Transport Policy 162 (2025) 353–363

Contents lists available at ScienceDirect

Transport Policy

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

The WCTRS global subway efficiency benchmarking task force inaugural report: The key findings, lessons learned, policy issues investigated

Tae Hoon Oum [a,*], Kun Wang [b], Chunyan Yu [c], Pierlugi Coppola [d], Luigi Castagna [d], Hironori Kato [e], Yiping Le [f], Xiao Luo [g], Suhui Gan [g], Kyungtaek Kim [h], Woojin Kim [h], Gopal Patil [i], Vinayak Gaur [i], Giovanni Circella [j], Hossain Mohiuddin [k], Maria Carolina Lecompte [l], Clark Lim [m]

[a] University of British Columbia, 2053 Main Mall, Vancouver, BC V6M 2E8, Canada
[b] Hong Kong Polytechnic University, China
[c] Embry Riddle Aeronautical University, Daytona Beach, FL, USA
[d] Politecnico di Milano, Department of Mechanical Engineering, Via G. La Masa 1, 20156, Milano, Italy
[e] Department of Civil Engineering, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan
[f] Program of Civil Engineering, Shibaura Institute of Technology, 3-7-5, Toyosu, Koto-ku, Tokyo, 135-8548, Japan
[g] Tongji University, Shanghai, China
[h] Korea Transport Institute, Seoul, South Korea
[i] Indian Institute of Technology Bombai, Mumbai, India
[j] University of California at Davis, Davis, CA, USA
[k] University of California at Davis CA, USA
[l] University of California at Davis, USA
[m] University of British Columbia, Canada

A R T I C L E   I N F O

Keywords:
Subway efficiency benchmarking
Variable input productivity (VIP) method
Policy interventions
WCTRS

A B S T R A C T

This study presents the key summary of the inaugural report of the WCTRS Global Subway Efficiency Benchmarking Task Force, focusing on evaluating the operating efficiency of subway systems across Asia, Europe, and North America. Given the increasing need for urban transit authorities to balance multiple objectives, such as reducing greenhouse gas emissions and enhancing public health, this benchmarking initiative provides a crucial assessment of their productivity and efficiency of operations. We employed the Variable Input Productivity (VIP) method intelligently computed by the translog multilateral indexing method, chosen over DEA due to its transitive properties and better suitability for policy applications. Our analysis measures the VIP index, specifically examining labor and soft input productivities, to gauge how efficiently subway systems utilize their variable resources. By adjusting for factors beyond management control through regression analysis, we derived Net VIP scores, offering a more accurate comparison across different cities.

The results reveal significant variations in efficiency levels, with some cities demonstrating remarkable productivity despite limited resources while others lag behind due to structural and operational challenges. Our findings underscore the importance of targeted policy interventions to enhance the efficiency of urban rail systems. Above all, the WCTRS Task Force members who volunteered their time and effort hope to raise awareness of efficiency and productivity as an important aspect of managing and operating the subways and other city transit systems.

* Corresponding author.
E-mail address: tae.oum@ubc.ca (T.H. Oum).

https://doi.org/10.1016/j.tranpol.2024.12.012
Received 29 May 2024; Received in revised form 14 December 2024; Accepted 15 December 2024
Available online 18 December 2024

1. Background and introduction

1.1. The WCTRS subway efficiency task force: background and objectives

The World Conference on Transport Research Society (WCTRS) was formed in 1976. Our 50th anniversary is only a few years away. In the 45+ years, the WCTRS has had significant progress and great success in improving its global reach, primarily to academics and researchers from 100+ countries, mainly through its triennial world conferences. There was a strong desire within the WCTR Society to initiate projects for contributing directly to the transport and supply chain industries while continuing to improve our contributions to the research community. Soon, our discussions began to focus on the topic of productivity with which the city authorities conduct their transit business. Since the transit authorities are typically organized by city governments as natural monopolies. In the absence of competition, they do not pay sufficient attention to improving the efficiency by which they conduct city transit business. This lack of efficiency concern is hidden and ignored, especially since cities and transit authorities need to deal with an increasing number of multiple objectives to pursue, including traffic congestion reduction, promoting alternative transport modes, pursuing sustainable economic development, reducing negative impacts of transport on public health; supporting climate change policy objectives, etc.

