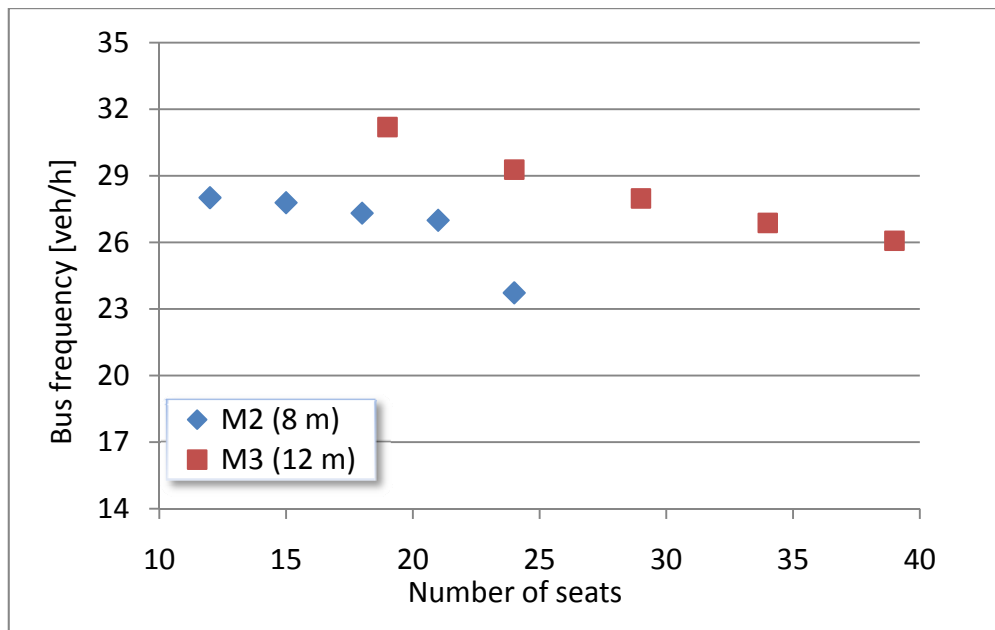


Figure 15: Optimal bus frequency for suboptimal numbers of seat

3.7 The second best scenario

The preceding analysis was undertaken by assuming that a congestion toll on cars is in place, as shown in Figure 10. In this section, the second best case in which there is no car toll is investigated¹⁵. We limit the analysis to a graphical comparison of relevant optimisation outputs between the first best and second best scenarios.

The second best bus fare is negative across the demand range tested and for all utility specifications (M1 to M4), i.e., the optimal decrease in bus fare to face a zero toll policy is larger than the optimal first best bus fare (between 10 and 40 cents)¹⁶. More generally, this result points to the potential optimality of very low public transport fares and large subsidies if car use is substantially under-priced (common in peak periods), in line with

¹⁵ The principles behind first best and second best pricing are extensively analysed in Small and Verhoef (2007) and Tirachini and Hensher (2012) among others.

¹⁶ A negative second best bus fare is also obtained by Ahn (2009).

earlier findings (Glaister and Lewis, 1978; Proost and Van Dender, 2008; Ahn, 2009; Parry and Small, 2009)¹⁷.

Figure 16 shows the difference between optimal toll and fare in the first best and second best scenarios for demand models M1, M2 and M3 (therefore, in the second best scenarios the curves are equal to the absolute value of the negative bus fare). The difference between toll and fare is lower in the second best scenario, as also found by Ahn (2009). In Ahn (2009), the second best bus fare does not decrease sufficiently to maintain the difference between fare and toll in the first best scenario because such a low bus fare would produce a greater than socially optimal amount of total trips; whereas in our framework the amount of total trips is fixed but the amount of motorised trips is not, and hence a low (negative in this case) bus fare attracts not only car users but also walkers to public transport (Kerin, 1992; Tirachini and Hensher, 2012). This explains that the second best bus fare is not so low as to maintain the first best toll-minus-fare difference.

