Research in Transportation Economics 29 (2010) 231–242

Contents lists available at ScienceDirect

Research in Transportation Economics

journal homepage: www.elsevier.com/locate/retrec

Comparing operator and users costs of light rail, heavy rail and bus rapid transit over a radial public transport network

Alejandro Tirachini [a,*], David A. Hensher [a], Sergio R. Jara-Díaz [b]

[a] Institute of Transport and Logistics Studies, The University of Sydney, Australia
[b] Civil Engineering Department, Universidad de Chile, Chile

Keywords:
Bus rapid transit
Light rail
Heavy rail
Frequency
Number of lines

A B S T R A C T

A model to compare three alternative forms of public transport – light rail, heavy rail and bus rapid transit - is developed for an urban network with radial lines emanating from the borders to the city centre. The theoretical framework assumes an operation aimed at minimising the total cost associated with public transport service provision, which encompasses both operator and users costs. The decision variables are the number of lines (network density) and the frequency per period for each mode. This approach has no prejudices a priori in respect of whether a specified delivery scenario is aligned with existing modal reputation. Rather, we establish the conditions under which a specific transit mode should be preferred to another in terms of the operator (supply) and user (demand) side offerings. The model is applied using data from Australian cities, suggesting that in most of the scenarios analysed a high standard bus service is the most cost-effective mode, because it provides lower operator costs (infrastructure, rolling stock and operating cost), access time costs (due to a larger number of lines) and waiting time cost (due to larger frequencies of operation). A rail mode, such as light rail or heavy rail, may have a lower total cost only if it is able to run faster than bus rapid transit, and the difference in speed is enough to outweigh the bus advantage on operator cost and access and waiting times.

1. Introduction

The provision of specific forms of urban public transport should be guided by the level of demand to be served, the quality of service desired and budget and operational constraints faced by the transit agencies and regulatory authorities. A systematic assessment of public transport needs should also be placed within a framework that recognises the strategic, tactical and operational (STO) elements of planning (Desaulniers & Hickman, 2007), in line with the STO paradigm promoted in the Thredbo Series. A key strategic decision in urban transport policy is the choice of technology for a transit system, in particular the choice amongst rail (light and heavy) and/or bus based systems, to provide service in a particular corridor and across a network. Trains and buses may be different in almost every aspect of the provision of a transit service, such as capital and operating costs, speed, frequency, distance between stops, infrastructure requirements, reliability and comfort; hence the decision must be as informed as possible. Moreover, careful consideration in the selection of a specific mode is also needed within the overarching aim of providing maximum net social benefit per tax payers dollar outlaid (commonly referred to a 'value for money').

This paper develops a framework within which to compare alternative forms of public transport for an urban network with radial lines emanating from the borders to the CBD, with the objective of minimising the total cost, defined as the user cost plus the operator cost. Both the number of lines and service frequency are decisions conditional on the choice of technology for the network. The problem focuses on the design of a trunk network considering three possibilities for the transit technology: Light Rail (LR), Heavy Rail (HR) and Bus Rapid Transit (BRT).[1]

Comparisons between trains, buses and other modes have been previously addressed by several studies, considering a single line (Meyer, Kain, & Wohl, 1965; Smith, 1973), and trunk lines with feeder routes, where the number of lines is fixed and given (Allport, 1981; Boyd, Asher, & Wetzler, 1978; Bruun, 2005; Dewees, 1976; Meyer et al., 1965). Boyd et al. (1978), Dewees (1976) and Smith (1973) found that buses are more cost effective than trains in all the

* Corresponding author.
E-mail address: alejandro.tirachini@sydney.edu.au (A. Tirachini).

[1] BRT can be defined as "a high-quality bus based transit system that delivers fast, comfortable, and cost-effective urban mobility through the provision of segregated right-of-way infrastructure, rapid and frequent operations, and excellence in marketing and customer service" (Wright & Hook, 2007).

0739-8859/$ – see front matter
doi:10.1016/j.retrec.2010.07.029

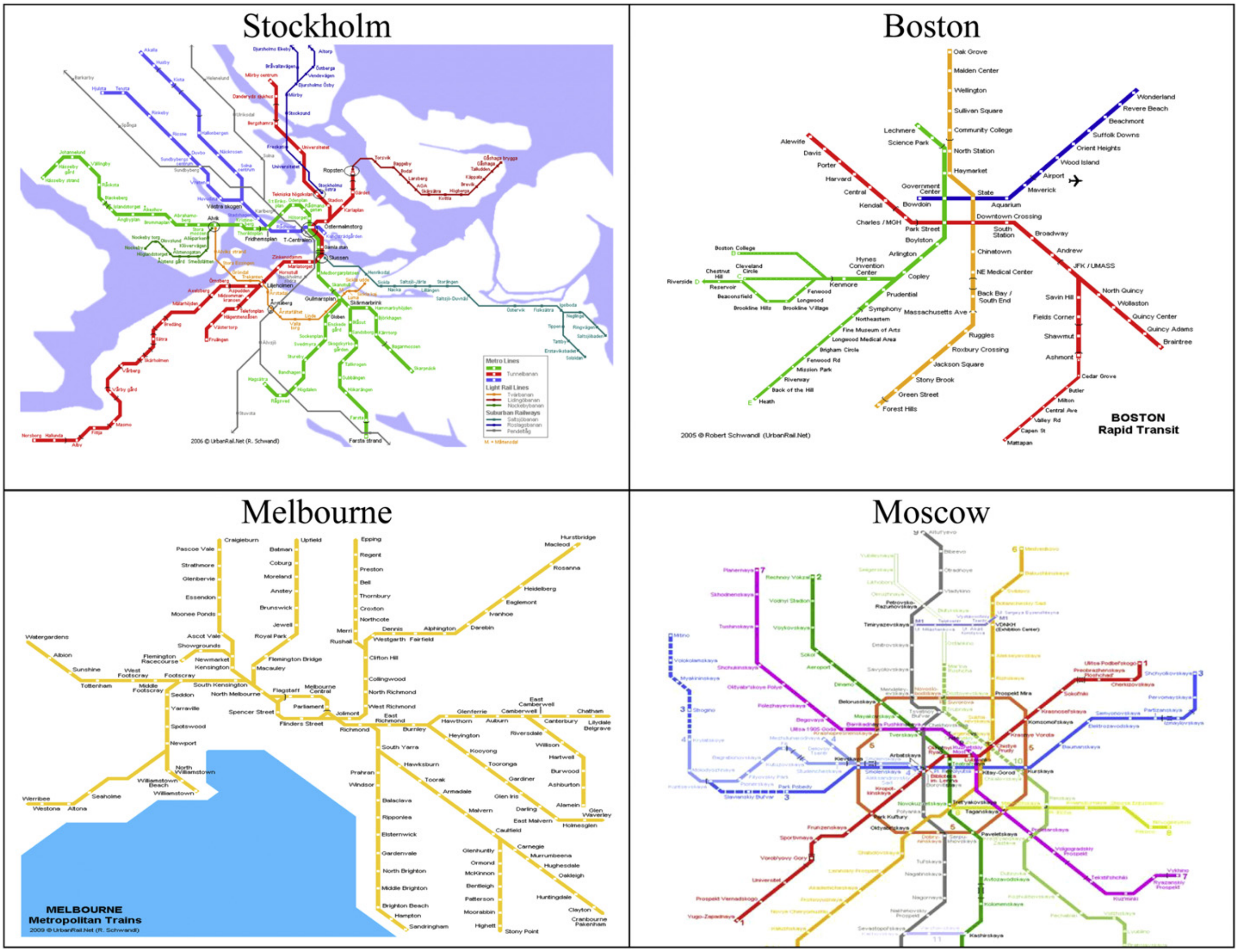

Fig. 1. Examples of radial rail transport networks. (source: www.urbanrail.net).

situations and demand levels considered, when comparing only operator costs (Smith, 1973), only users costs (Dewees, 1976) or both user and operator costs (Boyd et al., 1978). These results are generally explained by lower operator cost and/or access and waiting time costs provided by buses, that outweigh a potential advantage of speed from trains. On the other hand, other research finds different results depending on the demand level to be served; Meyer et al. (1965) estimate the operator cost for different modes, finding that for a demand up to 5000 pax/h one way, the car provides the lowest cost, while after that level, buses minimises cost, except in high density areas where trains could be more cost effective than buses for demands over 30,000 pax/h. In a similar fashion, Allport (1981) found that on the basis of operating costs only, bus is the least cost mode for low demand, with light rail transit the lowest cost in the middle demand range; but when users costs are included the bus advantage extends towards larger demand levels, due to the assumed bus advantage on access and waiting times (larger frequency and shorter distance between stations for buses). In a more recent study, Bruun (2005) develops a parametric cost model to compare operating costs for BRT and LR finding situations in which LR dominates BRT, considering that to accommodate peak demands, cars can be added or removed from a train, whereas in the case of buses, an increase of capacity is only achieved adding separate vehicles, both cases reporting different marginal costs.

