Australasian Transport Research Forum 2013 Proceedings
2 - 4 October 2013, Brisbane, Australia
Publication website: http://www.patrec.org/atrf.aspx

How does the efficiency performance of Sydney CityRail compare with international urban rail systems

Chi-Hong (Patrick) Tsai[a], Corinne Mulley[b]

[a,b] *Institute of Transport and Logistics Studies, Business School, The University of Sydney, 2006, Australia*
[a] *E-mail for correspondence: chi-hong.tsai@sydney.edu.au*

Abstract

The motivation for this paper comes from the fare determination judgement by Independent Pricing and Regulatory Tribunal (IPART) for CityRail which requires a certain level of service quality and quantity whilst improving the operation efficiency. There is limited information about the historical efficiency performance of CityRail and this gap is the focus of this paper.

This paper compares the operational performance of Sydney CityRail with 11 international urban rail systems using a Data Envelopment Analysis approach. The operating performance is examined using cost-efficiency and cost-effectiveness measures to understand the extent service inputs are efficiently used to generate service outputs, in terms of car-km operated and passengers carried. This research finds the operation of CityRail is efficient in terms of car-km operated, but the cost-effectiveness score is the lowest of the 12 systems being compared. Around 3 percent of employees as well as 10 percent of operating cost are identified as redundant service inputs in CityRail suggesting possible strategies for CityRail to improve its service effectiveness.

1. INTRODUCTION

The extent to which a transport system is considered to be efficient or service inputs could be saved in the production of the current level of outputs is of concern to both transport operators and regulators. Understanding the operational performance provides information on which to understand potential improvement in service quality and financial plans, as well as fare determination. The measurement of efficiency performance is often undertaken by comparing multiple transport systems, and such evaluation is a way to investigate how a transport operator performs in terms of the service outputs with respect to the use of service inputs.

The comparison of the efficiency performance is particularly important for urban rail systems because they are usually monopolies in their local markets. A monopoly market makes it more difficult for the transport regulators in measuring and monitoring the efficiency performance of the local rail operator. For example, the Independent Pricing and Regulatory Tribunal (IPART), the fare regulator of Sydney CityRail, requires information about CityRail’s efficiency performance so as to estimate its efficient operating cost as part of the fare determination, but such a comparison was only possible with a limited number of operators in Australia because of data limitations. (IPART, 2008).

Given the close relationship between efficiency performance and fare determination, a rigorous examination of performance across urban rail systems internationally can reveal how well Sydney’s CityRail is performing on the international stage. Despite having different spatial settings, rail based technology and rail regulation (particularly in relation to safety) is very similar across national and international boundaries which in turn facilitates effective comparisons.

This paper evaluates the efficiency performance of Sydney’s CityRail with another 11 international urban rail systems in Asia, Australia, Europe and North America using data from 2009 to 2011. Section 2 reviews the literature on the efficiency measurement and its related methodology. Section 3 introduces the DEA approach and the data of this paper with a preliminary analysis using a Partial Factor Productivity (PFP) methodology. Section 4 presents the research outcomes and Section 5 concludes this paper.

2. LITERATURE REVIEW

There is a lack of international evidence focussing on urban rail systems although some literature on the evaluation of public transport systems performance exists, most previous studies have been conducted within the context of Europe (Gathon and Pestieau, 1995, Cantos and Maudos, 2001, Pina and Torres, 2001, Mulley, 2003, Merkert et al., 2010, Karlaftis and Tsamboulas, 2012) and North America (Benjamin and Obeng, 1990, Chu et al., 1992, Karlaftis and McCarthy, 1997, Viton, 1997, Karlaftis, 2004). Sampaio et al. (2008) was one of the few studies that compared efficiency performance across continents (Brazilian and European metro systems). Anderson and Harris (2007) analysed the performance of 22 worldwide metro systems as part of the COMET and NOVA benchmarking groups but the focus of the evaluation was on passenger alighting and boarding rates with respect to the operating characteristics of the metro systems such as frequency and stop time. However, there is no Australian evidence in the literature nor an evidence base which compares the efficiency performance of Australian against other international rail operators. This paper addresses this research gap.

The definition of efficiency for public transport has been widely discussed since the 1980s. Fielding et al. (1985) proposed a benchmarking framework for public transport systems as illustrated in Figure 1. This framework distinguishes between efficiency and effectiveness in evaluating the performance of a public transport operator. Efficiency in this framework refers to the total service outputs usually measured by car-km travelled or car-hour operated with respect to service inputs (labour, fuel consumption, or operating cost), whereas effectiveness represents the service consumption by passengers such as number of passengers, passenger-km against service inputs. The ratio of service consumption to service outputs is defined as service-effectiveness. This framework has been used in related research (Chu et al., 1992, Viton, 1997, Karlaftis and Tsamboulas, 2012). The distinction between efficiency and effectiveness highlights the different aspects of performance evaluation from the operator's and consumer's perspectives.

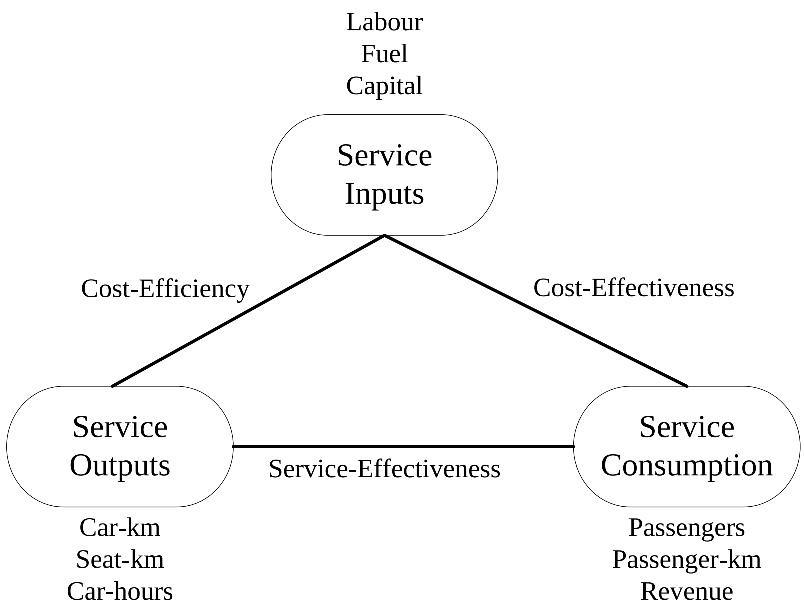

Figure 1. A framework for benchmarking the performance of public transport systems
Source: reproduced from Fielding et al. (1985)

