Benchmarking the Efficiency Performance of International Metro Systems

Chi-Hong (Patrick) TSAI [a], Corinne MULLEY [b]

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

Abstract:
This paper employs Data Envelopment Analysis to evaluate the efficiency performance of 15 international metro systems in Asia, Australia, Europe and North America in 2011. The research outcomes identify the relative efficiency and effectiveness performance, as well as the scale efficiency of the 15 metro systems. The findings suggest that the metro systems in Singapore, Hong Kong, Kaohsiung, London, Lisbon, San Francisco, and Chicago are technically efficient, but only Singapore, London, San Francisco, and Chicago reach the optimal scale of operation. This research also investigates the potential inefficient use of service inputs in terms of labour, rolling stock, and operating cost, with results suggesting the extent to which savings in these inputs could be made in generating the current level of service outputs.

Keywords: Efficiency Performance, Benchmarking, Data Envelopment Analysis, Metro Systems

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 benchmarking is a way to investigate how a transport operator performs in terms of the service outputs with respect to the use of service inputs.

Efficiency benchmarking is particularly important for rail-based metro systems because they are usually monopolies in their markets. A monopoly market makes it more difficult for the transport regulators in measuring and monitoring the efficiency performance of the local metro operator. Operators too find benchmarking with other rail-based systems a useful exercise and one which is not commercially threatening since it is uncommon that the transport market is shared by multiple metro operators within a metropolitan area. Moreover, rail technology is very similar across national and international boundaries which in turn facilitates effective benchmarking.

Although some literature on benchmarking public transport systems exists, most previous studies are 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). There is also a lack of international evidence focussing on rail-based urban metro systems. Sampaio et al. (2008) was one of the few studies that compared efficiency performance across continents (Brazilian and European metro systems) and Anderson and Harris (2007) have analysed the performance of 22 worldwide metro systems with a focus on passenger alighting and boarding rates with respect to the operating characteristics of the metro systems such as frequency and stop time. A comprehensive benchmarking on the efficiency performance of international metro systems is not evident in the literature.

This paper evaluates the efficiency performance of 15 international metro systems in Asia, Australia, Europe and North America using a Data Envelopment Analysis (DEA) approach. 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) method. Section 4 presents the research outcomes and Section 5 concludes this paper.

2. LITERATURE REVIEW

The efficiency performance of public transport systems has been receiving extensive attention internationally with the definition of the efficiency measures for public transport being 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 refers to the total service outputs (car-km travelled, car-hour operated) with respect to service inputs (labour, fuel consumption, or operating cost), whereas effectiveness represents the service consumption (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).

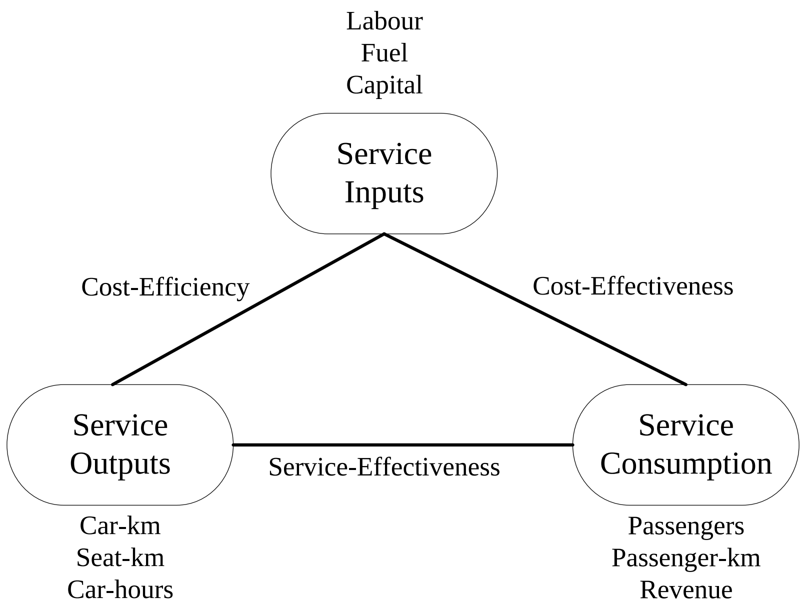

Figure 1. A framework for benchmarking the performance of public transport systems
Source: reproduced from Field et al. (1985)