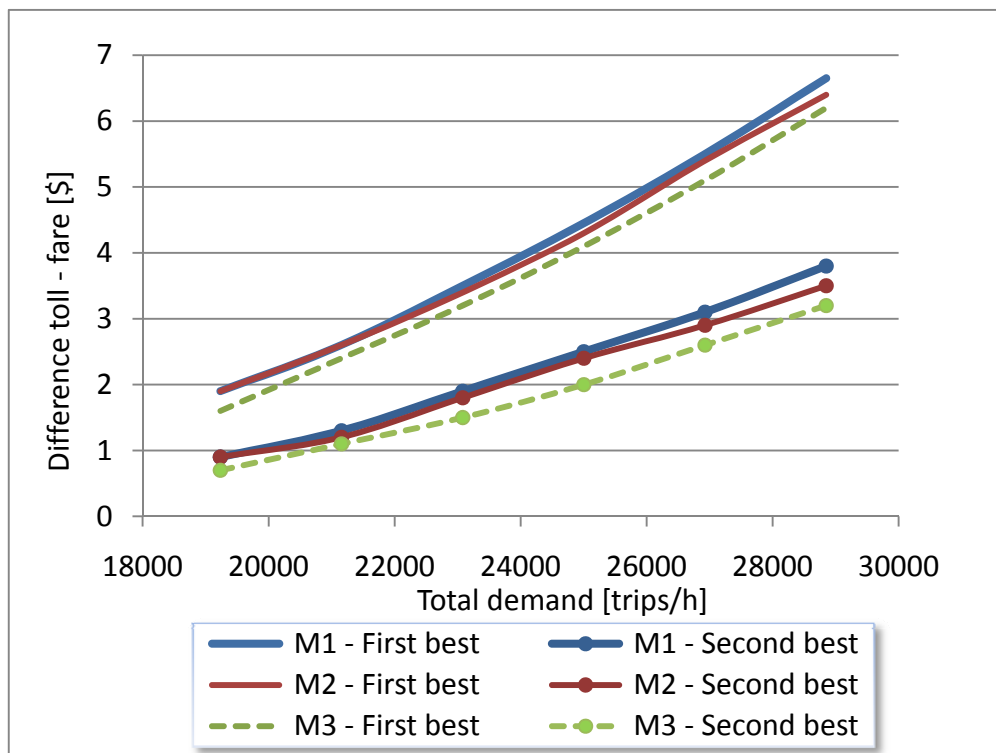


Figure 16: Optimal toll minus bus fare, first best and second best scenarios

Finally, optimal bus frequency is lower in the second best scenario as shown in Figure 17 (model M2 is eliminated for easiness of exposition), because the underpricing of car traffic generates a greater than optimal amount of car trips, and therefore increased congestion for both cars and buses. In addition to the rather unrealistic scenario with negative bus fares, in Figure 17 we have added a third scenario in which a budget constraint is applied that eliminates the subsidy required for the bus route in both the first and second best scenarios. In this “zero profit scenario”, bus fares escalate to \$0.50-0.60 for M1 and to \$0.70-0.75 for M3; optimal frequency reduces and falls within the

¹⁷ Basso and Silva (2014) show that the existence and level of optimal bus subsidies might be strongly influenced by the provision of dedicated bus lanes.

range 16-19 bus/h for M1 and 21-24 bus/h for M3. Both variables imply a considerable loss of bus passengers, as shown in Figure 17. This reduction in demand actually means that crowding levels in the zero-profit scenario do not significantly change with respect to the scenarios without budget constraints (for example, when total demand is 28,850 trips/h, average occupancy rate of buses is 36.9 % in the scenario “M3 – Second best”, versus 38.1 % in the scenario “M3 – second best – zero profit”).