The aim of this paper is to develop a model to identify the conditions under which BRT makes better sense than heavy or light rail and vice versa (Hensher, 2007; Litman, 2007), given the

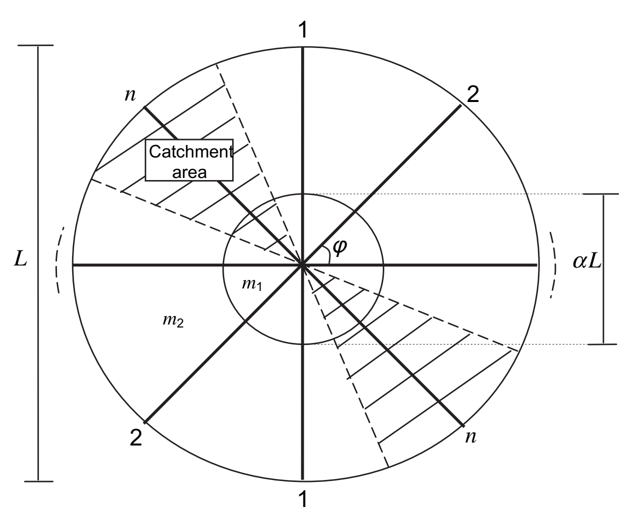

Fig. 2. Public transport network.

Table 1
Effect of frequency and number of lines on costs.

Variable	Influence on:			
	Access time	Waiting time	In-vehicle time	Operator cost
Frequency	No	Yes, waiting time increases with average headway, or decreases with frequency	Yes, if dwell time depends on passengers boarding and alighting at stations, which is a function of frequency	Yes, operator cost increases with frequency
Number of lines	Yes, distance to walk to a line decreases with the number of lines	No	Yes, if dwell time depends on passengers boarding and alighting at stations, which is a function of the number of lines	Yes, the operator cost increases with the number of lines

objective of total cost minimisation, i.e., considering total operator costs and access, waiting and in-vehicle time costs in the comparisons. We deal with a radial trunk network in line with the structure of many real urban public transport systems, characterised by a centre (usually the Central Business District or CBD) with a set of radial lines, as shown in the examples of Fig. 1. The frequency and number of lines are the decision variables. The design of a feeder network is not considered. Unlike previous literature, we optimise both frequency and number of lines, accounting for the effect of both variables in the access, waiting and in-vehicle time cost for users, and on the total operator costs (taking into account a combination of land, infrastructure and operating costs). In particular, in this model different number of lines provided by the modes will have an impact on the access time cost that has not been considered in the previous literature that compares bus and rail public transport systems. The negative effect of crowding on users is also taken into account as a mark up on in-vehicle time cost, which is also a novel aspect of transit modal comparisons.

The following sections of the paper are organised as follows. Section 2 presents the model assumptions and input formulae used to define the optimisation problem; in Section 3 several applications are undertaken to assess the implications of scenarios centred around specific treatments of operator and user costs; Section 4 sets out the main findings.

2. Model

2.1. Assumptions and definitions

The real systems in Fig. 1 are abstracted herein as a set of straight radial lines evenly spaced over a circular area of diameter L, like a bike wheel, as shown in Fig. 2. There are n lines of length L, and φ is the angle between lines. Demand is assumed to be cyclical and fixed over T periods. The city is divided into two areas, each of which has an evenly distributed demand: an *inner circle* or *CBD* of diameter αL, $0<\alpha<1$ with demand density m_{1t} per area unit [trips/km^2] for period t, and *the suburbs*, outside the CBD, where the demand density is m_{2t} [trips/km^2] in period t as shown in Fig. 2. The total demand inside (y_{1t}) and outside (y_{2t}) the CBD is defined by equation (1).

$$y_{1t} = \frac{\pi(\alpha L)^2}{4} m_{1t}, \quad y_{2t} = \frac{\pi L^2 (1-\alpha^2)}{4} m_{2t} \tag{1}$$

As the demand is evenly distributed both inside and outside the CBD, the demand catchment area for a line encompasses an arc of angle $\varphi/2$ to the left and $\varphi/2$ to the right of the line, and consequently, the demand per line is y_t/n, where $y_t = y_{1t} + y_{2t}$ is the total demand in period t. Fig. 2 depicts the catchment area for line n. The relation between φ and the number of lines n is given by $n\varphi = \pi(\text{rad})$.

The problem posed is to find optimal values for the number of lines and frequencies in each period, to take into account the coverage and accessibility provided by a particular transit mode throughout an urban area. Service frequency and the number of lines are assumed to affect three of the four components of the total cost function, *ceteris paribus*, as shown in Table 1.

2.2. User and operator cost functions and problem formulation

In what follows we set out the different cost components for a transit service. Access time cost is computed only at the origin, as no assumption is made about the distribution of destinations. Access time to the transit system has two parts, access *to the line* and access *along the line* (to the station). For simplicity, access to the line is assumed to be undertaken by walking in concentric circles around the centre, as shown in Fig. 3, i.e., for a person whose position is at a distance r from the centre and angle λ from the closest transit line, his or her walking distance is r times λ.

Given that demand is evenly distributed inside and outside the CBD, after integrating over the catchment area (see Appendix A1), the total access distance to the line for period t, d_{at}, is

$$d_{at} = \frac{\varphi L}{12}(\alpha y_{1t} + g(\alpha) y_{2t}) = \frac{\pi L}{12n}(\alpha y_{1t} + g(\alpha) y_{2t}) \tag{2}$$

where $g(\alpha) = 1 + ((\alpha^2)/\alpha + 1)$. For access time along the line, if d [km] is the distance between two consecutive stations or stops, passengers have to walk on average $d/4$ to reach the closest station. Hence, if D_t is the duration of period t [h], P_a is the value of access time savings [\$/h] and v is the passenger walking speed, the total cost associated with the access time, C_a, is:

$$C_a = P_a \sum_{t=1}^{T} \frac{D_t}{4v}\left[\frac{\pi L}{3n}(\alpha y_{1t} + g(\alpha) y_{2t}) + \mathrm{d} y_t\right] \tag{3}$$

For the waiting time cost formulation, we will distinguish between services with high and low frequency, as the behaviour of passengers is different in the two cases; when the frequency is high passengers usually arrive at the stations randomly at a constant rate, but when the frequency is low, generally there is a timetable of services and most of the users arrive at stations following the schedule in order to reduce their waiting time. There is a threshold "forget the timetable" headway that separates both regimes, which is considered to be between 10 and 15 min (RAND, 2006).[2] In the applications (Section 3) we will use that for frequencies greater than 5 veh/h, equivalent to an average headway up to 12 min, users arrive randomly at stations, in which case the average waiting time cost w_t can be modelled as a fraction $\varepsilon \geq 0.5$ of the headway:

[2] Note that for low frequency services there is still a number of passengers who arrive randomly at stations (mainly for not having timetabling information), case that will not be considered in this model.

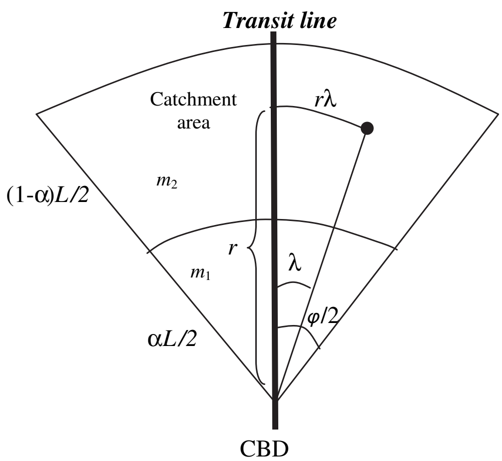

Fig. 3. Access distance to a particular transit line.