A number of methods used to evaluate efficiency and effectiveness performance have been developed in the literature (Oum et al., 1999, Merkert et al., 2010, Karlaftis and Tsamboulas, 2012). The simplest approach is the Partial Factor Productivity (PFP) method, which measures the ratio of a public transport system's output to a single input. The advantage of this approach is that it is easy to implement and understand and is thus often favoured by policy-makers. However, the PFP measures only one input against one output and, as a result, multiple Key Performance Indicators (KPIs) are produced without a rationale for combining into a single overall indicator. What is meant by overall performance is open to discussion if the PFP approach is used. A second method is the Total Factor Productivity (TFP) methodology as employed in some studies (Benjamin and Obeng, 1990, Karlaftis and McCarthy, 1997). The TFP approach generates a single index based on the ratio of an aggregate output and an aggregate input in quantities. However, Oum et al. (1999) have suggested that aggregation problems may occur when producing a single index from multiple inputs or outputs.

Other methodologies used in the literature are Stochastic Frontier Analysis (SFA) (Gathon and Pestieau, 1995, Cantos and Maudos, 2001, Karlaftis and Tsamboulas, 2012) and Data Envelopment Analysis (DEA) (Chu et al., 1992, Viton, 1997, Pina and Torres, 2001, Karlaftis, 2004, Sampaio et al., 2008, Merkert et al., 2010, Karlaftis and Tsamboulas, 2012). The SFA uses an econometric model to estimate a firm's productivity based on its service inputs. Traditional cost or production functions are typically used to estimate the frontier of a firm's productivity and thus to identify the relative efficiency amongst multiple firms in the dataset. This approach is data-demanding and ideally panel data are required to control for the unobserved heterogeneity (Karlaftis and Tsamboulas, 2012).

DEA has been commonly applied in the transport literature. It was introduced by Farrell (1957) and further developed by Charnes et al. (1978). DEA is a non-parametric approach using linear programming to identify the linear production frontier and an efficiency score for each firm in the sample. This approach has been widely employed because of its flexibility in selecting multiple inputs and outputs. It also allows for the assumption of variable returns to scale (Banker et al., 1984) which is one of the key properties of transport industry (Braeutigam, 1984, Karlaftis, 2004). Moreover, when historical data are available, efficiency may change over time as a result of institutional reform or technical change. DEA can identify these changes for an operator leading to an increase in understanding as to whether the operating strategy has improved efficiency performance.

3. DATA ENVELOPMENT ANALYSIS

3.1 Approach

The concept of DEA is illustrated in Figure 2. Each of the firms A, B, C, D and E produces a vector of outputs from a vector of service inputs. The curve ABCD is called the production frontier where firms A, B, C and D have produced the maximum outputs using the current level of inputs, so they are considered "technically efficient". Firm E, however, is technically inefficient because it could produce a higher level of output at $E_{V'}$ rather than the current output level S based on the existing input level Q, or it could reduce its input from Q to T to produce it current output level S. This technical inefficiency is measured by SE_V/SE.

Figure 2 also explains the concept of scale efficiency. A firm is "scale efficient" when it is at the optimal scale with constant returns to scale, that it, doubling the service inputs is expected to double the service outputs. When a firm is too small, doubling the service inputs will generate more than a doubling of the service outputs (increasing returns to scale or economies of scale). In contrast, a firm may show decreasing returns to scale (diseconomies of scale) when the scale of the firm is too big, and doubling the service inputs results in less than a doubling of outputs because the inputs cannot be efficiently used to increase outputs. A firm is "scale inefficient" when it is either increasing returns to scale or decreasing returns to scale.

In Figure 2, constant returns to scale is shown by the straight line OM from the origin, where the output is proportional to the service input. The curve ABCD shows variable returns to scale, where firm B is the only one that reaches a scale efficiency which represents the optimal scale within this sample. Firm A, C and D are scale inefficient although they are all technically efficient on the production frontier curve. For example, the scale inefficiency of firm A is given by the ratio of RA_c to RA, and this firm would approach optimal scale if it increased its size. In contrast, firm C and D are too big in scale and it need to decrease their size to approach the optimal scale. Firm E is neither technically efficient nor scale efficient, and its scale inefficiency is measured by the ratio of SE_c/SE_v.