Since the WCTRS is made up largely of academic researchers, many of us have been conducting independent research on urban transport issues, including subway and urban rail projects. Our academic research interest is consistent to improve consumer welfare by enhancing the productivity of all sectors of the economy, including urban transport. Our motivation is not to benefit any particular firm or transit authority and, thus, is different from profit-seeking consulting organizations. Thus, this Task Force is consistent with social benefit maximization by raising awareness of productive efficiency even when the transit authorities need to pursue multiple objectives. As cities focus on popular objectives of the day, such as reducing GHG emissions, they tend to forget they still need to do all of those activities most efficiently without wasting inputs.

We believe that the WCTRS has the credibility to do the subway efficiency benchmarking work properly because our members include the world's top academic experts, many of whom also have experience in advising cities, major firms and regulatory agencies. Since the WCTRS has urban transport experts from virtually all countries in the world, we are the most suitable organization for conducting the subway efficiency benchmarking project. In addition, since we are a non-profit academic society, we are in the best position to conduct the efficiency benchmarking work on a global basis, even in the absence of financial sponsors for the project.

1.2. Scope of the efficiency benchmarking study

The urban transport sector is very large and complex as it includes different modes of transportation involving a large number of municipal and city boundaries. As stated above, urban transit authorities pursue a number of multiple objectives involving multiple municipal boundaries. As such, it is difficult for any research organization to measure and evaluate efficiency of all aspects of an individual urban transportation system.

Therefore, in the first year, our Task Force has decided to focus on the efficiency with which transit authorities conduct their heavy rail/subway operations. However, in many cities, subway networks and operations are connected with complex urban rail networks. As such, we decided to measure the efficiency of the subway-urban rail network as a system since, in most cases, it is impossible to capture their accounting information (revenues and costs) separately between subway and urban-rail operations.

1.3. Focus on the operating efficiency (variable input productivity)

It is almost impossible to measure the true capital cost of subway-urban rail systems because the true capital costs include various forms of direct and indirect subsidies and tax credits from all levels of the government, capital cost allowance provisions and associated depreciation policies on capital assets, and other direct and indirect tax advantages. It would be nearly impossible to create a database for capital stock series comparable between cities, between countries and across different continents. Therefore, in the first year of the Task Force work, we decided to focus on the operating efficiency of each city's subway-urban rail system. The operating efficiency of the subway-urban rail system can be measured reliably by capturing the variable input productivity (VIP), which measures how efficiently the subway-urban rail system uses its variable inputs, i.e., labor and other variable inputs of all kinds.

2. Review of literature on key past work relevant to subway efficiency

The study of subway efficiency has garnered significant attention from both academic researchers and industry practitioners, with various methodologies employed to measure and benchmark the performance of urban rail systems. This section summarizes key empirical works only, highlighting their methodologies, findings, and contributions. The reason for this limited literature review is that literally more than 10,000 DEA papers have been published so far since the seminal paper by Charnes et al. (1978) was published.

2.1. Specific methodologies

2.1.1. Data Envelopment Analysis (DEA) approaches

Data Envelopment Analysis (DEA) has been a widely used method for assessing the efficiency of public transit systems. Studies by Oum et al. (1999) applied DEA to evaluate the performance of bus transit systems, which laid the groundwork for its application to subway systems. Zhu (2003) extended this approach to assess the efficiency of subway systems in China, identifying significant variations in efficiency levels across different cities. DEA's ability to measure the efficiency of firms (Decision Making Units, DMU) without the price data on inputs and outputs made it a popular choice, though it often faces criticism for its sensitivity to data noise and lack of transitivity in comparisons, as discussed in detail in Appendix A of this paper.