$$w_t = P_w \frac{\varepsilon_t}{f_t} \tag{4}$$

where P_w is the value of waiting time savings [$/h] and f_t is the frequency [veh/h] in period t. For example, if headways are regular, the average waiting time is half of the headway, i.e. $\varepsilon = 0.5$. On the other hand, when users arrive at stations according to the timetable of services (low frequency), there is a waiting time cost outside stations because departures are not at a user desired time (called "schedule delay"), a penalty that increases the longer is the headway between vehicles. As this *passive waiting time* can be spent at home or another place where passengers can assign their time to a more productive use or leisure, the opportunity cost or value of *passive waiting time* savings, P_h, is lower than the value of station or *active waiting time* savings, P_w. Thus, the average waiting time cost can be modelled as:

$$w_t = P_w\left(t_w + \mu\frac{\varepsilon_t}{f_t}\right) \tag{5}$$

$\mu = P_h/P_w$ is the ratio of the value of *home waiting time* savings to the value of *station waiting time* savings (for example, 0.33), and t_w is a fixed 'safety threshold' time that passengers spend waiting at stations before the expected arrival of the next vehicle. Then, the total waiting time cost C_w is derived as

$$C_w = P_w \sum_{t=1}^{T} D_t\left(t_{0t} + t_{1t}\frac{\varepsilon_t}{f_t}\right) y_t \tag{6}$$

where

$$t_{0t} = \begin{cases} 0 & \text{if } f_t \geq 5 \text{ veh/h} \\ t_w & \text{if } 0 < f_t < 5 \text{ veh/h} \end{cases}$$

$$t_{1t} = \begin{cases} 1 & \text{if } f_t \geq 5 \text{ veh/h} \\ \mu & \text{if } 0 < f_t < 5 \text{ veh/h} \end{cases}$$

Implicit in (4) is that the capacity constraint of the vehicles is not binding; in fact the frequency will be set to avoid overloading of vehicles (as will be seen in equation (13) below). It is also worth noting that, to account for service unreliability, a higher value of ε for a less reliable mode can be used in (6).[3]

For the estimation of the in-vehicle time, directional demands y_t^i (inbound) and y_t^o(outbound) are defined, such that $y_t^i + y_t^o = y_t$. The total travel time between terminals t_{dt} [h] is the sum of the running time between stations R_t[4] and the dwell time (time spent at stations boarding and alighting passengers), which is the product of the number of passengers that board and alight a bus (given by y_t^r/nf_t in direction r), and the boarding and alighting time per passenger β [h/pax]. Then t_{dt} per direction is:

$$t_{dt}^i = \frac{y_t^i}{n f_t}\beta + R_t \text{ and } t_{dt}^o = \frac{y_t^o}{n f_t}\beta + R_t \tag{7}$$

Users' in-vehicle time is modelled as a fraction of l_t/L of t_{dt}, where l_t [km] is the average trip length and L [km] is the route length (obviously, $l < L$). On the other hand, the in-vehicle time cost depends on the trip conditions, as an increasing amount of evidence shows that transit users are willing to pay more for a less crowded trip[5] (Maunsell and Macdonald, 2007; Pepper, Spitz, & Adler, 2003; Polydoropoulou & Ben-Akiva, 2001; Whelan & Crockett, 2009), often expressed in terms of the probability of getting a seat. This effect can be captured using a mark up on the value of in-vehicle time savings P_v [$/h] when vehicles get crowded and a number of passengers have to stand (Jara-Díaz & Gschwender, 2003; Kraus, 1991). To incorporate this crowding disutility, we can define a crowding penalty factor $\delta_t(\theta_t) \geq 1$ as a function of the average occupancy rate $\theta_t = k_t/K$, where K is the capacity of vehicles (seating + standing) [pax/veh] and k_t is the average load per vehicle [pax/veh], expressed per direction as follows:

$$k_t^i = \frac{l_t^i}{L}\frac{y_t^i}{n f_t} \text{ and } k_t^o = \frac{l_t^o}{L}\frac{y_t^o}{n f_t} \tag{8}$$

A quadratic form for the crowding penalty $\delta_t(\theta_t) = 1 - 0.2381\,\theta_t + 1.1905\,\theta_t^2$ will be used on the applications (Section 3) as justified and derived in the Appendix A2. With this, the in-vehicle time cost is modelled as:

$$C_v = P_v \sum_{t=1}^{T} D_t\left[\delta_t\left(\theta_t^i\right)\frac{l_t^i}{L}\left(\frac{y_t^i}{n f_t}\beta + R_t\right) y_t^i + \delta_t\left(\theta_t^o\right)\frac{l_t^o}{L}\left(\frac{y_t^o}{n f_t}\beta + R_t\right) y_t^o\right] \tag{9}$$

Assuming the running time R_t to be fixed may not represent a situation with congestion both *across* modes, i.e., other modes (notably cars) causing public transport congestion, or *within* a mode, in which buses or trains must slow down or stop because the bus or train ahead is impeding them to move. We are assuming that the three alternative modes under analysis, LR, HR and BRT, run in segregated rights of way, thus congestion across modes should be negligible (except at intersections in some cases). On the other hand, congestion between transit vehicles is more likely when the frequency is high in bus systems, but assuming that BRT and the rail modes are centrally controlled to avoid bunching, this problem can also be prevented to a large extent.

Operator cost C_o has four components. The first element is the cost per line c_0[$/line-day], in which a combination of land and infrastructure capital and maintenance costs can be included. The second component c_1 [$/veh-day] is the rolling stock capital cost,

[3] Unreliability in vehicle running times leads to bunching of vehicles on un-scheduled services and unexpected delays on scheduled operation, effects that have an impact on longer waiting times.

[4] R_t is the time necessary to cover distance L, it includes the acceleration and deceleration times at stations, opening/closing doors and other station dead times, but does not consider passenger boarding and alighting times.

[5] In unpublished research Hensher found a 0.7 mark up on the VTTS per person hour to ensure the availability of a seat 45 percent of the time.

which is given by the fleet requirement in the peak period. The third component c_{2t} [$/veh-h] is the unit operator cost per vehicle-hour, including personnel cost. Finally, the fourth part c_{3t} [$/veh-km] is the unit operator cost per vehicle-kilometre and comprises running costs such as fuel consumption, lubricants, tyres, etc. Other cost components such as overhead costs (scheduling, rostering, supervision, depot-related costs, human resources, etc.) will be also considered in the applications of the model. Define B_t as the fleet size requirements of period t; $\eta > 1$ as a factor that accounts for a reserve fleet to deal with unexpected breakdowns and maintenance (e.g., $\eta = 1.05$ meaning that 5% of vehicles are not used and kept at depots); V_t is the commercial speed (cycle time divided by total route length) [km/h]. Thus:

$$C_o = c_0 n + c_1 \eta \max_t B_t + \sum_{t=1}^{T} D_t (c_{2t} B_t + c_{3t} V_t B_t) \tag{10}$$

The fleet size requirement per period is given by the frequency, the number of lines and the total travel time to complete a cycle (cycle time) $t^i_{dt} + t^o_{dt}$:

$$B_t = n f_t \left(t^i_{dt} + t^o_{dt} \right) = y_t \beta + 2n f_t R_t \tag{11}$$

Using (11) and t_{ct} as $2L/V_t$ in the last term of (10), the operator's cost final expression is:

$$\begin{aligned} C_o = {} & c_0 n + c_1 \eta \max_t (y_t \beta + 2n f_t R_t) + \sum_{t=1}^{T} D_t [c_{2t} y_t \beta \\ & + 2n f_t (c_{2t} R_t + c_{3t} L)] \end{aligned} \tag{12}$$

After deriving the access, waiting and in-vehicle time costs plus the operator cost, the total cost C_t can be formulated as the sum of equations (3), (6), (9) and (12), which is the function that must be minimised to obtain the optimal value of the frequency and number of lines. The capacity constraint plays a role in setting a minimum value for the frequency as

$$f_{\min t} = \max\left\{ \kappa^i y^i_t, \kappa^o y^o_t \right\} \frac{1}{n \upsilon K} \tag{13}$$

where κ^i and κ^o are the fractions of passengers that transverse the section of the line with maximum load in directions i and o, respectively, and υ is a safety factor introduced to have spare capacity to absorb random variations on demand (for example, $\upsilon = 0.8$). On the other hand, frequencies are also constrained by the maximum feasible frequency $f_{\max}$ of each mode, given by the capacity of

Table 2
Modal parameters.

Parameter	Units	BRT	LR	HR
c_{0l} (Land)	$/line-day	51,781	51,781	77,671
c_{0i} (Infrastructure)	$/line-day	59,556	115,201	201,602
c_{0m} (Maintenance)	$/line-day	3579	13,057	22,413
c_1	$/veh-day	158	701	1721
c_2	$/veh-h	42	73	130
c_3	$/veh-km	1.42	1.83	3.32
Speed	km/h	30	35	40
Capacity	pax/veh	101	190	750
β	s/pax	0.33	0.33	0.11
s	km	1	1	1
$f_{\max}$	veh/h	120	80	40

Source: ATC (2006) and Evans (2005) as explained in the Appendix. Figures in 2007 Australian Dollars (Euro 1 ≈ AUD 1.7). Crew and operating cost are assumed equal for all periods based on evidence from Sydney.

Table 3
Periods and demand densities.