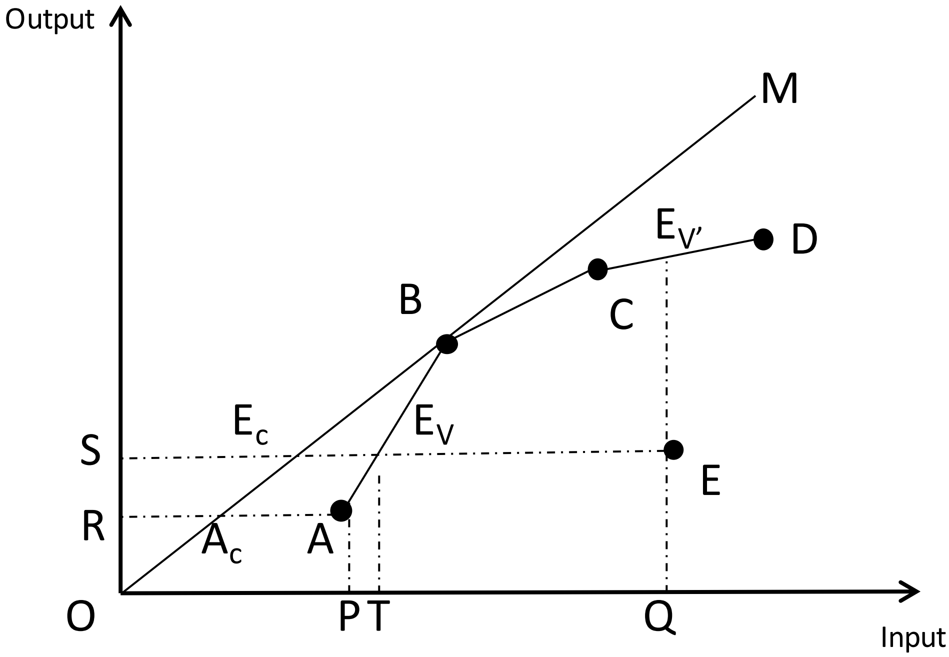

Figure 2. The Production Frontier and Efficiency Measurement

As reviewed above, DEA is a well-developed methodology for evaluating efficiency performance of multiple firms based on their service inputs and outputs in the transport sector. It is processed through a sequence of linear-programming solutions. Assuming a firm's objective is to minimise service inputs for a given level of outputs and there is no scale efficiency (i.e. constant returns to scale), the linear program can be presented as in Equation (1):

Minimise θ_n with resect to $w_{1,} ..., w_N$,

Subject to:

$$\begin{aligned}
&\sum_{j=1}^{N} w_j y_{ij} - y_{in} \geq 0 && i = 1, ..., I \\
&\sum_{j=1}^{N} w_j x_{ij} - \theta_n x_{kn} \leq 0 && k = 1, ..., K \\
&w_j \geq 0 && j = 1, ..., N
\end{aligned} \tag{1}$$

where θ_n is the efficiency score for the n_{th} firm subject to the constraints listed above. The linear programming problem is solved by minimising θ_n from the N firms in the sample producing I different outputs using K different inputs. y_{in} and x_{kn} are the total amount of outputs and inputs for firm n. w_j is the weight applied across different firms.

The linear-programming problem specified in Equation (1) indicates that the efficiency score of a firm is minimised subject to the constraints. The weighted outputs of all firms in the sample must be more than each of the output produced by any single firm (first constraint), and the weighted inputs of all firms must not exceed the input for any other single firm (second constraint). The third constraint limits the weights to being non-negative. The efficiency score represents the smallest proportion of inputs that a firm can use to produce its existing level of outputs. This is a relative performance score and a score of one means that the firm has reached "technical efficiency".

An important advantage of the DEA methodology is the identification of scale efficiency. For public transport systems, the returns to scale is usually considered to be variable rather than constant, that is, the ratio of the service outputs to inputs is not constant but instead it varies with the size of the firm. As introduced in Banker (1984), DEA can incorporate variable returns to scale by introducing a convex restriction as a constraint:

$$\text{Minimise} \quad \theta_n \text{ with resect to } w_{1,}..., w_N,$$

Subject to:

$$\begin{aligned}
&\sum_{j=1}^{N} w_j y_{ij} - y_{in} \geq 0 && i = 1,..., I \\
&\sum_{j=1}^{N} w_j x_{ij} - \theta_n x_{kn} \leq 0 && k = 1,..., K \\
&w_j \geq 0 && j = 1,..., N \\
&\sum_{j=1}^{N} w_j = 1
\end{aligned} \tag{2}$$

where the weights are restricted to a sum of one when allowing for variable returns to scale.

3.2 Data

This paper analyses 12 urban rail systems across Asia, Australia, Europe, and North America. The urban rail systems included in the comparison are: Barcelona, Hong Kong, Kaohsiung, Lisbon, London, Madrid, Montreal, Singapore, Sydney, Taipei, Toronto and Washington DC.

The data are collected from publicly available financial reports or annual reports between 2009 and 2011 from the operators' official websites. The three years of data allow for an investigation of changes in efficiency performance due to possible technical changes or organisational reforms during the study period. The service outputs are categorised into efficiency measures and effectiveness measures as suggested by Fielding et al. (1985). The selection of the service inputs for this research follows the literature (Chu et al., 1992, Viton, 1997, Karlaftis, 2004, Karlaftis and Tsamboulas, 2012), which has generally suggested to use car-km travelled as the efficiency measure and the number of passengers as the effectiveness measure, where a car refers to the carriage (or coach) of a train.

The service inputs for this paper are defined by labour, rolling stock, and operating cost. Labour is measured by the total number of employees of the system operator. All employees regardless part-time or full-time appointments are included to make this measure consistent across the 12 operators. Rolling stock is defined by the total number of cars owned by the operator. Operating cost is the cost required for system operation. Asset or capital-related cost such as depreciation and amortisation is not included in the operating cost. The operating costs are converted to Australian Dollars based on 2011 conversion rate. In addition, a Purchasing Power Parity (PPP) conversion factor is used to standardise the operating cost in each country to adjust the difference in consumer price and the level of living expenses. For example, the labour price, which is a component of operating cost, would be cheaper in some Asian cities than Australia, so it is necessary to standardise the consumer price to make the cost-efficiency and cost-effectiveness comparable across countries. The PPP conversion factor is acquired from World Bank[1] and adjusted based on the Australia index (PPP index=1 for Australia).

The descriptive statistics of all the variables are summarised in Table 1. The total number of observations is 34 due to data unavailability of Sydney in 2010 and Lisbon in 2011. Although there appears to be a substantial difference in scale for the 12 urban rail systems (shown by the large variation in car-km and patronage) the efficiency and effectiveness performance are compared by an overall ratio of service outputs to inputs. For example, Sydney CityRail is a commuter rail system which has a larger scale of network and thus the operation requires more employees and larger fleet size, but it is also expected to generate more service outputs in terms of car-km operated and passengers carried, given the larger scale of inputs. Therefore, the efficiency and effectiveness performance across the 12 urban rail systems can still be compared on the same grounds. The DEA approach only focuses on the internal operational performance which is the aim of this study although it is recognised that other external factors such as population density and financial subsidy might have a significant impact on service outputs. An investigation of the impact of these external factors requires a post-DEA regression analysis which is outside the scope of this paper.