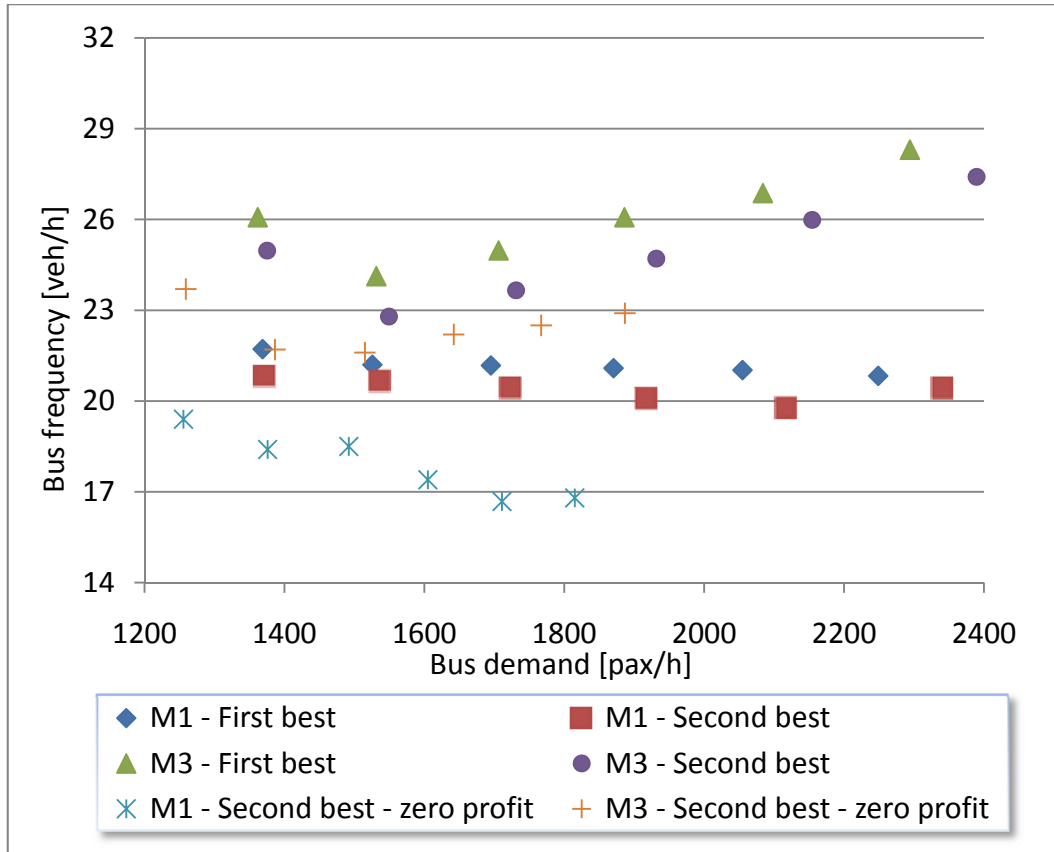


Figure 17: Optimal bus frequency, first best and second best scenarios

4. Conclusions

In this paper we have introduced a social welfare maximisation model with disaggregated origin destination demand and multiple travel alternatives, with the aim of optimising the design of urban bus routes including pricing decisions for both bus and car. The influence of bus crowding is highlighted as we analyse its impact on both the design of the bus service and the congestion level on the road. The consideration of crowding externalities as increasing the discomfort of public transport users pushes towards having bigger and more frequent buses (Jara-Díaz and Gschwender, 2003), which in turn may worsen both bus and traffic congestion on shared roads. The number of seats in buses is introduced as a decision variable for the first time in a microeconomic public transport model; the number of seats is the result of the trade-off between passengers’ comfort (that drives the number of seats up) and vehicle capacity (which might be increased by removing seats). The model is applied to the Military Road corridor in North Sydney and results are discussed in several scenarios with different demand levels and modelling assumptions.

A number of results stand out. The consideration of crowding externalities (at both seating and standing) imposes a higher optimal bus fare, and consequently, a reduction

of the optimal bus subsidy. Optimal bus frequency results from a trade-off between the level of *congestion inside buses*, i.e., passengers' crowding, and the level of *congestion outside buses*, i.e., the effect of frequency on slowing down both buses and cars in mixed-traffic. In particular, optimal bus frequency is quite sensitive to the assumptions regarding crowding costs, the impact of buses on traffic congestion and the overall congestion level, as the crowding externality puts pressure on operators to provide more frequent services, which in turn may add to both bus and car congestion. We find that if crowding matters, bus frequency should increase (for a given bus size) with demand even under heavy congestion, however that might not be the case if the crowding externality is not accounted for, in which case an increase of total demand might be met by a decrease of both frequency and number of seats per bus, at the expense of crowding passengers inside buses and making more passengers stand while travelling.

We find that the existence of a crowding externality implies that buses should have as many seats as technically possible, up to a minimum area that must be left free of seats. If for any other reason planners decide to have buses with fewer seats than optimal (e.g., to increase bus capacity), bus frequency (and the number of buses itself) should be increased to compensate for discomfort imposed on public transport users. Future research should test the optimality of providing the maximum number of seats in the following two cases: (i) in routes with high public transport demand, in which an active capacity constraint may push the number of seats down, and (ii) if the standing disutility is low relative to sitting, in which case, in principle, it might be optimal to provide less seats than the maximum feasible, if the benefit of an increased capacity for the operator is larger than the comfort disbenefit for users.