2.1.2. Stochastic Frontier Analysis (SFA)

Stochastic Frontier Analysis (SFA) has also been utilized to measure subway efficiency. Studies by Coelli et al. (2005) and Viton (1997) applied SFA to urban transit systems, focusing on the impact of environmental variables on efficiency. Lan and Lin (2002) employed SFA to evaluate the technical efficiency of metro systems in Taiwan, incorporating factors such as scale economies and operational environments. Recently, Castagna et al. (2024) used SFA to measure and compare the efficiency of the European subways. SFA's ability to separate inefficiency from random error provides a robust framework for efficiency analysis, though it requires strong assumptions about the functional form of the production frontier.

2.1.3. Productivity index utilizing translog multilateral indexing method

The translog multilateral indexing method, as proposed by Caves et al. (1982), offers a flexible approach to measuring productivity that accounts for transitivity, making it suitable for benchmarking studies. Andrikopoulos et al. (1993) applied this method to compare the productivity of public transport systems across different countries, highlighting its advantages over non-parametric approaches like DEA. More recently, Graham (2008) utilized the translog index to analyze the efficiency of European metro systems, finding that it provides a more nuanced understanding of productivity changes over time.

2.1.4. Comparative empirical benchmarking studies

Benchmarking studies conducted by organizations such as the International Association of Public Transport (UITP) and the American Public Transportation Association (APTA) have provided comprehensive comparisons of subway system performance. The UITP's 2015 report on global metro benchmarking highlighted key performance indicators (KPIs) such as energy efficiency, service frequency, and passenger satisfaction. Similarly, the APTA's annual benchmarking reports offer valuable insights into the operational and financial performance of North American transit systems, though they often lack the methodological rigor seen in academic studies.

2.1.5. Integrated approaches

Recent studies have sought to integrate multiple methodologies to provide a more comprehensive analysis of subway efficiency. For instance, Boame (2004) combined DEA and SFA to evaluate the efficiency of Canadian urban transit systems, addressing the limitations of each method individually. Similarly, Haworth et al. (2015) employed a mixed-method approach, incorporating both quantitative and qualitative data to assess the efficiency of Australian metro systems, emphasizing the importance of contextual factors in performance measurement.

2.2. Transport for London (TfL)'s underground benchmarking reports

Transport for London (TfL), a statutory agency of the London government, has begun to publish its Rail and Underground International Benchmarking Reports since 2014. At first, the reports have been published for four consecutive years, from 2014 to 2017. After that, only the 2023 report was released. TfL's reports aim to compare some key indicators of London's urban rail systems (London Underground and Docklands Light Railway) with other members of CoMET (community of metros) and Nova Group of Metros, covering major metropolises around the world. Specifically, such benchmarking work covers cities in North America (e.g., Toronto, Montreal, New York, etc), Europe (e.g., London, Lisbon, Madrid, Barcelona, Paris, Berlin, Brussels, etc) and Asia (e.g., Beijing, Shanghai, Shenzhen, Hong Kong, Singapore, etc). A major modification after 2017 was that in addition to the underground and light rail systems, the suburban railways and bus networks were also added to the benchmarking analysis. Therefore, the 2017 and 2023 reports are actually focused on benchmarking the entire public transportation system of London with those of other counterparts. The sample size of the 2023 report is expanded to 48, with more selected cities in North America, Northeast Asia, and Europe and the transportation systems in Africa and Middle-East countries were incorporated for the first time.

With regard to the methodology for benchmarking, TfL's reports rely on various simple indicators, and they do not come up with any aggregate efficiency index. The calculated indicators include "cost efficiency", "labor productivity", "reliability", "carbon emission", and "staff and customer safety". For instance, the cost efficiency indicator in TfL's report is calculated by the level of government subsidy required, total operating cost and maintenance cost per car kilometer, and the labor productivity indicator refers to car kilometers per total staff and contract hours. Such indicators are based on simple summary statistics that could reflect the subway's relative performance in specific dimensions. While being very valuable and important empirical work, they are not sufficient to compare the operating efficiency comprehensively between cities, between countries or between continents because price differentials of inputs and outputs between cities, between countries, and between continents are not taken into account in the TfL study.