Period	Time of day	Duration (h)	m_1 [trips/h-km^2]	m_2 [trips/h-km^2]	y_1 [trips/h]	y_2 [trips/h]
1	06:00–07:00	1	400	60	7238	41,326
2	07:00–09:00	2	950	110	17,191	75,764
3	09:00–15:00	6	600	30	10,857	20,663
4	15:00–16:00	1	1550	90	28,048	61,989
5	16:00–18:00	2	2300	60	41,620	41,326
6	18:00–0.00	6	400	10	7238	6888

stations and safety considerations. Accordingly, defining the fleet size as $B = \max_t (y_t \beta + 2n f_t R_t)$, the optimisation problem is:

$$\begin{aligned} \min_{f_1 \ldots f_T, n, B} C_t = {} & P_a \sum_{t=1}^{T} \frac{D_t}{4v} \left[\frac{\pi L}{3n} (\alpha y_{1t} + g(\alpha) y_{2t}) + d y_t \right] \\ & + P_w \sum_{t=1}^{T} D_t \left(t_{0t} + t_{1t} \frac{\varepsilon_t}{f_t} \right) y_t \\ & + P_v \sum_{t=1}^{T} D_t \left[\delta_t \left(\theta^i_t \right) \frac{l^i_t}{L} \left(\frac{y^i_t}{n f_t} \beta + R_t \right) y^i_t \right. \\ & \left. + \delta_t (\theta^o_t) \frac{l^o_t}{L} \left(\frac{y^o_t}{n f_t} \beta + R_t \right) y^o_t \right] + c_0 n + c_1 \eta B \\ & + \sum_{t=1}^{T} D_t [c_{2t} y_t \beta + 2n f_t (c_{2t} R_t + c_{3t} L)] \end{aligned}$$

subject to

$$\begin{aligned} & y_t \beta + 2n f_t R_t \le B \quad \forall t \in \{1, \ldots, T\} \\ & \max\{\kappa^i y^i_t, \kappa^o y^o_t\} \tfrac{1}{n \upsilon K} \le f_t \le f_{\max} \quad \forall t \in \{1, \ldots, T\} \\ & n \in N \end{aligned}$$

which must be solved numerically. The algorithm chosen to find the solution, implemented in Matlab, has two parts: first, the problem is solved assuming that n is a real number,[6] then the solution for n is rounded to its lower and upper closest integer (n_l and n_u, respectively), and the problem is solved again using n_l and n_u (as given values), choosing the solution that provides the lowest total cost.

3. Application

3.1. Assumptions

We conduct several experiments to analyse the performance of Bus Rapid Transit (BRT), Light Rail (LR) and Heavy Rail (HR), under different operating conditions and assumptions. Table 2 presents the mode-specific parameters. The calculation of the operator cost unit values is based on data from ATC (2006) and Evans (2005), and is representative of urban transport systems in Australia (see Appendix A3 for details).

Increasing operator costs are assigned as we move from bus to rail modes, with HR defined as the fastest mode, followed by LR and then BRT. In addition, we assume the same value of travel time savings (VTTS) for the three modes: $P_a = 12.5$ $/h (Value of access time savings), $P_w = 15$ $/h (Value of waiting time savings), $P_v = 10$ $/h

[6] Other authors that optimise frequency and spacing between parallel lines over a rectangular area (e.g., Chang and Schonfeld, 1991), assume that the spacing between lines (width of the total area divided by the number of lines) is a real number. Unless the resulting spacing is a multiple of the width of the total area (equivalent to say that there is an integer number of lines), there will be a problem in the location of at least one line. This is a minor technical issue that have no impact on results, but should be considered in the present model as we take the number of lines as the variable, because it is easier to visualize than the angle between lines.

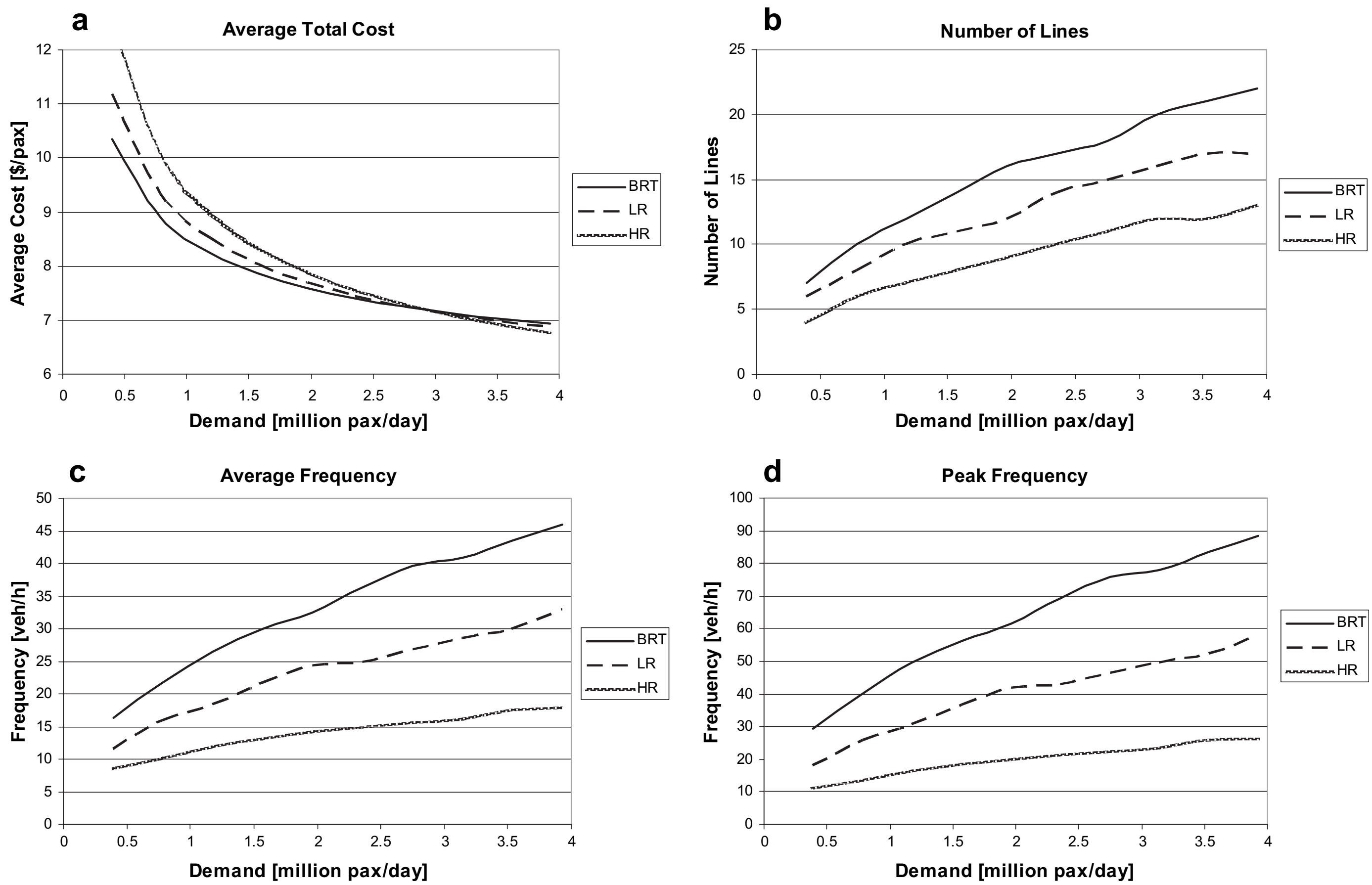

Fig. 4. Results Case 1.

(Value of in-vehicle time savings). For waiting time, we assume $\varepsilon = 0.5$, $\mu = 0.33$ and $t_w = 4$ min (consequently, expressions (4) and (5) give the same result for the assumed threshold frequency of 5 veh/h).

The route length is $L = 30$ km,[7] the average trip length is $l = 10$ km, and the walking speed is $v = 4$ km/h. The demand densities m_{1t} and m_{2t} [trips/h-km^2] are estimated drawing on data from Sydney, obtained from the simulation program TRESIS (described in Hensher, 2008). The CBD is the "Inner Sydney" zone in TRESIS, whose length is approximately 5 km, therefore we use $\alpha = 16.7\%$ (5/30). An operation of 18 h per day is assumed, which is divided in six periods by TRESIS, as shown in Table 3. The estimated m_{1t} and m_{2t} account for trips in both trains and buses from the current Sydney public transport system. Using the figures in Table 3, the total number of trips per day is 764,278. It is also assumed that the demand in each direction is half of the total demand and the maximum load is 50 percent of the demand, across all periods and directions (parameter κ in equation (13)).[8]

The parameter η (for the calculation of the fleet size) is 1.05 and the safety factor υ is 0.8.