The methodology used to evaluate operating performance has also been discussed and reviewed (Oum et al., 1999, Merkert et al., 2010, Karlaftis and Tsamboulas, 2012). The most used 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, but the PFP measures only processes one input against one output, and as a result multiple Key Performance Indicators (KPIs) would be produced without a single overall indicator for benchmarking. 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). The outputs of DEA also provide information about a firm's scale efficiency. This methodology is adopted in this paper.

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.

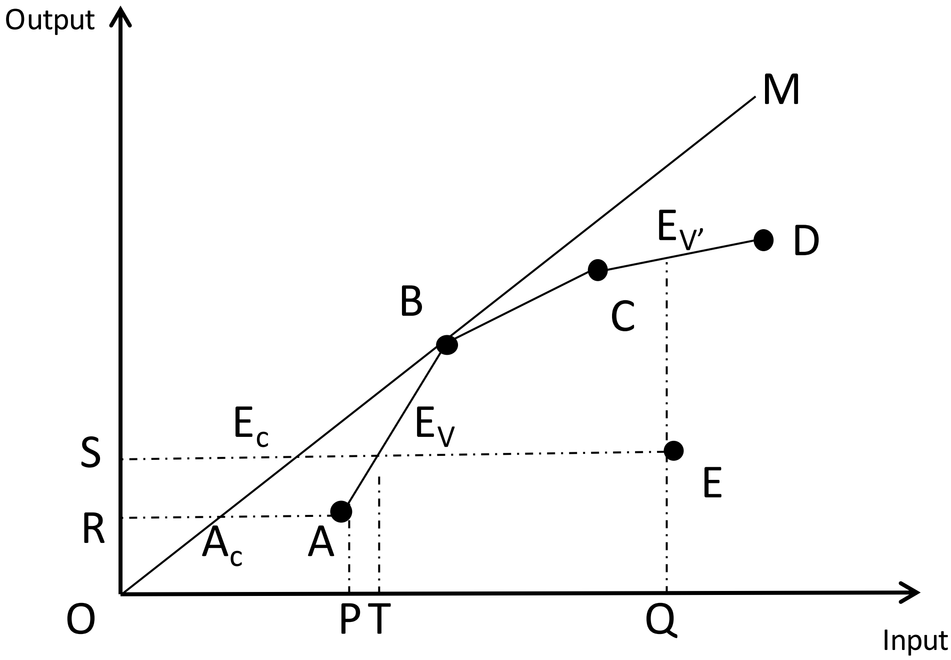

Figure 2. The Production Frontier and Efficiency Measurement

As reviewed above, DEA is a well-developed methodology for benchmarking 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:

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

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:

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

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

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

3.2 Data

This paper analyses 15 rail-based metro systems across Asia, Australia, Europe, and North America. The data are collected from publicly available financial reports or annual reports in 2011 from the operators' official websites. This research attempted to collect data from more metro system operators worldwide, but most of the operators do not publicly release their financial or operating data. The set of 15 metro systems is the largest sample size this research could achieve with a consistent dataset without missing variables.

The service outputs are categorised into efficiency measures and effectiveness measures as suggested by Field 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. Both variables are based on 2011 annual data.

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 15 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 India 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 other limitation of the data is that some system operators in this dataset are multi-modal operators, so the financial data and the number of employees are not separately specified for the metro sector in their annual reports. These systems include Singapore, Montreal, Toronto, Hong Kong, London, Chicago, and Washington DC, and the number of employee and the operating cost of these systems are weighted by the car-km operated by the different modes.