Table 1. Descriptive Statistics of the Dataset

Variable	Unit	Obs	Mean	S.D.	Min	Max
Car-km	km (million)	34	151	134	12	507
Patronage	passengers (million)	34	501	370	43	1,366
Employee	persons	34	7,277	5,820	1,279	19,064
Car	cars	34	1,323	1,101	126	4,243
Operating cost	Australian dollars (million)	34	1,330	1,267	100	4,863

[1] Available at http://data.worldbank.org/indicator/PA.NUS.PPPC.RF/

The other limitation of the data is that some system operators in this dataset are multi-modal operators such as Montreal, Toronto, Hong Kong and Washington DC, and the financial data and the number of employees is not segmented by sector in the published reports. This paper first explored possible methods to allocate the number of employees and operating cost to the urban rail sector by investigating some of the multi-modal operators with segmented data publicly available. Data from Singapore, London and Barcelona are summarised in Table 2. Using operated car-km as a base, it can be seen that the ratio of rail employees to total employees (employee weights) are lower than the car-km weight in Singapore and Barcelona, but substantially larger in London. On the other hand, the cost weights are both lower than the car-km weight in Barcelona and London. It is not clear how the number of employees and operating cost should be allocated to the rail sector only for those multi-modal operators without segmented data and further exploration of weighting schemes is required but outside the scope of this paper.

Table 2. Segmented Operating Data of Singapore, Barcelona and London

City	Year	Rail Car-km (million)	Total Car-km (million)	**Car-km weight**	Rail Employees	Total Employees	**Employee weight**	Rail Cost (million)	Total Cost	**Cost weight**
Singapore	2011	100	180	**0.56**	3,217	6,565	**0.49**	n/a	675	**n/a**
	2010	92	170	**0.54**	3,259	6,651	**0.49**	n/a	624	**n/a**
	2009	85	162	**0.53**	3,051	6,226	**0.49**	n/a	607	**n/a**
Barcelona	2011	901	133	**0.68**	3,723	8,015	**0.46**	404	688	**0.59**
	2010	88	130	**0.67**	3,764	7,370	**0.51**	376	644	**0.58**
	2009	79	128	**0.62**	3,576	7,152	**0.50**	314	589	**0.53**
London	2011	507	996	**0.51**	15,585	18,839	**0.83**	2,178	4,911	**0.44**
	2010	476	962	**0.49**	17,239	20,822	**0.83**	2,050	4,842	**0.42**
	2009	486	969	**0.50**	17,944	21,619	**0.83**	2,301	4,908	**0.47**

3.3 Preliminary Analysis

A preliminary analysis is conducted to examine the data quality and to identify possible outliers in the sample. This section of analysis uses the PFP method which evaluates some KPIs based on the service inputs and outputs from the dataset. The three dimensions of performance measurement as proposed by Fielding et al. (1985) are presented from Figure 3 to Figure 5.

Figure 3 shows the cost-efficiency in terms of operating cost per car-km travelled. Note that the costs are adjusted to Australian Dollars and standardised by the PPP index, so cost-efficiency can be directly compared across countries. Of the 12 urban rail systems examined, Washington DC is the least efficient system, in which the operating cost is around 12 to 14 dollars per car-km. The most efficient system is Singapore with around 5 dollars per car-km. The historical trend for the 12 urban rail systems does not show noticeable change. The operating cost per car-km in Barcelona, Hong Kong, Montreal, Toronto and Washington DC have been slightly increased from 2009 to 2011, whereas Singapore and Taipei demonstrated an efficiency improvement given their decreasing operating costs per car-km.