The simulations are undertaken over a range of demand levels, obtained by expanding the demand densities in Table 3 up to five times (total demand of 3,821,389 trips per day), to represent other urban contexts with a larger number of transit trips than Sydney. Several cases are developed; the first two differ in terms of components inputs defining the fixed cost per line:

Case 1: Unit cost c_0 includes land acquisition, infrastructure development and maintenance costs ($c_0 = c_{0l} + c_{0i} + c_{0m}$ in Table 2).
Case 2: Unit cost c_0 includes infrastructure development and maintenance costs only ($c_0 = c_{0i} + c_{0m}$). It represents a case in which new busways or railways are built using available land (for example, existing avenues and highways) and no land has to be bought.

Arguably, reality lies between Cases 1 and 2, as for the development of new high standard transit lines in consolidated urban areas some land has to be acquired, but by no means the whole area required by a new railway or busway, since generally large parts are traced over existing public space. We choose Case 2 to continue testing other scenarios and modelling assumptions:

Case 2A: The crowding penalty factor is not considered
Case 2B: The BRT and LR running speeds are increased to find threshold speed
Case 2C: The running speed is equal for all modes.

[7] In general, big cities around the world have an outer metropolitan area that exceeds 30 km of length, encompassing less populated suburbs and satellite towns, which are commonly served by high-speed commute trains that may also reach a central station in the CBD. In this paper we focus on a trunk public transport network for the core metropolitan area, i.e., the interaction between a long distance public transport service and the local trunk network is not treated.

[8] More refined information on demand may be desirable (different values of k per direction and period, different values of the directional demands y_t^i and y_t^o); certainly this could change the optimal values of the variables but not the general conclusions obtained in this paper.

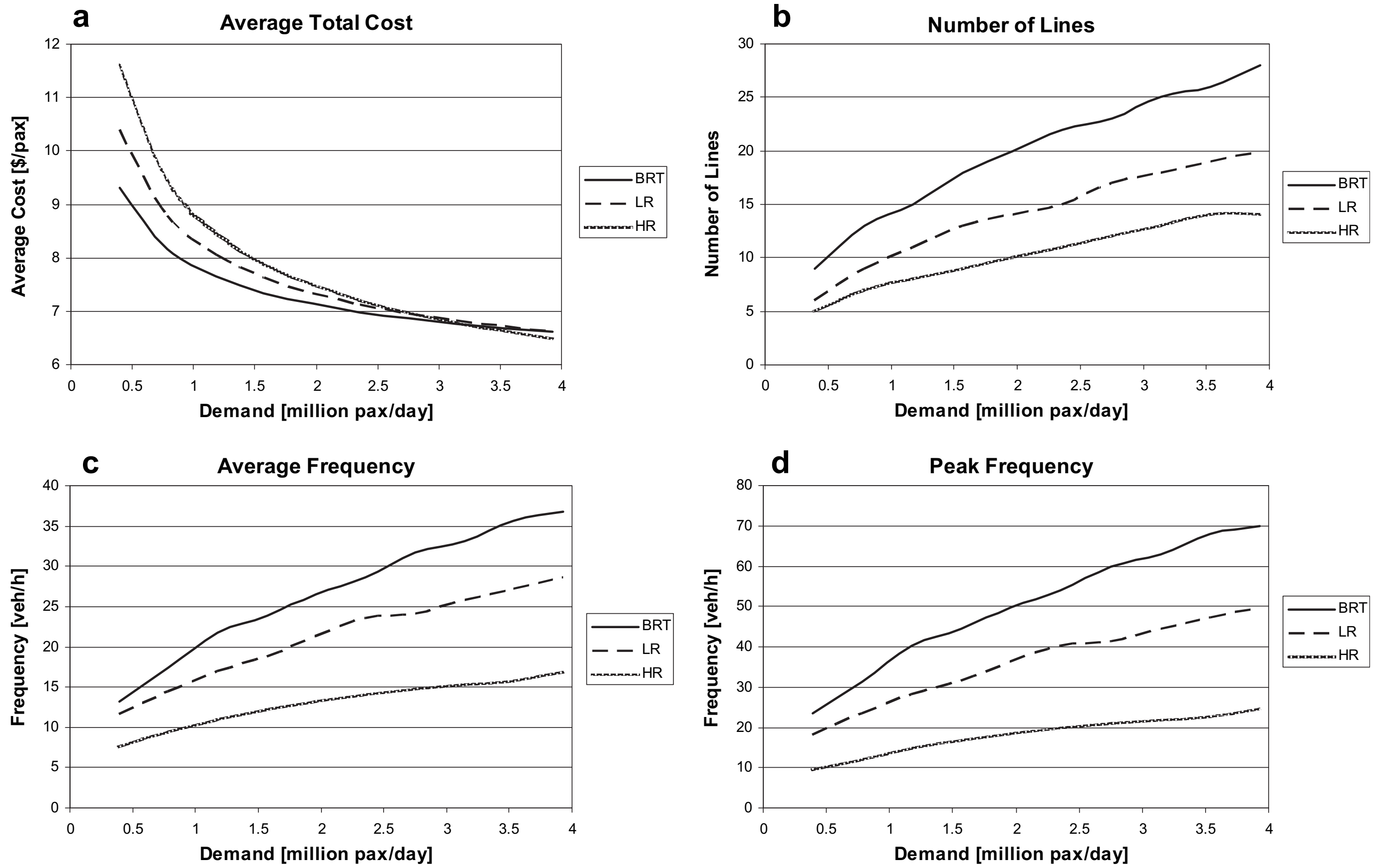

Fig. 5. Results Case 2.

Case 2D: Case 2C + a decrease of 25 percent in the operator cost of rail modes
Case 2E: Case 2D + fixed dwell time at stations and BRT assumed less reliable than LR, which is less reliable than HR.

3.2. Total cost, number of lines and frequencies

In what follows the main results for every case are presented and interpreted.

3.2.1. Case 1: all fixed inputs are allowed for

The average total cost evolution can be observed in Fig. 4a. It suggests that for demands lower than 2.8 million-pax/day, BRT is the most cost-effective mode, whilst for larger demands HR emerges as the best alternative, with LR always in second place. Over the whole range of demand, BRT provides more lines (Fig. 4b) and larger frequencies (Fig. 4c for the average frequency along all periods, Fig. 4d for the peak frequency corresponding to Period 2), followed by LR in all these aspects, which means that users perceive lower access and waiting time costs with a BRT service. The result for total cost reveals that for low demand it is more convenient to have a less expensive mode, more frequent (lower waiting time) and with more lines (lower access time), whereas for high demand the higher running speed assigned to the rail modes becomes more important, outweighing the advantages of BRT. The order of the frequencies and number of lines is dominated by the operator cost, lower for BRT according to Table 2.

The average total cost functions on Fig. 4a decreases with patronage. It shows that, under this framework, public transport service provision presents economies of scale. This happens because the optimal response to a demand increase is to raise the frequency and number of lines less than proportionally to demand, which in turns diminishes average access, waiting, in-vehicle and operator costs, phenomenon known as the Mohring Effect, after Mohring (1972). In an earlier paper where we compared the performance of different transit modes over a single line (Tirachini, Hensher, & Jara Diaz, 2010), we found that due to the capacity constraint of a corridor (expressed through a maximum mode frequency), it is possible that after certain demand level, average total cost increases, producing diseconomies of scale. This is not observed on this model as when demand grows, the number of lines can be adjusted as well (apart from the frequency), keeping user cost – particularly crowding levels – low.

3.2.2. Case 2: land acquisition fixed input costs are excluded

When the land acquisition cost is disregarded, the only noticeable difference on total costs is that the BRT advantage is slightly enlarged until 3.2 million-pax/day, due primarily to the larger land cost necessary for HR compared to BRT and LR (Table A3.1). Comparing frequencies and number of lines with Case 1, a difference is also obtained in the magnitude of the figures, since when the land cost is not considered, it is relatively less expensive to provide more lines (relative to increase the frequency), therefore the model responds by increasing the number of lines (by 27 percent for BRT, 12 percent for LR and 13 percent for HR, in average), and decreasing the frequencies (by 19 percent for BRT, 9 percent for LR and 8 percent for HR, in average), compared to Case 1. The average occupation rates (expression 7) across all periods and demand levels are 34 percent for BRT, 32 percent for LR and 20 percent for HR.

3.2.3. *Case 2A: Case 2 without the crowding penalty factor*

When the crowding penalty is not considered, BRT emerges as the most cost-effective mode across all the demand range tested (Fig. 6a), followed by LR in second place, which is a different outcome compared to Case 2. This is explained by the fact that the crowding penalty factor increases the relative weight of in-vehicle time cost in the total cost function, which is assumed to be the rail advantage (due to larger running speeds); in other words, if crowding plays a role in the (dis)utility of travelling, it is more important to have faster modes than if it does not, even to the point of changing the order of the most cost-effective modes for large demands.