Table 1. Descriptive Statistics of the Dataset

Variable	Unit	Observation	Mean	S.D.	Min	Max
Employee	persons	15	6,891	5,401	1,329	19,064
Car	cars	15	1,140	711	126	2,545
Operating cost	AU dollars (million)	15	3,107	4,498	135	15,000
Car-km	km (million)	15	154	125	14	507
Passengers	persons (million)	15	467	380	50	1,366

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

3.3 Preliminary Analysis

A preliminary analysis is conducted to examine the data quality and to identify possible outliers in the sample. This 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 Field 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 the cost-efficiency can be compared. Of the 15 metro systems examined, Kaohsiung and Washington DC are the least efficient systems, in which the operating costs are around 14 dollars per car-km. The most efficient systems are Singapore and Chicago with less than 6 dollars per car-km.

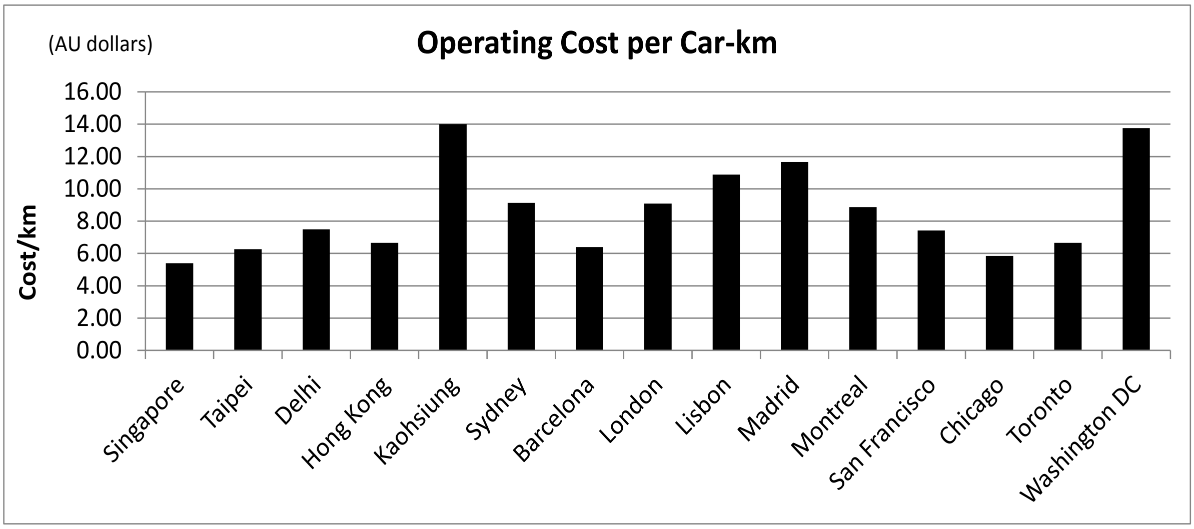

Figure 3. Cost-Efficiency of 15 Metro Systems

The cost-effectiveness measure is presented in Figure 4, where Sydney and San Francisco appear to be the least cost-effective systems with around 8 to 9 dollars of operating cost per passenger carried in 2011. The Asian metro systems in Singapore, Taipei, Delhi, and Hong Kong, together with Barcelona, Lisbon, and Toronto, perform better than other systems in terms of cost-effectiveness. Figure 4 also shows that there is more variation in the cost-effectiveness measure across the 15 systems than the cost-efficiency measure in Figure 3. This is because the number of passengers is less influenced by the operating cost. Instead, other factors might affect the patronage such as land use density and public transport fares.