In 2017, some new indicators have been included, such as fare affordability, demand growth rate, frequency, capacity and accessibility. The latest 2023 report further considers the financial and demand recovery of the subway system from the COVID-19 pandemic.

2.3. Summary comments on literature

The past empirical work on subway efficiency measurement and benchmarking is significant, with each methodology offering unique strengths and limitations. While DEA and SFA have been popular choices, the productivity index computational method using the translog multilateral indexing procedure has gained traction for its flexibility and robustness. Comparative benchmarking studies by industry organizations complement academic research, providing practical insights for transit authorities. This study builds on this rich literature by employing the translog index to provide a transitive and comprehensive benchmarking of subway systems across Asia, Europe, and North America, offering valuable policy recommendations based on empirical findings.

Despite the significant amount of work were done to measure and evaluate the Urban transport and urban rail systems, still the subway and urban rail operating authorities do not pay significant attention to improving the efficiency (productivity) with which they operate. Therefore, the WCTRS has formed a Task Force to raise the importance of the operating efficiency of the subway and urban rail systems.

3. Methodology chosen for efficiency measurement and benchmarking

3.1. Choosing the efficiency measurement methodology

Since our benchmarking work should be empirically driven, we need to use a simple methodology that is easily understandable by the users of our report, i.e., industry, government agencies, regulators, and even concerned citizens. Any sophisticated methodology, such as the stochastic frontier cost or production function approach or the stochastic frontier DEA (Data Envelopment Analysis) indices approach, would not work because the industry and government sectors do not use such methodologies.

Literally, tens of thousands of academic papers used numerous versions of DEA methods. Conversely, the central banks of all countries, government agencies, policy institutes, industry associations, and international organizations such as the IMF and World Bank use index number theory for measuring productivity changes and for computing price indices for all industries. The question is which one serves our purpose best. A version of the DEA method or productivity Index number method?

After serious examinations, our Task Force decided to use the productivity index number method for the reasons described below. The version of the productivity index approach used by our task force is 'translog multilateral indexing procedure proposed by Caves, Christensen and Diewert entitled "Multilateral Comparisons of Outputs, Inputs, and Productivity Using Superlative Index Numbers" (Caves et al., 1982). This procedure allows us to produce a "superlative index number", which has several important properties of an index number especially when one needs to aggregate multiple outputs and multiple inputs an organization or industry produces. All of the subways and urban rail systems produce multiple outputs using multiple inputs.

The most serious deficiency of the DEA approach for efficiency ranking of the cities is that the DEA index lacks 'transitivity', which is the most important property of an index number: i.e., DEA may not produce a 'TRANSITIVE" index. Let us explain this issue in detail by briefly explaining how the DEA approach would compute the efficiency ranking of the cities. The DEA forms an 'efficiency frontier' by the most efficient cities in the output-input space. All cities on the 'efficiency frontier' are considered 100% efficient (DEA score of 1.0). All other cities' efficiency scores are measured as the shortest Euclidian distances from their respective output-input points to the efficiency frontier. If we add another efficient city to the data set, it will change the DEA's efficiency frontier and, thus, likely change the relative ranking of all cities that are not 100% efficient. This is a critical and unacceptable problem as an efficiency benchmarking method. This problem of DEA generating intransitive index is clearly shown in Appendix A.

3.2. Approach for measuring the variable input productivity (VIP)

As stated in Section 1.3 (Focus on Operating Efficiency), at least in the first few years, our Task Force has decided to focus on measuring 'Variable Input Productivity' (VIP) of the subways and urban rail systems, which is really operating efficiency of the system. In fact, VIP captures how efficiently the transit system uses its variable inputs in day-to-day operations. Since the efficiency of using variable inputs is largely controlled by current management, VIP is a major way to evaluate the efficiency by which current transit management operates the system.