Because in-vehicle time cost has a lower weight in this case, operator cost is relatively more important than in Case 2, the optimal number of lines and frequencies are expected to decrease. Consequently, the average frequency drops by 26 percent for BRT, 21 percent for LR and 9 percent for HR (average values over the full demand range, see Figs. 5c and 6c), whereas the number of lines is less sensitive, decreasing only by 1 or 2 lines for BRT, while for the rail modes in some cases the number of lines does not change and in others there is only one line less (compare Figs. 5b and 6b). With this, the average occupancy rate per vehicle is increased across all periods as expected. In fact, in this scenario 65 percent of cases for BRT and 35 percent of cases for LR resulted in an optimal frequency that was not high enough to accommodate all demand, having to be increased to satisfy the capacity constraint (13). In this case the average occupancy rates across all periods are 49 percent for BRT, 43 percent for LR and 24 percent for HR, figures that are between 4 and 15 percent higher than when the crowding penalty is considered (Case 2).

3.2.4. *Case 2B: Case 2 with BRT and LR running speeds increased to find threshold speed*

At this stage it is worth noting that the resulting average total cost and consequent relative order of modes is highly dependent on the speed assigned to the modes (30, 35 and 40 km/h for BRT, LR and HR respectively). To analyse the effect of speed on the previous results, we introduce a variation of Case 2 by assuming that the BRT running speed is increased to 31 km/h and the LR speed lifts to 36 km/h (keeping the HR speed on 40 km/h), the average total cost for the three modes is as in Fig. 7a, which depicts BRT as the most cost-effective mode along almost the entire range of demand, with the three modes having the same average total cost when total demand is close to 3.5 million-pax/day. Therefore, we have found the threshold speed difference in the trade-off between the bus advantage on operator cost, access time and waiting time cost and the rail advantage on in-vehicle time: for the rail modes to provide the lowest total cost (for high levels of demand), they need to be able to operate at more than 5 km/h faster than BRT in the case of LR, and 9 km/h faster in the case of HR, in other words, the benefit of BRT from having a lower total operator cost is translated as an advantage of at least 5 and 9 km/h in running speed, in comparison to the rail based systems.

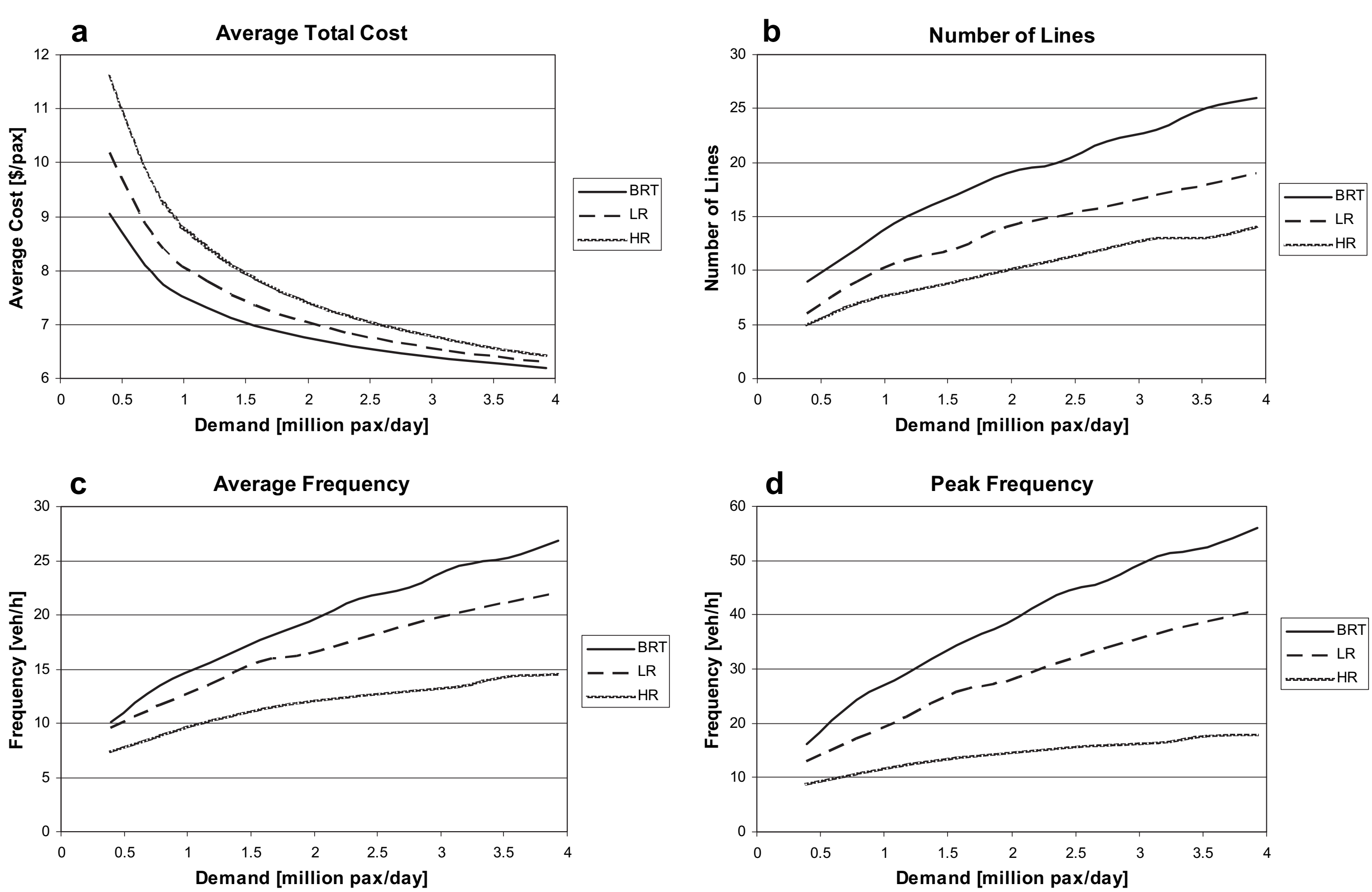

Fig. 6. Results Case 2A.

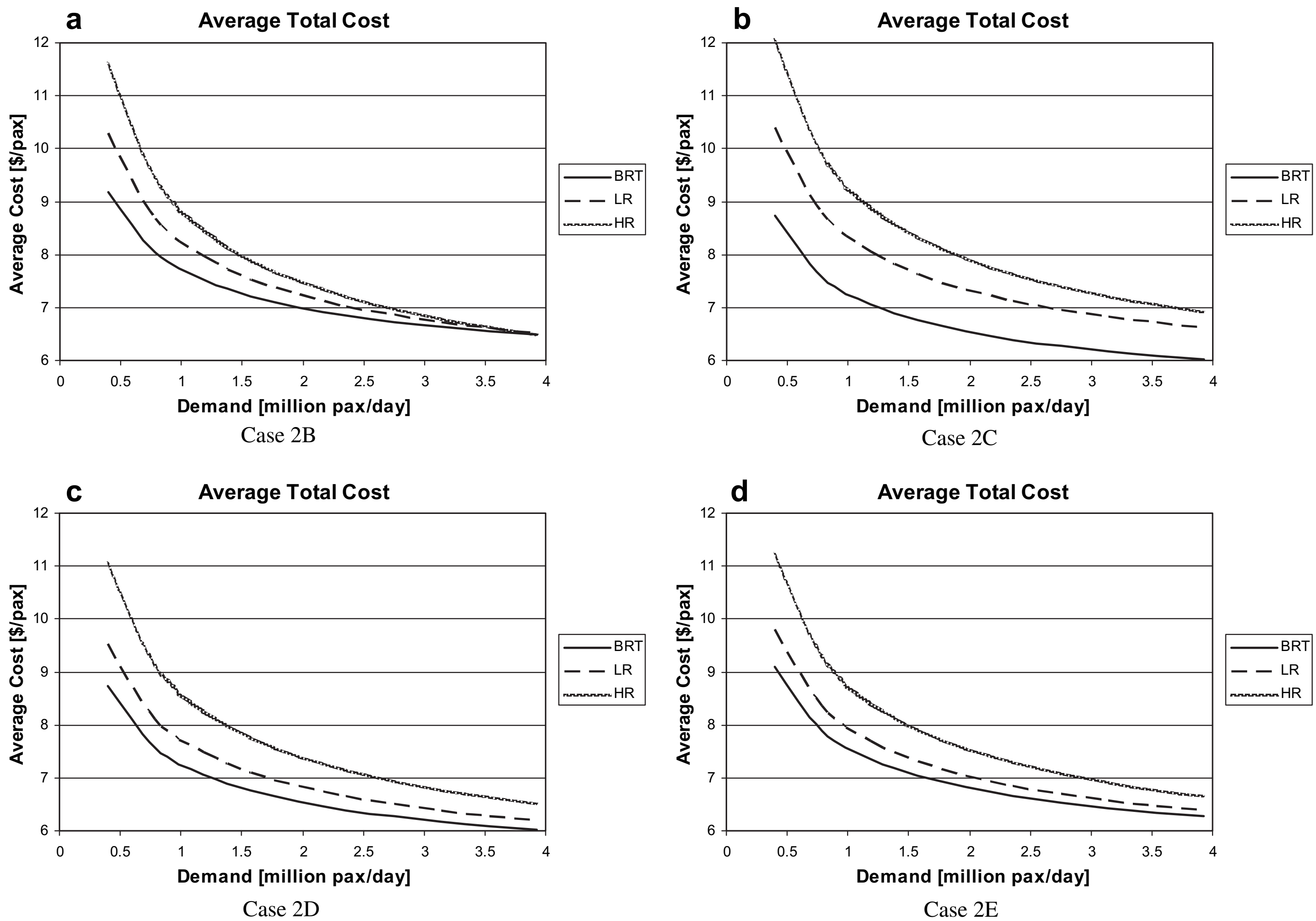

Fig. 7. Average total cost, Cases 2B to 2E.