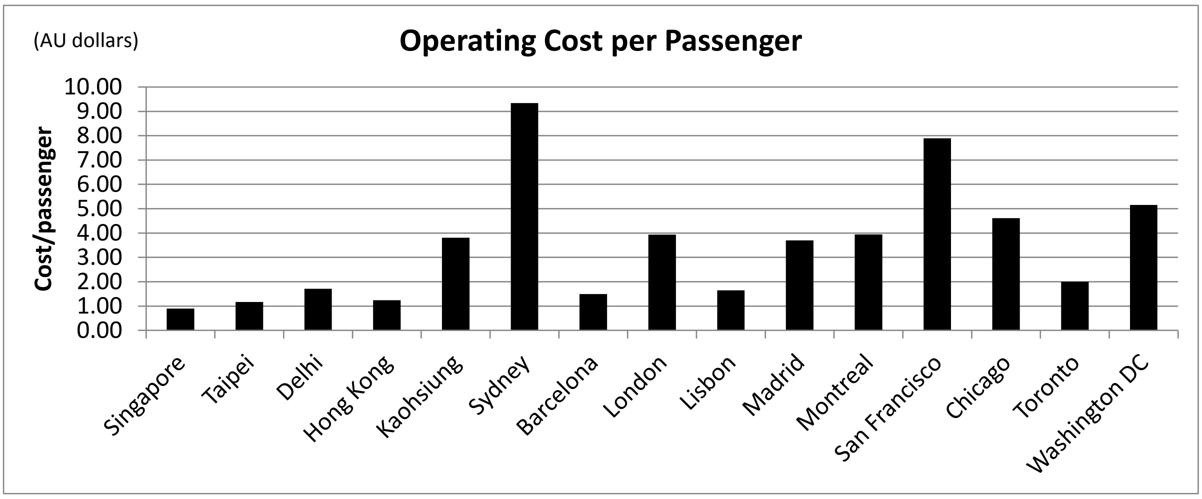

Figure 4. Cost-Effectiveness of 15 Metro Systems

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 metro systems in Asia generally have more passengers per car-km as well as in Lisbon. This is because most Asian metro systems have a higher passenger patronage than North-American and Australian systems with respect to the car-km operated. Lisbon also has a markedly high service-effectiveness because its network size is around 39.6 km which is relatively small, as compared to other systems, and this also results in a smaller level of car-km operated in Lisbon.