Regarding the relevance of non-motorised modes in urban mobility, in a corridor of 3.4 km, an increase in total transport demand worsens traffic congestion which increases the choice of walking relative to its motorised alternatives (with optimised bus service, fare and toll). This suggests that at least for short trips, improving the travel conditions of non-motorised modes is a wise strategy to tackle worsening congestion problems in cities.

Including bus riders sitting as a source of disutility produces important changes in some variables (optimal bus fares, bus size, frequency, subsidy). However, optimising the transport corridors under different assumptions on the relevance of crowding externalities to users produced similar values of other outcomes such as social welfare, consumer surplus and modal splits across models M1 to M4. More research is needed to assess if this finding holds in other contexts, for example, with a superior bus modal share (over 20 percent).

The analytical framework presented in the paper can be extended in several ways. Departure time decisions can be included by introducing a multi-period framework with modal and time-of-day substitution (e.g., Glaister and Lewis, 1978; De Borger and Wouters, 1998; Proost and Van Dender, 2008). Extensions to larger corridors or networks would make more apparent the need to introduce distance-based tolling and more complex public transport fare structures (by section or origin destination pair), extensions that can be added in the present framework by setting fares per kilometre. Note that a differentiation in price by distance is more attractive if destination choice is included as well; otherwise the model would overestimate the revenue from long-

distance travellers, some of whom would try to travel shorter distances. Our model also ignores issues of income distribution (Dodgson and Topham, 1987; Mayeres and Proost, 1997), heterogeneity in the value of time savings (Verhoef and Small, 2004) and that road capacity may also be endogenous (De Borger and Wouters, 1998; Arnott and Yan, 2000), all dimensions that are promising venues of further research for the analysis of marginal cost pricing and optimal public transport design including congestion and crowding externalities. Finally, regarding the number of seats as a variable, the bus operator cost structure could include differences in capital and maintenance costs due to alternative standing and seating configurations.

Acknowledgements

Part of this research was developed while the first author was associated with ITLS-Sydney. This study has been supported by the Australian Research Council Discovery Program Grant DP120100201 titled: 'Valuation of Service Reliability and Crowding under Risk and Uncertainty: Neglected Drivers of Demand for Public Transport' and by the Complex Engineering Systems Institute, Chile (Grants ICM P-05-004-F, CONICYT FBO16). This paper is also a contribution to the research program of the Volvo Research and Education Foundation Bus Rapid Transit Centre of Excellence. The authors are grateful to Stef Proost, Bruno de Borger, two anonymous referees and participants of the TML-KUL seminar (Leuven, May 2012), the Kuhmo-Nectar Conference on Transportation Economics (Berlin, June 2012) and the Conference on Advanced Systems for Public Transport – CASPT12 (Santiago de Chile, July 2012) for helpful comments. Remaining errors are the authors' responsibility alone.

References

- Ahn, K., 2009. Road pricing and bus service policies. *Journal of Transport Economics and Policy* 43(1), 25-53.
- Arnott, R., Yan, A., 2000. The two-mode problem: second-best pricing and capacity. *Review of Urban & Regional Development Studies* 12(3), 170-199.
- ATC, 2006. National Guidelines for Transport System Management in Australia – Volume 4: Urban Transport. Australian Transport Council.
- Basso, L. J., Silva, H. E., 2014. Efficiency and substitutability of transit subsidies and other urban transport policies. *American Economic Journal: Economic Policy*. Forthcoming.
- Basso, L. J., Guevara, C. A., Gschwender, A., Fuster, M., 2011. Congestion pricing, transit subsidies and dedicated bus lanes: Efficient and practical solutions to congestion. *Transport Policy* 18(5), 676-684.
- Chang, S. K., Schonfeld, P. M., 1991. Multiple period optimization of bus transit systems. *Transportation Research Part B* 25(6), 453-478.