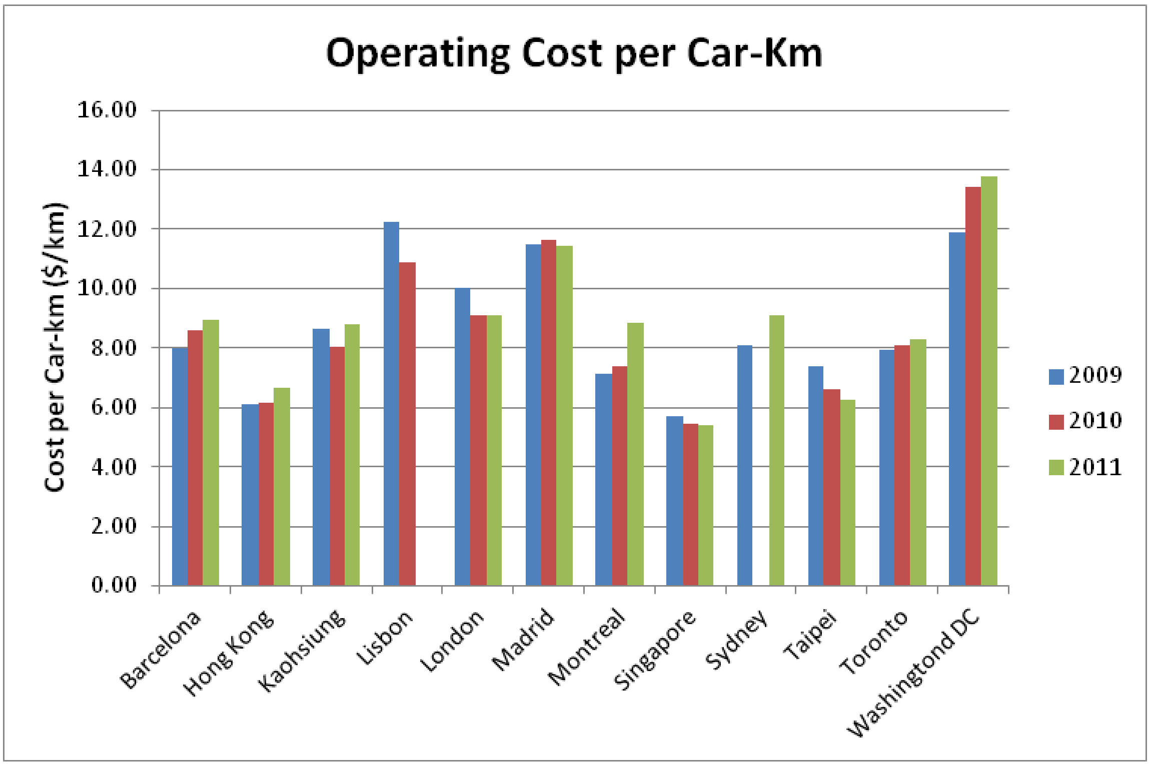

Figure 3. Cost-Efficiency of 12 Urban Rail Systems

The cost-effectiveness measure is presented in Figure 4, where Sydney stands out as the least cost-effective system with around 8 to 9 dollars of operating cost per passenger carried. The Asian urban rail systems in Singapore, Taipei, Hong Kong, perform better than other systems in terms of cost-effectiveness, with around 1 dollar of operating cost per passenger. Figure 4 also shows that there is more variation in the cost-effectiveness measure across the 12 systems than the cost-efficiency measure in Figure 3. This is because the number of passengers is less influenced by operating cost. Instead, other factors might affect the patronage such as land use density and public transport fares. In terms of the historical trend, the operating cost per passenger in Sydney and Washington DC have been increasing since 2009, whereas London has a substantial reduction in operating cost per passenger after 2009.

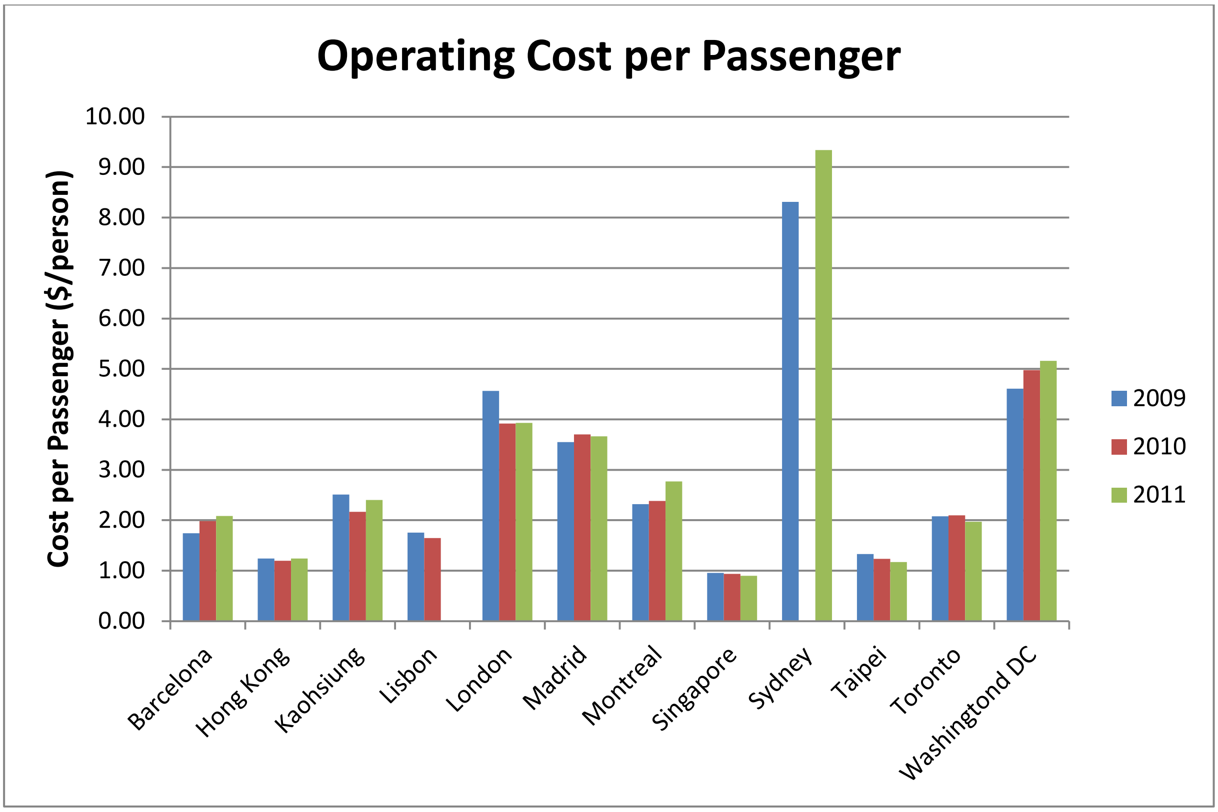

Figure 4. Cost-Effectiveness of 12 Urban Rail Systems

The third performance measure is service-effectiveness which is the ratio of number of passengers to the car-km travelled. As shown in Figure 5, the urban rail system of Lisbon and those in Asia generally have more passengers per car-km. Lisbon has a markedly high service-effectiveness as a result of its' small network size relative to its' patronage leading to the good performance in service-effectiveness. As compared to cost-efficiency and cost-effectiveness, Figure 5 shows service-effectiveness to have less variation for each of the systems over the three years probably as a result of outputs (passengers and car-km) being less influenced by potential external factors such as economy or price inflation than the input variables.