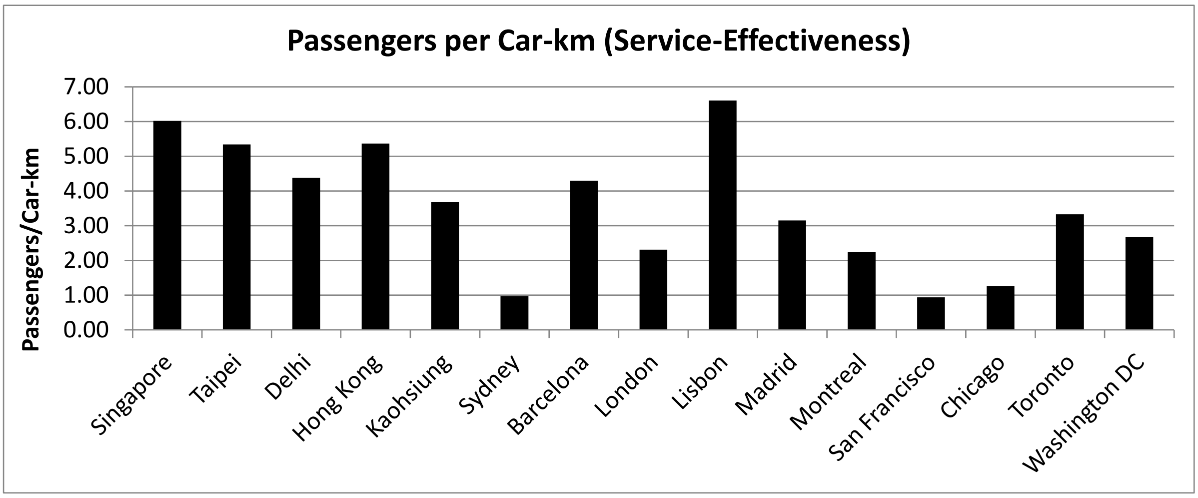

Figure 5. Service-Effectiveness of 15 Metro Systems

From the discussion of the preliminary analysis above, the general operating characteristics and performance can be evaluated based on each of the performance measures. However, as reviewed in Section 2, the limitation of this PFP approach is that there is no an 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. The next section presents the results from the DEA method, 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 2, in which the scores for efficiency and effectiveness of the 15 metro 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 the inputs of the number of employee, the number of cars, and operating cost.

Looking at the efficiency scores in Table 2, Singapore, Hong Kong, Kaohsiung, London, San Francisco, and Chicago each has a score of unity which suggests that these 6 systems have reached technical efficiency in terms of the car-km produced from the quantity of service inputs. Washington DC has the lowest score of efficiency at 0.59, indicating that the service inputs could be reduced by 41 percent to generate its current level of car-km. Other systems score between 0.72 and 0.93 suggesting their service inputs could be reduced by between 7 percent and 28 percent to reach optimal efficiency.

For the effectiveness scores, Singapore, Hong Kong, Kaohsiung, and Lisbon have a score of unity and are therefore the most cost-effective metro systems. San Francisco and Chicago, which score unity for efficiency, have much lower scores for effectiveness at 0.53 and 0.35 respectively. This indicates that these two systems generate a high level of car-km but do not attract sufficient passengers with respect to their service inputs. This result leads to the service-effectiveness measure in Figure 5, where San Francisco and Chicago both have low service-effectiveness ratios. Likewise, Sydney is ranked bottom for effectiveness performance although it scores 0.93 for efficiency. This is because the metro system in Sydney is a commuter rail system with a substantially larger network size (2,232 km) than other metro systems in the dataset and thus is able to produce a large amount of car-km whilst the patronage is much lower than other systems.

Table 2. The Results of Data Envelopment Analysis

City	Region	Efficiency		Effectiveness	
		Score	Rank	Score	Rank
Singapore	Asia	1	1	1	1
Taipei	Asia	0.87	12	0.78	7
Delhi	Asia	0.93	9	0.80	6
Hong Kong	Asia	1	1	1	1
Kaohsiung	Asia	1	1	1	1
Sydney	Australia	0.93	7	0.22	15
Barcelona	Europe	0.89	11	0.71	8
London	Europe	1	1	0.97	5
Lisbon	Europe	0.93	8	1	1
Madrid	Europe	0.75	13	0.53	10
Montreal	North America	0.72	14	0.41	12
San Francisco	North America	1	1	0.53	11
Chicago	North America	1	1	0.35	13
Toronto	North America	0.91	10	0.57	9
Washington DC	North America	0.59	15	0.35	14

The comparison between efficiency and effectiveness is illustrated in Figure 6, where each metro system is located according to its efficiency and effectiveness score. It can be observed that the correlation between efficiency and effectiveness is positive and reasonably linear, which corresponds to the finding in Karlaftis (2004) where efficient metro systems tend also to be effective. The exceptions to this are Sydney, Chicago, and San Francisco as discussed above. In general, the Asian metro systems have higher effectiveness performance than other regions, given all the Asian metro systems are located on the top right hand side of the scatter plot, together with London and Lisbon.