- De Borger, B., Wouters, S., 1998. Transport externalities and optimal pricing and supply decisions in urban transportation: a simulation analysis for Belgium. *Regional Science and Urban Economics* 28(2), 163-197.
- De Borger, B., Mayeres, I., Proost, S., Wouters, S., 1996. Optimal pricing of urban passenger transport: a simulation exercise for Belgium. *Journal of Transport Economics and Policy* 30(1), 31-54.
- Dodgson, J. S., Topham, N., 1987. Benefit-cost rules for urban transit subsidies: an integration of allocational, distributional and public finance issues. *Journal of Transport Economics and Policy* 21(1), 57-71.
- Fernández, R., Planzer, R., 2002. On the capacity of bus transit systems. *Transport Reviews* 22(3), 267–293.
- Fernández, R., Valenzuela, E., Gálvez, T., 2000. Incorporación de la capacidad y rendimiento de paraderos en el programa TRANSYT ("Incorporation of capacity and performance of bus stops in the program TRANSYT", in Spanish). Proceedings of the XI Panamerican Conference of Transport and Traffic Engineering, Gramado, Brasil.
- Fernández, R., del Campo, M., Swett, C., 2008. Data collection and calibration of passenger service time models for the Transantiago system. European Transport Conference, The Netherlands, 6-8 October.
- FTA, 2009. Characteristics of Bus Rapid Transit for Decision-Making. Project No. FTA-FL-26-7109.2009.1, Federal Transit Administration, U.S. Department of Transportation.
- Gibson, J., Baeza, I., Willumsen, L. G., 1989. Bus-stops, congestion and congested bus-stops. *Traffic Engineering and Control* 30(6), 291-302.
- Glaister, S., Lewis, D., 1978. An integrated fares policy for transport in London. *Journal of Public Economics* 9(3), 341-355.
- Gonzales, E. J., Daganzo, C. F., 2012. Morning commute with competing modes and distributed demand: User equilibrium, system optimum, and pricing. *Transportation Research Part B* 46(10), 1519-1534.
- Gonzales, E. J., Daganzo, C. F., 2013. The evening commute with cars and transit: Duality results and user equilibrium for the combined morning and evening peaks. *Transportation Research Part B* 57, 286-299.
- Hensher, D. A., 2010. Knowing your costs. Notes for the Certificate of Transport Management, Institute of Transport and Logistics Studies, The University of Sydney.
- Hensher, D. A., Rose, J. M., Collins, A., 2011. Identifying commuter preferences for existing modes and a proposed Metro in Sydney, Australia with special reference to crowding. *Public Transport* 3(2), 109-147.
- Jansson, J. O., 1980. A simple bus line model for optimization of service frequency and bus size. *Journal of Transport Economics and Policy* 14(1), 53-80.

- Jansson, K., 2010. Public transport policy with and without road pricing 5th Kuhmo-Nectar Conference on Transport Economics, Valencia, Spain, July 8-9.
- Jara-Díaz, S. R., 2007. *Transport Economic Theory*, Elsevier.
- Jara-Díaz, S. R., Videla, J., 1989. Detection of income effect in mode choice: Theory and application. *Transportation Research Part B* 23(6), 393-400.
- Jara-Díaz, S. R., Gschwender, A., 2003. Towards a general microeconomic model for the operation of public transport. *Transport Reviews* 23(4), 453 - 469.
- Jovicic, G., Hansen, C. O., 2003. A passenger travel demand model for Copenhagen. *Transportation Research Part A* 37(4), 333-349.
- Kerin, P. D., 1992. Efficient bus fares. *Transport Reviews* 12(1), 33 - 47.
- Koshy, R. Z., Arasan, V. T., 2005. Influence of bus stops on flow characteristics of mixed traffic. *Journal of Transportation Engineering* 131(8), 640-643.
- Kraus, M., 1991. Discomfort externalities and marginal cost transit fares. *Journal of Urban Economics* 29(2), 249-259.
- Li, Z.-C., Lam, W. H. K., Wong, S. C., 2012. Modeling intermodal equilibrium for bimodal transportation system design problems in a linear monocentric city. *Transportation Research Part B: Methodological* 46(1), 30-49.
- Maunsell, F., Macdonald, M., 2007. Rail overcrowding, reliability and frequency. Report for Centro, the West Midlands Passenger Transport Executive.
- Mayeres, I., Proost, S., 1997. Optimal Tax and Public Investment Rules for Congestion Type of Externalities. *The Scandinavian Journal of Economics* 99(2), 261-279.
- Mohring, H., 1972. Optimization and scale economies in urban bus transportation. *American Economic Review* 62(4), 591-604.
- Oldfield, R. H., Bly, P. H., 1988. An analytic investigation of optimal bus size. *Transportation Research Part B* 22(5), 319-337.
- Parry, I. W. H., Small, K. A., 2009. Should urban transit subsidies be reduced? *American Economic Review* 99(3), 700-724.
- Proost, S., Van Dender, K., 2004. Marginal social cost pricing for all transport modes and the effects of modal budget constraints. In: *Road Pricing: Theory and Evidence*. G.Santos (Ed.). Elsevier, 159-177.
- Proost, S., Van Dender, K., 2008. Optimal urban transport pricing in the presence of congestion, economies of density and costly public funds. *Transportation Research Part A* 42(9), 1220-1230.
- RTA, 2011. Key Roads Performance Report. Roads and Traffic Authority, NSW Government, September 2011.