3.2.5. *Case 2C: Case 2 with running speed equal for all modes*

If, however, we set the running speed and distance between stations equal across the three modes, the only advantage of LR and HR is removed, and therefore BRT provides the lower total cost by a significant margin, as depicted in Fig. 7b, due to its lower total operator cost and larger frequencies and number of lines.

3.2.6. *Case 2D: Case 2C + a decrease of 25 percent in the operator cost of rail modes*

A scenario with a lower rail operator cost is developed. All LR and HR operator cost components are assumed to drop by 25 percent. The results on total cost after this adjustment are given in Fig. 7c, where it can be observed that BRT prevails as the most cost-effective mode across the full demand range tested, i.e., even if the infrastructure, rolling stock and operating cost of rail modes are 25 percent less than the assumed values in Table 2, BRT still provides the lowest cost if it is able to operate at the same running speed.

3.2.7. *Case 2E: Case 2D + fixed dwell time at stations and BRT assumed less reliable than LR, which is less reliable than HR*

The reliability of running times can be different for bus and rail modes, depending on the operating conditions and degree of segregation of the transit system (for example, existence of an intersection shared with cars and other modes, which causes friction, and consequently, a lower predictability or running times). This issue is translated in the model as the inability of keeping regular headways, which has an impact on larger average waiting times and frequencies, through the factor ε. We assign $\varepsilon_{BRT} = 0.7$, $\varepsilon_{LR} = 0.6$ while ε_{HR} is kept at 0.5. On the other hand, we have assumed that the passenger dwell time at stations is fixed and independent of demand (situation more common in rail systems with uncrowded conditions). As in Case 2D, all modes have the same speed and distance between stations and rail operator cost is decreased by 25%. The result presented in Fig. 7d suggests that these two adjustments have little impact on total cost comparison, relative to Case 2D (Fig. 7c). This is because in most of the cases the resulting optimal frequency is high, over 5 veh/h, which means that the waiting time cost is relatively low relative to other components of the total cost function, then an increase on the parameter ε has little impact on the overall result. In addition, the time spent at stations boarding and alighting passengers is also little compared against the running time, then a change in the specification of that component of the travel time has, again, a small impact.

4. Conclusions

In the present paper we have developed a microeconomic model to compare the performance of Bus Rapid Transit, Light Rail and Heavy Rail, taking into account four elements: access time cost, waiting time cost, in-vehicle time cost and operator cost (including not only operating costs, but also land and infrastructure capital costs). The application of the model to a number of scenarios gives many important results, namely:

- BRT is the mode that provides greater frequencies and number of lines in all cases, which implies lower waiting and access time costs for users. This result is driven by the lower operator costs of the bus system, relative to the rail technologies.
- Out of the four cost components considered, BRT has the advantage on operator costs, access and waiting times; hence the only possible advantage of the rail is in the operational speed. In order to establish when trains outweigh the advantages of BRT, the threshold speed differences between rail modes and BRT can be calculated. Using the value of the parameters as in Case 2B, for LR (HR) to be more cost effective than BRT, it needs to run at least 5 (9) km/h faster, when the BRT speed is 31 km/h. If the difference in speeds is lower than this margin, BRT provides the lowest total cost for the entire range of demand.
- That said, if HR is able to provide a running speed considerably higher than BRT (for example, HR speed being 10 km/h faster than BRT as in Cases 1 and 2), there are situations where an HR outweighs BRT, but only for high levels of demand (e.g., over 3.2 million-pax/day on Case 2), a result that is consistent with the existant literature discussed in Section 1. On the other hand, the high rail capital cost makes it a very unattractive investment for low level of demand (e.g., below 2 million-pax/day) in any case.
- When the running speed is the same across all modes, BRT is the most cost-effective mode for all demand levels, even after assuming a 25 percent drop in the operator cost for rail modes.
- Accounting for crowding implies higher frequencies and number of lines and a lower occupancy rate for the three modes, as expected. The crowding penalty gives more weight to the in-vehicle time cost in the total cost function, making it more important to have faster modes.

A full comparison of transit technologies in a real context should consider other aspects as well, such as external and environmental costs, land use impacts, vehicle comfort, traffic disruption during the construction period and so on. Some of these elements have been argued as the advantages of trains over buses, in order to justify rail investments in the past, even when compared against buses providing a train-like service such as BRT in Curitiba, Bogotá and Brisbane (see Hensher & Golob, 2008). As noted by Smith (1973, p. 31), "If there are valid non-economic reasons for preferring a new railway, these reasons ought to be clearly identified and weighted against the advantages of the best bus alternative". In this paper we have presented a number of scenarios representative of plausible real situations, which suggest that if buses run at the same (or comparable) speed than trains, these "other reasons" must be able to outweigh the benefit of having more frequent services, better coverage and a lower total operator cost, provided by a high standard bus based system, in comparison to current LR and HR technologies. The challenge we face is convincing the political process to stop focussing on the ideological commitment to specific modal technologies and to base infrastructure and service provision decisions on value for money.

Acknowledgments

This research was partially funded by Fondecyt, Chile, grant 1080140, and the Institute for Complex Engineering Systems, grants ICM: P-05-004-F and CONICYT: FBO16.

Appendix A1. Access distance equation

The total access distance *to the line* is calculated using polar coordinates, as users are assumed to walk in concentric circles. Recalling that an area element in polar coordinates is $dA = r\,dr\,d\varphi$, the total walk distance on half of the catchment area (Fig. 3) is:

$$\bar{d}_a = \int_0^{\theta/2}\int_0^{\alpha L/2} r\varphi\, m_1\, r\,dr\,d\varphi + \int_0^{\theta/2}\int_{\alpha L/2}^{L/2} r\varphi\, m_2\, r\,dr\,d\varphi$$
$$= \frac{\theta^2 L^3}{3\cdot 2^6}\left[\alpha^3 m_1 + \left(1-\alpha^3\right) m_2\right]$$

If we sum all over the circle, the total walking distance is $d_a = 4n\,\bar{d}_a$, expression equal to equation (2), after using the definitions of y_1 and y_2 in (1).

Appendix A2. Crowding penalty factor estimation

A factor $\delta_t(\theta_t) \geq 1$ is used as a mark up in the calculation of the in-vehicle time cost. When the occupancy rate is lower than the seating capacity, the crowding disutility is not active, i.e. $\delta_t(\theta_t) = 1$, but as long as passengers have to stand, the perception of the trip changes, and the factor δ_t increases in line with the occupation rate until reaching a value of around 2 for urban trips, when a vehicle is full (Whelan & Crockett, 2009).

Table A2.1. Piecewise crowding factor.[a]

Occupancy rate θ_t	Crowding factor δ_t
0.1	1
0.2	1
0.3	1
0.4	1.1
0.5	1.2
0.6	1.3
0.7	1.4

[a] The crowding factors in Table A3.1 have to reflect average conditions; for instance, when the average occupancy is 70%, in some segments the vehicle circulates full (crowding penalty around 2) but in other segments it circulates with a occupancy under seating capacity, therefore, a value δ_t between 1 and 2 is chosen.

A simple and intuitive functional form for the crowding penalty is a piecewise linear function.[9] As in this paper we work with average occupancy (θ_t) instead of maximum occupancy, a maximum average occupancy of $\theta_t = 0.7$ was assumed by design, in order to have room to accommodate the maximum load. On the other hand, it is assumed that the crowding effect is active when θ_t reaches a value of 0.3, and from that point it increases linearly until the maximum value of 0.7, as shown in Table A2.1

Finally, a quadratic function is chosen to interpolate the piecewise linear function of Table A2.1, in order to avoid practical problems in the implementation of the solution algorithm. The estimated parameters are as follows:

$$\delta_t(\theta_t) = 1 - 0.2381\,\theta_t + 1.1905\,\theta_t^2 \quad \left(R^2 = 0.985\right)$$

Numerical testing showed that the solution is not dependent on the quadratic form adopted, as the results on optimal frequencies, number of lines and occupancy rates do not vary significantly

[9] Using a Stated Preference study, Whelan and Crockett (2009) estimate piecewise linear, power, exponential and Gompertz functions for the crowding factor, finding that the more complex functions do not yield results significantly different from the piecewise linear.