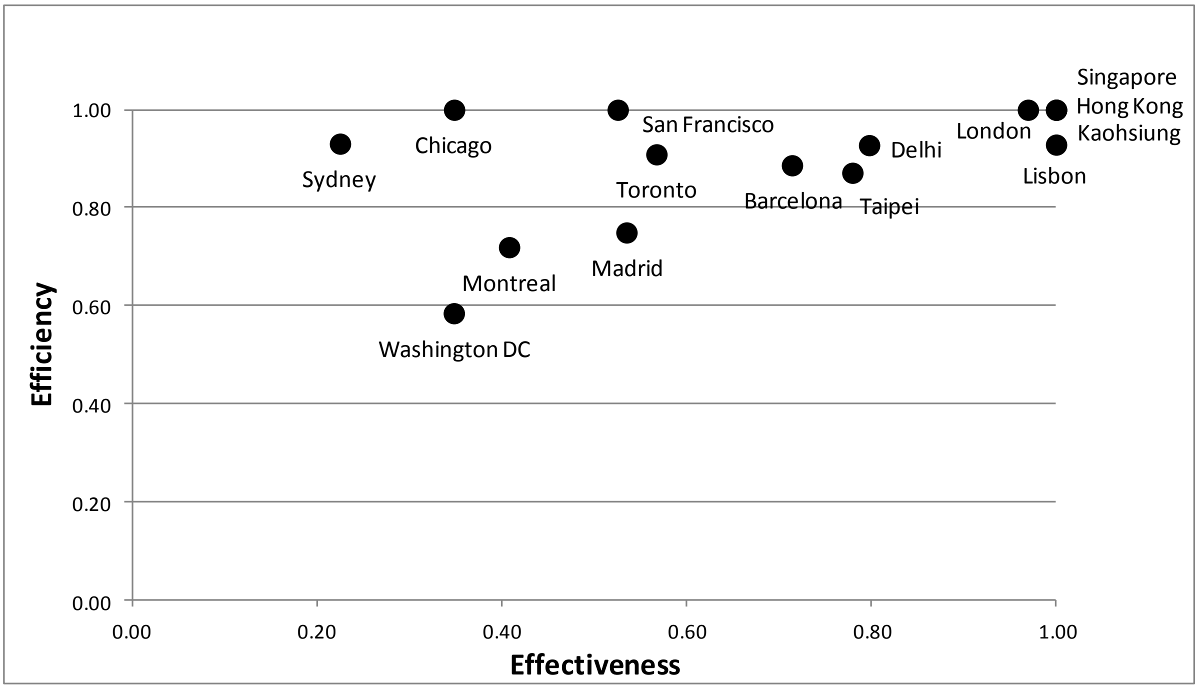

Figure 6. A Comparison of Efficiency and Effectiveness Scores

As reviewed in Section 2, the advantage of DEA is its capability of processing mulitple inputs and outputs to generate an overall measure for benchmarking. The discussion on efficiency and effectivenss above investigates the operating performance and characteristics for each system. The overall performance measure is discussed below by employing both car-km and the number of passengers as the outputs of the DEA model. These results generate a combined score for the performance benchmarking and identify the slack of inputs and is presented in Table 3.

Singapore, Hong Kong, Kaohsiung, London, Lisbon, San Francisco, and Chicago have a score of unity for the overall performance measure. The system with the lowest overall score is Wasingtiong DC at 0.59. Other systmes score between 0.72 to 0.93. The other practical outcome of DEA is the identification of the inefficient use of service inputs or "slack". The slack is the redudant quantity of a service input over and above technical inefficiency. For example, Sydney scores 0.93 for its overall performance with 4,497 employees being identified as slack. This means Sydney could reduce all its service inputs by 7 percent and produce its current level of outputs. In addition, 4,497 employees could be removed after all the service inputs are reduced by 7 percent. Those systems with employee slack identified are

Tiapei, Delhi, Sydney, Montreal, Toronto, and Washington DC. Car slack is evident in Taipei, Barcelona, and Madrid, indicating these systems have redudant car fleets in operation. The only system with a cost slack identified is Madrid, which shows 344 million (AU dollars) could be saved.