- Small, K. A., Verhoef, E. T., 2007. *The Economics of Urban Transportation*, Routledge, Taylor & Francis Group.
- TDC, 2010. 2008/09 Household Travel Survey. Summary Report, Transport Data Centre, Transport NSW.
- Tirachini, A., 2012. Multimodal pricing and the optimal design of bus services: new elements and extensions. PhD thesis, The University of Sydney. Available at <http://ses.library.usyd.edu.au/handle/2123/8584>.
- Tirachini, A., 2014. The economics and engineering of bus stops: Spacing, design and congestion. *Transportation Research Part A* 59, 37-57.
- Tirachini, A., Hensher, D. A., 2011. Bus congestion, optimal infrastructure investment and the choice of a fare collection system in dedicated bus corridors. *Transportation Research Part B* 45(5), 828-844.
- Tirachini, A., Hensher, D. A., 2012. Multimodal transport pricing: first best, second best and extensions to non-motorized transport. *Transport Reviews* 32(2), 181-202.
- Tirachini, A., Hensher, D. A., Jara-Díaz, S. R., 2010. Restating modal investment priority with an improved model for public transport analysis. *Transportation Research Part E* 46(6), 1148-1168.
- TRB, 2003. Transit capacity and quality of service manual. TCRP Report 100.
- Verhoef, E. T., Small, K. A., 2004. Product differentiation on roads. *Journal of Transport Economics and Policy* 38(1), 127-156.
- Wardman, M., Whelan, G. A., 2011. Twenty years of rail crowding valuation studies: evidence and lessons from British experience. *Transport Reviews* 31(3), 379-398.
- Whelan, G., Crockett, J., 2009. An investigation of the willingness to pay to reduce rail overcrowding. International Conference on Choice Modelling, Harrogate, England, April.
- Wichiensin, M., Bell, M. G. H., Yang, H., 2007. Impact of congestion charging on the transit market: An inter-modal equilibrium model. *Transportation Research Part A* 41(7), 703-713.
- Wright, L., Hook, W., 2007. *Bus rapid transit planning guide*, 3rd edition, Institute for Transportation and Development Policy, New York.
- Zhao, X.-m., Gao, Z.-y., Jia, B., 2007. The capacity drop caused by the combined effect of the intersection and the bus stop in a CA model. *Physica A: Statistical Mechanics and its Applications* 385(2), 645-658.

Appendices

A1 Estimation of demand models M1 to M4

The estimation of parameters for commuting and specification tests are presented in Table A1.1 ($n=1932$ observations):

Table A1.1: Estimation of parameters, MNL models

Parameter	M1	M2	M3	M4
Access time β_a	-0.016 (-1.22)	-0.017 (-1.33)	-0.017 (-1.33)	-0.017 (-1.33)
Headway β_h	-0.0088 (-2.71)	-0.010 (-3.06)	-0.010 (-3.07)	-0.010 (-3.06)
Travel time bus β_{vb}	-0.019 (-5.09)	-0.013 (-3.45)	-0.004 (-0.58)	-0.006 (-1.25)
Egress time β_e	-0.055 (-4.31)	-0.058 (-4.54)	-0.059 (-4.59)	-0.059 (-4.61)
Travel time car β_{va}	-0.016 (-3.11)	-0.018 (-3.41)	-0.018 (-3.37)	-0.018 (-3.37)
Cost β_c	-0.062 (-5.50)	-0.064 (-5.63)	-0.064 (-5.62)	-0.064 (-5.63)
MSC train α_t	-3.393 (-5.16)	-3.455 (-5.24)	-3.473 (-5.26)	-3.476 (-5.27)
MSC bus α_b	-4.131 (-5.67)	-4.275 (-5.82)	-4.313 (-5.86)	-4.315 (-5.86)
MSC metro α_m	-2.526 (-4.30)	-2.444 (-4.14)	-2.465 (-4.17)	-2.460 (-4.16)
$t_{vm} \times \text{den stand } \beta_{den}$		-0.004 (-4.48)	-0.003 (-2.82)	
$t_{vm} \times \text{prop seat } \beta_{seat}$			-0.013 (-1.70)	
$t_{vm} \times (\text{den stand})^2$ β_{den2}				-0.0005 (-2.41)
$t_{vm} \times (\text{prop seat})^2$ β_{seat2}				-0.013 (-2.46)
Specification tests				
Log-likelihood	-1283.4	-1273.1	-1271.7	-1271.7
Adjusted ρ^2 (relative to ASCs)	0.107	0.113	0.113	0.113
Likelihood ratio test with respect to M1		20.53 ($> \chi_{1,0.001}$ =10.83)	23.39 ($> \chi_{2,0.001}$ =13.82)	23.34 ($> \chi_{2,0.001}$ =13.82)
Likelihood ratio test with respect to M2			2.86 ($< \chi_{1,0.05}$ =3.84)	