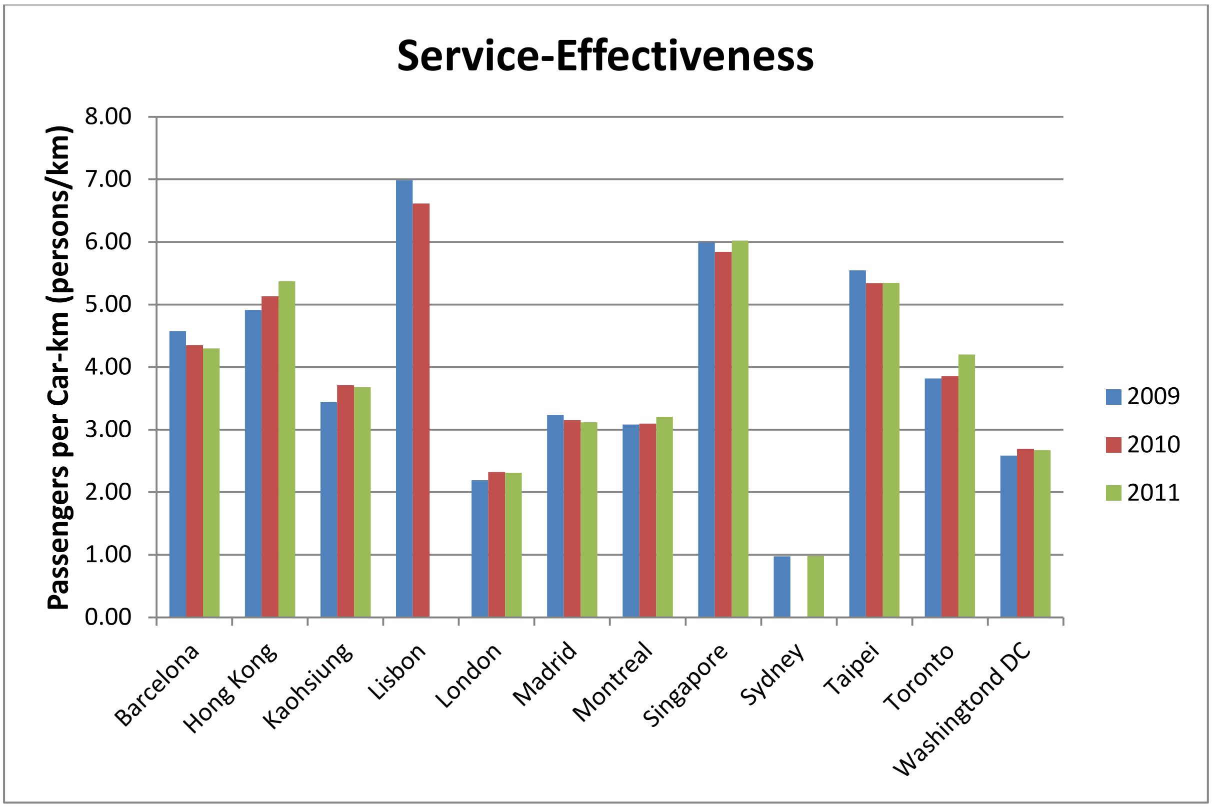

Figure 5. Service-Effectiveness of 12 Urban Rail Systems

In general, this section of analysis does not identify noticeable outliers in the dataset, other than Sydney's CityRail for cost effectiveness and does not find substantial historical change in the efficiency measures. However, as reviewed in Section 2, the limitation of this PFP approach is that there is no overall measure to evaluate multiple inputs and outputs. From a comparison of the three measures discussed above, it is inconclusive which system performs better than the others overall and provides a strong incentive to undertake analysis using DEA. The next section presents the results from the DEA methodology which is able to compare the various systems based on multiple service inputs and outputs.

4. DEA RESULTS

4.1 Efficiency Performance

The results of DEA are summarised in Table 3, in which the scores for efficiency and effectiveness of the 12 urban rail systems are identified. The efficiency score is the DEA which uses car-km as a single output whereas the effectiveness is analysed based on the number of passengers as a single output with respect to inputs of the number of employees, the number of cars, and operating cost.

Table 3. The Results of Data Envelopment Analysis

	Efficiency Score				Effectiveness Score			
	2009	2010	2011	AVG	2009	2010	2011	AVG
Barcelona	0.77	0.78	0.81	0.79	0.65	0.64	0.65	0.64
Hong Kong	1.00	1.00	0.96	0.99	0.95	1.00	1.00	0.98
Kaohsiung	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Lisbon	0.97	0.97	n/a	0.97	1.00	0.99	n/a	0.99
London	0.92	0.98	1.00	0.97	0.71	0.79	0.96	0.82
Madrid	0.83	0.82	0.83	0.82	0.55	0.48	0.50	0.51
Montreal	0.77	0.75	0.72	0.75	0.47	0.46	0.46	0.46
Singapore	0.96	1.00	1.00	0.98	0.96	0.98	1.00	0.98
Sydney	1.00	n/a	0.99	1.00	0.22	n/a	0.23	0.23
Taipei	0.75	0.82	0.88	0.82	0.71	0.74	0.77	0.74
Toronto	0.76	0.76	0.73	0.75	0.55	0.55	0.57	0.56
Washington DC	0.67	0.64	0.65	0.65	0.35	0.34	0.35	0.35

Looking at the efficiency scores in Table 3, Kaohsiung and Sydney each have an average score of unity which suggests that these two systems are on the technical efficiency frontier in terms of the car-km produced from the quantity of service inputs. It is interesting to note that, of the 12 systems being compared, Kaohsiung has the smallest system (39.12 km) and Sydney has the largest system (2,224 km) in terms of network size. Showing technical efficiency in terms of car-km operated with respect to the three input variables indicates that the network size of the urban rail systems is properly controlled by the DEA process and the efficiency score is not being deterministically influenced by system scale. Washington DC has the lowest average score of efficiency at 0.65, indicating that the service inputs could be reduced by 35 percent and still generate the current level of car-km. Other systems score between 0.75 and 0.99 on average suggesting service inputs could be reduced by between 25 percent and 1 percent on average to reach optimal efficiency.

Considering the effectiveness scores it can be seen that Kaohsiung is the only urban rail system to have an average score of unity. Hong Kong, Lisbon and Singapore score averages of 0.98 and 0.99 respectively suggesting a good performance in operating effectiveness. Sydney is ranked bottom for its effectiveness performance with an average effectiveness score of 0.23. This implies that, although Sydney has produced the optimal level of car-km given its service inputs (good efficiency), its service is ineffective in terms of its patronage.

Using data for multiple years in the DEA analysis means that changes in the performance over time can be investigated for each of the urban rail systems. From Table 3, most systems show little variation in their efficiency and effectiveness scores over the three years of study period. There are a couple of exceptions with the effectiveness score for London substantially increasing from 0.71 in 2009 to 0.96 in 2011. According to the 2011 annual report, this improvement was achieved by a reduction of staff numbers in 2011 which resulted in a saving of operating cost. The second system showing more variation was Taipei where the efficiency score increased from 0.75 in 2009 to 0.88 in 2011 as a result of an 11.4 km of network extension which opened after 2009. The extension line has enabled the operator of

Taipei to provide more efficient services with only a marginal increase in service inputs.