Table 3. The Combined Score and Input Slack of the Metro Systems

City	Combined Score	Employee Slack	Car Slack	Cost Slack
Singapore	1	0	0	0
Taipei	0.87	520	62	0
Delhi	0.93	1,378	0	0
Hong Kong	1	0	0	0
Kaohsiung	1	0	0	0
Sydney	0.93	4,497	0	0
Barcelona	0.89	0	68	0
London	1	0	0	0
Lisbon	1	0	0	0
Madrid	0.82	0	803	344
Montreal	0.72	298	0	0
San Francisco	1	0	0	0
Chicago	1	0	0	0
Toronto	0.91	875	0	0
Washington DC	0.59	500	0	0

4.2 Scale Efficiency

The DEA employed in this analysis assumes variable constant returns to scale. That is, the metro systems in the sample are allowed not to be at the optimal scale in terms of the service inputs. The scale efficiency score in Table 4 refers to the difference between the current scale efficiency and the optimal scale.

Table 4 shows that Singapore, London, San Francisco, and Chicago are at the optimal scale of operation with constant returns to scale identified. Most systems show increasing returns to scale, especially Kaohsiung and Lisbon in which the scale of service inputs could be increased by around 35 percent to reach scale efficiency. This is because Kaohsiung and Lisbon are the smallest systems amongst the 15 metro systems in terms of network size, so it is reasonable that their current operation scales are substantially smaller than the optimal scale. Only three systems show decreasing returns to scale, which are Taipei, Hong Kong, and Madrid. 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 metro systems have scale economies as suggested by the literature (Braeutigam, 1984, Karlaftis, 2004).

Table 4. Scale Efficiency of the Metro Systems

City	Scale Efficiency Score	Returns to Scale
Singapore	1.00	Constant returns to scale
Taipei	0.99	Decreasing returns to scale
Delhi	0.99	Increasing returns to scale
Hong Kong	0.90	Decreasing returns to scale
Kaohsiung	0.63	Increasing returns to scale
Sydney	0.99	Increasing returns to scale
Barcelona	0.96	Increasing returns to scale
London	1.00	Constant returns to scale
Lisbon	0.65	Increasing returns to scale
Madrid	0.96	Decreasing returns to scale
Montreal	0.96	Increasing returns to scale
San Francisco	1.00	Constant returns to scale
Chicago	1.00	Constant returns to scale
Toronto	0.99	Increasing returns to scale
Washington DC	0.96	Increasing returns to scale

5. CONCLUSION

This paper applies the PFP approach and the DEA methodology to measure the efficiency performance of 15 international metro systems. Comparing the results between the two approaches, it is clear that DEA gives a more comprehensive indicator of the efficiency performance by processing multiple inputs and outputs. The relative efficiency and effectiveness elements identify the extent to which a metro operator could reduce its service inputs and yet produce the current level of output. The PFP approach essentially compares KPIs based on one input and one output and, although easy to understand, have no theoretical base as to which KPI should be justified or preferred as the overall indicator for benchmarking.

The results of the DEA show that the metro systems in Singapore, Hong Kong, Kaohsiung, London, Lisbon, San Francisco, and Chicago are technically efficient, but only Singapore, London, San Francisco, and Chicago reach the optimal scale. By distinguishing efficiency and effectiveness, most Asian metro systems are seen to be more effective as a result of higher patronage. In contrast, metro systems in North America and Sydney have a worse performance in effectiveness as compared to efficiency. The research outcomes also identify the potential slack of service inputs and the scale efficiency for each of the metro system. This finding gives practical information for the operators or regulators about how the system could be more efficiently operated, by either reducing the quantity of service inputs or adjusting the scale of the system.

It is important to note that the performance scores presented in this paper are relative measures, which are subject to the operation performance within the sample. It is inevitable that there may be some unobserved heterogeneity that cannot be controlled by the DEA approach. A larger sample size would be preferable to reduce the sensitivity of the research results. The major contribution of this paper is to benchmark rail-based metro systems across various regions. The results allow these metro operators to compare their performance to other metro systems particularly where a country has a limited number of metro systems to conduct domestically based benchmarking.

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.

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.