Productivity and Efficiency: We first measure and compare two partial input productivities: labor input productivity and so-called the "soft" input productivity. Labor (input) productivity is the aggregate output index divided by the labor quantity index. Obviously, the labor quantity index is either the number of full-time equivalent employees or computed by dividing the total labor expenses (including wages, pension, training and other human resource (HR) expenses) by the average total cost per employee. The 'soft input' cost includes all other variable input costs, which consists of the uncountably large number of variable inputs a firm uses, including the costs of outsourced services, consultant services, legal services, utility costs, travel expenses, non-labor building and equipment maintenance and repair services, stationaries, and all other purchased materials. In practice, firms or agencies do not report each of these hundreds of expense items separately. Therefore, the soft input costs can only be captured as the 'catch-all' variable input cost as follows:

Soft input costs = [Total operating cost] minus [all capital input-related costs (such as interest costs, depreciation and amortization, all other capital-related costs)] minus [all labor input-related costs (wages, employee training costs, pension costs, etc.)]

The soft input cost for subway-urban rail systems varies from city to city and country to country, but can account for 25%–75% of the subway-urban rail's total variable input costs. Therefore, it is very important to consider this soft cost input in evaluating subway and urban rail system's operating efficiency performance.

This soft input cost (expense amount) needs to be divided by the soft input price index in order to compute the soft input quantity index. Since soft input includes an uncountably large number of items, we use the consumer price index of the city (region) as a proxy for the soft input price index. Therefore, our soft input quantity index is computed by dividing the (catch-all) soft input expenses by the city's consumer price index.

Methods of computing output and input quantity indices: our preliminary investigation led us to conclude that the subway-urban rail system produces enough information in their annual reports and financial reports for the two revenue items: (a) Passenger kilometers (PK) produced in each year (passenger output) and/or number of passengers they served and (b) other service outputs which include advertising revenues, station space rental revenues, and other revenues. Since other service output revenues are captured as an amount, they also have to be divided by the city's consumer price index in order to compute the other output quantity index.

These two outputs (the passenger service output index and the other output quantity index) are aggregated by using the translog multilateral index aggregation procedure discussed above. The exact technical formula for this translog multilateral aggregation procedure is presented in Appendix A.

The labor quantity index and the soft cost input quantity index are aggregated into the multilateral variable input quantity index similar way as the output quantity index aggregation just discussed.

Computing (Gross) Variable Input Productivity Index (Gross VIP): The variable input productivity (VIP) index is then computed as the ratio of the aggregate output quantity index to the aggregate variable input quantity index. This procedure establishes 'transitive comparison' across all observations (time-series data of cross-section of cities) via a series of binary comparisons between each of the observations and the geometric mean of the data, as is shown explicitly in the translog multilateral aggregate indexing formula proposed by Caves et al. (1982):

$$ln\frac{X_i}{X_j} = \sum \frac{W_{ki} + \overline{W}_k}{2} ln\frac{X_{ki}}{\tilde{X}_k} - \sum \frac{W_{ki} + \overline{W}_k}{2} ln\frac{X_{ki}}{\tilde{X}_k} \tag{A-1}$$

where Xi is the aggregate index of input (or output) for i-th observation, X_{ki} is k-th input (output) for i-th observation, the W_{ki} are weights, a bar over a variable indicates the arithmetic mean and a tilde over a variable indicates the geometric mean. Revenue shares are often used as the weights in output aggregation (with the assumption of constant returns to scale), while cost shares are used as weights in input aggregation.

3.3. Net (residual) variable input productivity (net VIP)

As stated above, the gross VIP index is computed by dividing the aggregate output quantity index by the variable input quantity index. This gross VIP index shows how efficiently the city's subway-urban rail system utilizes variable inputs (labor and soft inputs) at a given level of capital infrastructure and facilities in the city.

However, the measured gross Variable Input Productivity (VIP) can be influenced over time by the level of automation of the subway system in that city, which, in turn, can be influenced by the level of capital investment made in the system over time.