The comparison between efficiency and effectiveness is illustrated in Figure 6, where each system is located according to its efficiency and effectiveness scores. It can be observed that there is an obvious positive correlation between efficiency and effectiveness. Moreover, most observations follow a linear trend which corresponds to the finding in Karlaftis (2004) where efficient systems tend also to be effective. The exception is Sydney with high efficiency but low effectiveness. In general, the urban rail systems in Asian cities have higher effectiveness performance than other regions, given all the Asian systems are located on the top right hand side of the scatter plot, together with London and Lisbon.

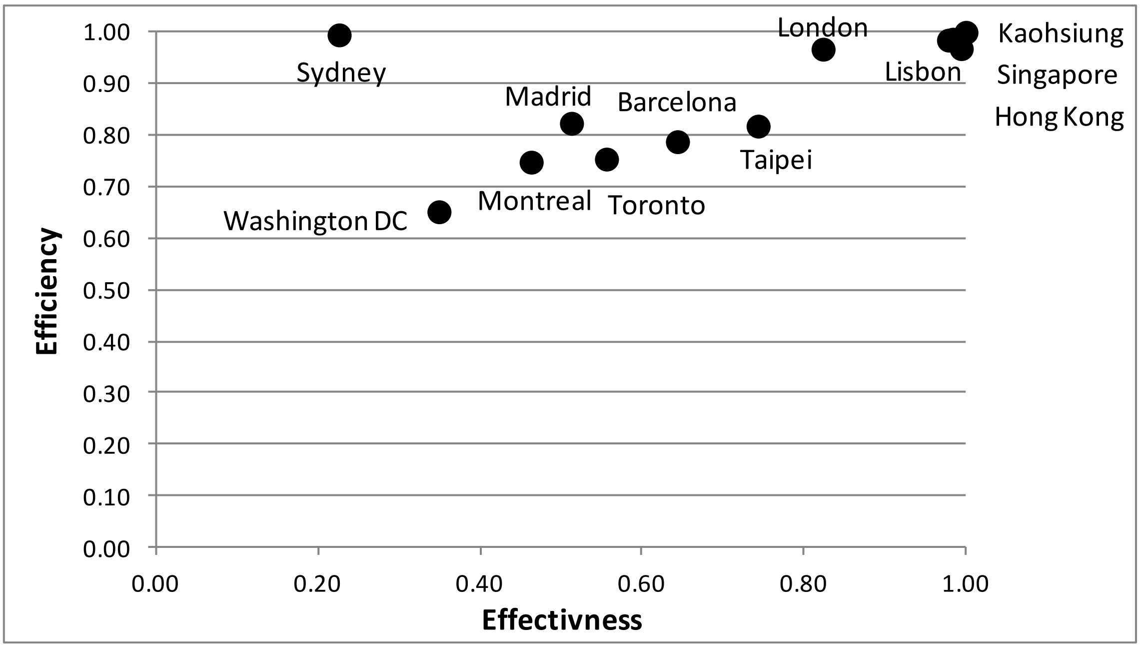

Figure 6. A Comparison of Efficiency and Effectiveness Scores

One of the outcomes of DEA analysis is the idenification of the inefficient use of service inputs or “slack”. Table 4 summarises the input slacks for each of the systems identified using car-km and patronage as multiple outputs in the DEA model. For Sydney, there are 473 employees and 304 million dollars of operating cost identified as redundant inputs, which corresponds to about 3 percent of employees and 10 percent of operating cost in 2011. London which had a major reduction in staff numbers in 2011, as discussed above, has successfully reduced its employee and cost slack to zero in 2011. Thus DEA analysis can be used as a reference point in their operational stragegies by providing evidence as to the extent of redundancy in their service inputs.

Table 4. The Combined Score and Input Slack of the Urban Rail Systems

	Employee Slack			Car Slack			Cost Slack		
	2009	2010	2011	2009	2010	2011	2009	2010	2011
Barcelona	0	0	0	44	0	0	49	110	160
Hong Kong	0	0	0	0	0	0	0	0	0
Kaohsiung	10	0	0	0	0	0	5	0	0
Lisbon	0	0	0	0	0	3	0	0	2
London	1,095	912	0	0	278	0	114	0	0
Madrid	0	0	0	311	377	314	356	363	353
Montreal	483	474	0	23	4	0	0	0	13
Singapore	38	219	0	72	56	0	0	0	0
Sydney	0	0	473	0	0	0	0	0	304
Taipei	194	693	489	0	97	61	0	0	0
Toronto	352	322	291	0	0	0	0	0	0
Washington DC	0	0	0	0	0	0	147	202	233

4.2 Scale Efficiency

The DEA employed in this analysis assumes variable constant returns to scale (RTS). That is, the urban rail systems in the sample are allowed not to be at the optimal scale in terms of their service inputs. The scale efficiency score from 2009 to 2011 in Table 5 refers to the difference between the current scale efficiency and the optimal scale.

Table 5 shows that London and Singapore are at the optimal scale of operation with constant returns to scale identified in 2011. Most systems show increasing returns to scale, especially Kaohsiung and Lisbon suggesting the scale of service inputs could be increased by around 30 to 35 percent to reach the scale efficiency. This is because Kaohsiung and Lisbon are small systems relative to the group of 12 urban rail systems in this study in terms of network size. Only the three systems of Taipei, Hong Kong, and Sydney show decreasing returns to scale in 2011. However, their scale efficiency scores are higher than 0.90 which suggests they are still close to the optimal scale. The identification of scale efficiency in this analysis confirms that most urban rail systems have scale economies as suggested by the literature (Braeutigam, 1984, Karlaftis, 2004). London again shows the most noticeable change in scale efficiency over time: decreasing returns to scale in 2009 and 2010 but constant returns to scale in 2011 after the reduction in staff numbers. Sydney and Taipei, although showing changes in RTS over time, are not considered to experience dramatic scale efficiency changes given that their scale efficiency are close to unity over the three years.