Note: t -tests in bracket below parameter estimates. Time in minutes, cost in \$ (AUD).

Focusing on the goodness-of-fit measures, the log-likelihood and adjusted ρ^2 statistics relative to a model with alternative specific constants (ASCs) only, demonstrate that the three crowding models (M2-M4) outperform the model with no crowding (M1), but the difference in overall fitness amongst the crowding models is not significant. In fact, M2, M3 and M4 have the same adjusted ρ^2 value, and a likelihood ratio test indicates that M2, M3 and M4 are significantly superior than M1 at the 99.9 percent confidence level. Models M2 and M3 are not statistically different at 95 percent confidence level¹⁸, nevertheless from a behavioural perspective, the alternative crowding cost specifications do provide differences on the estimation of value of travel time savings¹⁹.

Parameters for the utility functions (2) to (7) are taken from Table A1.1, with the exceptions of the time parameter for walking and the mode specific constants, which are estimated as follows. First, walking as a travel alternative was not considered in the survey of the main stated choice experiment from 2009 in Sydney, described in Section 2.2 (Hensher *et al.*, 2011); therefore a reasonable value for the disutility of travel time while walking has to be supplemented. To this end, a secondary intra-CBD model described in an internal 2009 report by Hensher and Rose is used, in which walking was an alternative to public transport modes and taxi for short CBD trips; in the intra-CBD model, it is found that the time parameter of walking (β_{ve}) is 1.86 times greater than the in-vehicle time parameter for bus (β_{vb})²⁰. Thus, we assume a constant value of β_{ve} across models, equal to 1.86 times β_{vb} on M1 (because the latter is an average value of β_{vb} for all crowding conditions); therefore, $\beta_{ve} = 1.86 \cdot -0.019 = -0.035$.

Second, mode specific constants for demand models M1 to M4 are calibrated to represent the current Sydney modal split of trips shorter than 5 kilometres: 62.5 percent car, 31.6 percent walk, and 5.9 percent bus (TDC, 2010). The current bus frequency of 16 bus/h in the morning peak is used, with a fare of \$2.10 and no car toll. The car specific constant is fixed at zero. With these two considerations for the time parameter for walking and the mode specific constants, the estimated parameters used in Section 3 are presented in Table 2.

A2 Estimation of queuing delay function

To estimate the queuing delay of buses we use the bus stop simulator IRENE (Gibson *et al.*, 1989; Fernández and Planzer, 2002). A total of 265 simulations were run encompassing all bus sizes (8-, 12-, 15- and 18-metre long buses) and bus stop with one, two and three linear berths, for a range of frequencies from 20 to 220 bus/h and dwell

¹⁸ This would suggest that if we use the density of standing to characterise crowding costs, the inclusion of the availability of seats as a variable that influences modal choice is not statistically relevant for commuting. However, the occupancy level that triggers a crowding disutility is usually lower for non-commuters than for commuters; in fact, some studies find that the crowding disutility is activated with occupancy levels as low as 40 percent of the seat capacity for leisure travelers (for a review see Wardman and Whelan, 2011). Therefore, it is relevant to include in the analysis models in which the availability of seats is also a source of utility.

¹⁹ Models M2 and M4 cannot be compared with a likelihood ratio test because they are not nested.

²⁰ This figure is in the order of the values estimated by Jovicic and Hansen (2003) for Copenhagen (1.36 and 2.32 for the ratio β_{ve}/β_{vb} for purposes commuting and education, respectively, considering walking and cycling altogether as a non-motorised mode, and trips up to 30 minutes long).

times between 10 and 65 seconds. Buses are assumed to arrive at a constant rate at stops (no bus bunching) and bus stops are isolated from traffic lights. Estimated parameters are presented in Table A2.1. For more details see Tirachini (2014).