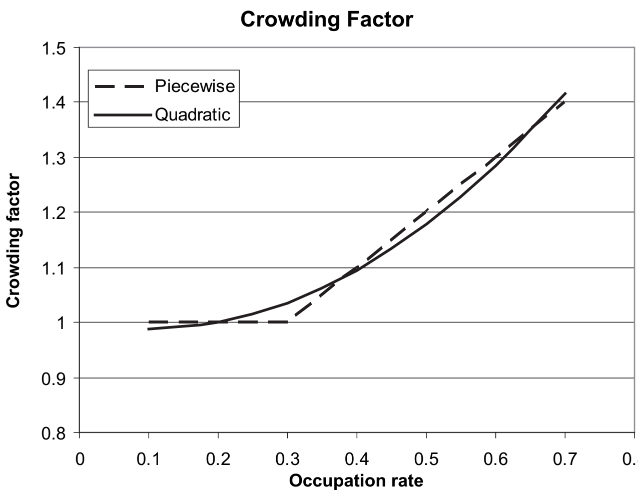

Fig. A2.1: Crowding factor as a function of the average vehicle occupancy.

by adopting the quadratic form or the second piece of the linear function ($\delta_t(\theta_t) = 0.7 + \theta_t$) shown in Fig. A2.1.

Appendix A3. Operator cost items

The figures presented are derived from values recommended in ATC (2006), updated to year 2007. The discount rate used is 7%. Fixed costs are based on annuity calculations. All values in Australian dollars (Euro 1 ≈ AUD 1.7).

I. Fixed costs
(1) Infrastructure capital cost: It corresponds to railways and dedicated busways. In the model it is expressed in $/day-line, derived from unit values in $/km, as in Table A3.1.
(2) Land: Market value of required land. It is highly variable depending on the location, transaction costs, etc. 9 million $/hectare is the figure used to estimate the land costs per mode, in the function of the area of the infrastructure facilities, as in Table A3.1.
(3) Rolling stock capital costs: an electric-power train three-car single deck unit is considered for HR (based on trains in Melbourne and Brisbane), while BRT are 18 m long articulated buses (ATC, 2006, p. 61). The original LR cost in ATC (2006) is 4.5 million $/veh, corresponding to five-module articulated trams which cannot be coupled (as the Combino vehicles in Melbourne); however, this value was corrected because it is too high compared to values suggested in other sources (Bruun, 2005; Evans, 2005). Based on the value recommended by Evans (2005), 2.5 million $/veh (2004 US dollars), the LR cost used is calculated as 3.49 million $/veh (2007 Australian dollars). A residual value of 5% of the initial price is considered for the rolling stock.

II. Operating costs
(4) Crew costs: wages and direct on-costs for on-vehicle staff (drivers, guards, etc.).
(5) Direct operating costs: these cover vehicle fuel and power and vehicle maintenance.
(6) Infrastructure operations and maintenance costs: operation and maintenance of track, right-of-way, signalling, communications and so on. For simplicity, it is expressed in $/day-line, although in practice some costs vary with usage.
(7) Overhead operating costs: covers all other operating costs not included in the other three categories, such as scheduling, rostering, supervision, depot-related costs, head office cots, human resources and so on. Although overhead costs may be expressed in various ways, they are considered as a percentage of all other operating costs (included in the values used for 4, 5, and 6), as in Table A3.2.

With this, the unit cost c_0 is computed considering (1), (2) and/or (6) depending on the case, c_1 as (3), c_2 as (4) and c_3 as (5).

Table A3.1: Fixed operator cost calculation.

Mode	Infrastructure cost			Land cost	
	Cost [million$/km]	Asset life [years]	Cost (1) [$/line-day]	Wide [m]	Value (2) [$/line-day]
HR	35	100	201,602	15	77,671
LR	20	100	115,201	10	51,781
BRT	10	50	59,556	10	51,781
Mode	Rolling stock cost				
	Capacity [pax/veh]	Cost [million$/veh]	Asset life [years]	Cost (3) [$/veh-day]	
HR	750	8.32	35	1721	
LR	190	3.49	35	701	
BRT	101	0.62	20	158	

Source: Author's estimates based on information in ATC (2006) and Evans (2005).

Table A3.2: Operating cost.

Mode	Crew (4) [$/veh-h]	Operating (5) [$/veh-km]	Maintenance (6) [$/day-line]	Overhead (operating)
HR	130	3.3	22,413	14%
LR	73	1.8	13,057	17.5%
BRT	42	1.4	3579	21%

Source: Author's estimates based on information in ATC (2006)

References

Allport, R. J. (1981). The costing of bus, light rail transit and metro public transport systems. *Traffic Engineering and Control, 22*, 633–639.

ATC. (2006). *National guidelines for transport system management in Australia*. In *Urban transport, Vol. 4*. Australian Transport Council. http://www.atcouncil.gov.au/documents/NGTSM.aspx Retrieved on January 2009 from.

Boyd, J. H., Asher, N. J., & Wetzler, E. S. (1978). Nontechnological innovation in urban transit: a comparison of some alternative. *Journal of Urban Economics, 5*, 1–20.

Bruun, E. (2005). Bus rapid transit and light rail: comparing operating costs with a parametric cost model. *Transportation Research Record, 1927*, 11–21.

Chang, S. K., & Schonfeld, P. M. (1991). Multiple period optimization of bus transit systems. *Transportation Research Part B: Methodological, 25*, 453–478.

Desaulniers, G., & Hickman, M. D. (2007). *Public transit. Handbooks in operations research and management science*. In *Transportation. Barnhart and Laporte, Vol. 14*. Elsevier. pp.69–127.

Dewees, D. N. (1976). Urban express bus and railroad performance: some toronto simulations. *Journal of Transport Economic and Policy, 10*, 16–25.

Evans, J. (2005). Capacity and cost comparisons of rapid transit modes. In *Institute of Transportation Engineers annual meeting, Melbourne, Australia*.

Hensher, D. A. (2007). Sustainable public transport systems: moving towards a value for money and network-based approach and away from blind commitment. *Transport Policy, 14*, 98–102.

Hensher, D. A. (2008). Climate change, enhanced greenhouse gas emissions and passenger transport – what can we do to make a difference? *Transportation Research Part D: Transport and Environment, 13*(2), 95–111.

Hensher, D. A., & Golob, T. F. (2008). Bus rapid transit systems – a comparative assessment. *Transportation, 35*(4), 501–518.

Jara-Díaz, S. R., & Gschwender, A. (2003). Towards a general microeconomic model for the operation of public transport. *Transport Reviews, 23*(4), 453–469.

Kraus, M. (1991). Discomfort externalities and marginal cost transit fares. *Journal of Urban Economics, 29*(2), 249–259.

Litman, T. (2007). Evaluating rail transit benefits: a comment. *Transport Policy, 14*, 98–102.

Maunsell, F., & Macdonald, M. (2007). Rail overcrowding, reliability and frequency. Report for Centro, the West Midlands Passenger Transport Executive.

Meyer, J. R., Kain, J. F., & Wohl, M. (1965). *The urban transportation problem*. Cambridge, Massachusetts: Harvard University Press.

Pepper, J., Spitz, G., & Adler, T. (2003). Customer perspectives on multilevel coaches for increasing rail system capacity. *Transportation Research Record, 1838*, 19–29.

Polydoropoulou, A., & Ben-Akiva, M. (2001). Combined revealed and stated preference nested logit access and mode choice model for multiple mass transit technologies. *Transportation Research Record, 1771*, 38–45.

RAND. (2006). Benefits of punctuality improvements for the Paris suburban railway network. http://www.rand.org/pubs/technical_reports/TR271/ (Technical report). Retrieved on January 2009 from.

Smith, E. (1973). An economic comparison of urban railways and express bus services. *Journal of Transport Economic and Policy, 7*, 20–31.

Tirachini, A., Hensher, D. A., & Jara Diaz, S. (2010). Restating modal investment priority with an improved model for public transport analysis. *Transportation Research E, 46*(6), 1148–1168.

Whelan, G., & Crockett, J. (2009). An investigation of the willingness to pay to reduce rail overcrowding. In *International conference on choice modelling, Harrogate, England, April*.

Wright, L., & Hook, W. (2007). *Bus rapid transit planning guide* (3rd ed.). New York: Institute for Transportation and Development Policy.