Table 5. Scale Efficiency of the Urban Rail Systems

	Scale Efficiency			RTS[1]		
	2009	2010	2011	2009	2010	2011
Barcelona	0.92	0.98	0.99	1	1	1
Hong Kong	0.92	0.93	0.91	-1	-1	-1
Kaohsiung	0.69	0.70	0.72	1	1	1
Lisbon	0.63	0.66	n/a	1	1	n/a
London	0.93	0.89	1.00	-1	-1	0
Madrid	0.99	0.99	0.99	1	1	1
Montreal	0.98	0.98	0.96	1	1	1
Singapore	0.98	0.99	1.00	1	1	0
Sydney	1.00	n/a	0.98	0	n/a	-1
Taipei	0.99	1.00	0.99	1	1	-1
Toronto	0.97	0.97	0.97	1	1	1
Washington DC	0.98	0.97	0.97	1	1	1

[1]RTS=-1: decreasing return to scale;
RTS= 0: constant return to scale;
RTS= 1: increasing return to scale;

5. CONCLUSION

This paper applies the PFP and DEA methodologies to measure the efficiency and effectiveness performance of Sydney CityRail relative to another 11 urban rail systems. Comparing the results between the PFP and DEA, it is clear that DEA gives a more comprehensive understanding of performance by processing multiple inputs and outputs. The relative efficiency and effectiveness elements identify the extent to which an operator could reduce its service inputs and yet produce the current level of output. The PFP approach only compares KPIs based on one input and one output and, although easy to understand, has no theoretical base as to which KPI should be justified or preferred as the overall indicator for performance evaluation.

The results of the DEA analysis show that Sydney CityRail has a good efficiency performance in terms of the car-km operated. However, it scores the lowest of all 12 systems for its effectiveness. This finding suggests that, as such a large-scale urban rail system, Sydney CityRail is unable to attract the patronage it needs given its service inputs. This DEA analysis also identifies considerable employee slack and cost slack in CityRail, suggesting the efficiency and effectiveness performance could be improved by reducing both staff numbers and operating cost.

For other international urban rail systems, Kaohsiung, Singapore, and Hong Kong perform equally well in both efficiency and effectiveness scores. London has made substantial improvement since 2009 by reducing its staff numbers. The network size of these efficient and effective systems varies from 39 km in Kaohsiung to 402 km in London, suggesting that technical efficiency and effectiveness can be achieved regardless the operation scale.

It is important to note that the performance scores presented in this paper are relative measures and are subject to the operational performance of the sample. It is inevitable that there will be some external factors influencing the efficiency and effectiveness performance as well as unobserved heterogeneity that is not controlled by the DEA methodology. A post-DEA regression analysis would examine those impacts and this is to be undertaken in future research. Some systems in this paper are multi-modal operators in which the number of employees and operating cost are weighted by car-km and this too is the subject of further investigation in the future using sensitivity analysis if further data acquisition does not provide the solution. Ideally more Australian evidence is needed to compare Sydney CityRail to heavy rail systems in other Australian capital cities which may share more common operational characteristics and external factors such as land use density, but this will require further research with more data provided by the rail operators.

REFERENCES

Anderson, R. J. Harris, N. G. (2007) An international comparison of urban rail boarding and alighting rates. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit,* 221(4), 521-526.

Banker, R. D., Charnes, A. Cooper, W. W. (1984) Some models for estimating technical and scale inefficiencies. *Management Science,* 30, 1078-1092.

Benjamin, J. Obeng, K. (1990) The effect of policy and background variables on total factor productivity for public transit *Transportation Research Part B,* 24(1), 1-14.

Braeutigam, R. R. (1984) A firm specific analysis of economies of density in the U.S. railroad industry. *The Journal of Industrial Economies,* 18, 3-20.

Cantos, P. Maudos, J. (2001) Regulation and efficiency: the case of European railways. *Transportation Research Part A,* 35(5), 459-472.

Charnes, A., Cooper, W. W. Rhodes, E. (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research,* 2, 429-444.

Chu, X., Fielding, G. J. Lamar, B. W. (1992) Measuring transit performance using data envelopment analysis. *Transportation Research Part A,* 26(3), 223-230.

Farrell, M. J. (1957) The measurement of productive efficiency. *Journal of the Royal Statistical Society,* 120, 449-460.

Fielding, G. J., Babitsky, T. T. Brenner, M. E. (1985) Performance evaluation for bus transit. *Transportation Research Part A,* 19(1), 73-82.

Gathon, H.-J. Pestieau, P. (1995) Decomposing efficiency in its managerial and regulatory components. *European Journal of Operation Research,* 80(3), 500-507.

Independent Pricing and Regulatory Tribunal 2008) *Review of CityRail fares, 2009-2012*. Independent Pricing and Regulatory Tribunal, Sydney, Australia.

Karlaftis, M. G. (2004) A DEA approach for evaluating the efficiency and effectiveness of urban transit systems. *European Journal of Operational Research,* 152(2), 354-364.

Karlaftis, M. G. McCarthy, P. S. (1997) Subsidy and public transit performance: A factor analytic approach. *Transportation,* 24(3), 253-270.

Karlaftis, M. G. Tsamboulas, D. (2012) Efficiency measurement in public transport: Are findings specification sensitive? *Transportation Research Part A,* 46(2), 392-402.

Merkert, R., Smith, A. S. J. Nash, C. A. (2010) Benchmarking of train operating firms – a transaction cost efficiency analysis. *Transportation Planning and Technology,* 33(1), 35-53.

Mulley, C. (2003) The benchmarking of the internal efficiency of local public transport. *Trasporti Europeii,* 4, 13-19.

Oum, T. H., Waters, W. G. Yu, C. (1999) A survey of productivity and efficiency measurement in rail transport. *Journal of Transport, Economics and Policy,* 33(1), 9-42.

Pina, V. Torres, L. (2001) Analysis of the efficiency of local government services delivery. An application to urban public transport. *Transportation Research Part A,* 35(10), 929-944.

Sampaio, B. R., Neto, O. L. Sampaio, Y. (2008) Efficiency analysis of public transport systems: Lessons for institutional planning. *Transportation Research Part A,* 42(3), 445-454.

Viton, P. A. (1997) Technical efficiency in multi-mode bus transit: A production frontier analysis. *Transportation Research Part B,* 31(1), 23-29.