Table A2.1: Queuing delay parameters

Parameter	Estimate	Std. Error
b_0	-2.952	0.887
b_{l1}	0.061	0.020
b_{d1}	2.185	0.530
b_{d2}	-1.903	0.495
b_{d3}	-2.044	0.510
b_f	23.089	0.723
b_{l2}	0.361	0.046
b_{d4}	1.807	0.091
b_{d5}	-0.374	0.093
b_{d6}	-0.627	0.087
R^2	0.921	
Sample size	265	

A3 Dwell time estimation

When there are multiple doors to board and alight, passengers can choose a door to get on and off buses, and the spatial dispersion of their decision will determine how long the boarding and alighting times are. It seems unreasonable to suppose that passengers will distribute uniformly across doors if middle or back doors have closer access to more seats than, say, the front door. We assume that the middle doors would attract a number of passengers that is 50 percent higher than that of the front or back doors (for example, for buses with two doors, the rear door is placed towards the centre of the bus, and is therefore assumed to attract 60 percent of the boarding demand, leaving 40 percent boarding through the front door, next to the driver). The same assumption is made regarding alighting. With this, the proportions p_b and p_a of passengers boarding and alighting at the busiest door (necessary for estimation of dwell times in equation 15), is given in Table A3.1

Table A3.1: proportion of passengers boarding and alighting at the busiest door

Number of doors	$TnBn$ $p_a (= p_b)$	$TnB1$ p_a
1	100%	
2	60%	100%
3	43%	60%
4	30%	38%

Boarding and alighting times per fare collection system and boarding policy are taken from Fernández *et al.* (2008) and Tirachini (2012):

Table A3.2: Average boarding and alighting times per passenger

Parameter	Time [s/pax]
Boarding time cash	10.74
Boarding time magnetic strip	2.94
Boarding time contactless card	2.05
Boarding time off-board payment	1.46
Alighting time	1.46

A4 Operator cost

The station infrastructure cost depends on the amenities provided, quality of shelter and overall design, ranging from \$15,000 for a simple shelter to \$150,000 or more for stations with passenger enclosure, at-level boarding, retail services and detailed passenger information (FTA, 2009). In this paper, we assume that the cost increases linearly with bus length: \$50,000 (8 m. bus), \$75,000 (12 m. bus), \$100,000 (15 m. bus) and \$125,000 (18 m. bus), values that are amplified by 25 percent if off-board payment is provided.

There are two vending machines per station, and four fare collection readers in case of off-board payment. Fare collection costs are taken from Wright and Hook (2007), the cost of a fare collection reader is \$750 (coins), \$1,750 (magnetic strip) and \$2,500 (contactless card), while the cost per vending machine is \$10,000 (magnetic strip) and \$15,000 (smart card). For on-board payment methods, two fare collection readers per boarding door are considered. The cost of software is \$100,000 for coin payment, \$300,000 for magnetic strip and \$500,000 for contactless card. Bus driving cost is \$29.9 (Hensher, 2010) and running costs are 0.6, 0.9, 1.0 and 1.1 \$/veh-km for 8, 12, 15 and 18 metre long buses, respectively (ATC, 2006), values that are increased by 21 percent to account for overhead costs. The cost of buses is \$160,000 (8 m.), \$370,000 (12 m.), \$520,000 (15 m.) and \$700,000 (18 m.). The estimated parameters in Tables A4.1 and A4.2 are adjusted to 2011 Australian Dollars assuming 20 years of asset life for buses, 15 years for stations and 5 years for software, card readers and vending machines; one year is equivalent to 2947 peak hours of operation for a typical urban bus service in Australia.

Table A4.1: Cost items related to bus size

Bus size [m]	Bus cost [\$/bus-h]	Driver cost [\$/bus-h]	Station cost [\$/station-h]	Operating cost [\$/bus-km]
8	5.1	37.6	4.4	0.9
12	11.9	37.6	6.5	1.3
15	16.9	37.6	8.7	1.4
18	22.0	37.6	10.9	1.6

Table A4.2: Cost items related to fare collection technology

Technology	Software cost [\$/h]	Card reader [\$/h]	Vending machine [\$/h]
Coin	12.1	0.1	0.0
Magnetic strip	36.3	0.2	2.4
Contactless card	60.5	0.3	3.6
Off-board	60.5	0.3	3.6