More importantly, the gross VIP is not appropriate to compare directly between different cities, countries and/or continents because the gross VIP level is influenced by the following factors, which are beyond the management control of the transit authority.

- Scale (size) of the subway-urban rail network (because of economies of scale)
- Population and/or traffic density (economies of density)
- Number of rail cars per train
- Passenger density per station
- Interface between the subway-urban rail system and the rest of the transport modes
- Percentage of transit users vs. car drivers
- How much resources (input) does the transit authority use on environmental issues, including CO2 reduction and climate change policy? (in fact, how successful they are in reducing CO2 emissions).
- Etc.

Because of these and other factors beyond current management's control, it is impossible to compare the gross VIP index even between different cities within the same country.

In order to calculate the Variable Input productivities (VIPs) across different cities, it is important to take into account the factors beyond management control by applying statistical controls. One approach is to decompose the gross VIP into various sources including each of the factors (variables) beyond the management control.

Our task force used the regression analysis of the gross VIP on the variables beyond management control as many as we can measure in order to decompose the gross VIPs into those factors, and use the residuals of the regression as the measure of Net Variable Input Productivity (Net VIP). This Net VIP is far more comparable between cities and over time than the gross VIP values. Please note that in all of our regression analyses, we converted all continuous dependent and independent variables into natural logarithmic values before running regressions.

In sum, we compare these Net Variable Input Productivity (Net VIP) as the measure of the operating efficiency of the subway-urban rail system.

4. The cities included in our benchmarking and the summary of benchmarking

Although our Task Force would have liked to include all of the cities of the world with subway and/or heavy urban rail systems, our inauguration year project was constrained by the following reality.

- availability of qualified volunteer task force members who are capable and willing to contribute time and effort for collecting and doing preliminary data analysis;
- availability of trustworthy data from the city transit authorities, especially in the form of published data via annual reports and yearly audited financial reports such as balance sheets, income and expense statements, etc.
- For statistical reasons, our Task Force chose to use unbalanced panel data (cross-section and time-series of all years data of since 2000); we had to exclude the cities with only a few years of data (mostly new systems in India, for example)

Based on these constraints, our Task Force chose to include only the cities for which we were able to collect consistent data which are publicly available. The list of the cities, key data and the Net Variable Input Productivity (Net VIP) results, our subway operating efficiency measures, are reported in Tables 1a and 1b.

The winning cities with the highest Net VIP scores and the cities with the most improved Net VIP scores are briefly described by continent and country.

Table 1a
Cities included, key outputs, Inputs and Efficiency Index (Net VIP, 2019).

	System/Region	Variable Input Quantity Measures		Output Quantity Measures		Efficiency
		Number of Employees	Soft Cost Input Index (base city = 1.00)	Passenger-Km (millions)	Other Revenues Output Index (=1.00)	Net VIP (2019)
Europe (2019)	Madrid	6439	1.55	4587	3.99	1.08
	Lisbon	1435	0.15	878	0.23	1.00
	Bilbao	767	0.11	603	0.08	1.00
	Barcelona	3469	1.00	2060	1.00	0.91
	Bucharest	4445	0.34	2520	0.53	0.89
	London	16,485	3.66	11,754	4.43	0.89
Japan (2019)	Kobe	574	0.03	921	0.05	1.11
	Nagoya (Metropolitan)	35,302	2.73	84,938	1.76	1.07
	Tokyo	13,447	1.00	29,552	1.00	1.00
	Sapporo	598	0.08	1369	0.06	0.94
	Yokohama	965	0.05	1854	0.03	0.93
	Kyoto	657	0.04	763	0.06	0.91
	Fukuoka	574	0.06	904	0.05	0.89
	Osaka	4975	0.17	5486	0.22	0.89
	Tokyo (Metropolitan)	86,313	7.69	240,074	6.36	0.87
	Sendai	431	0.03	472	0.03	0.85
	Nagoya	2721	0.11	3065	0.19	0.85
	Osaka (Metropolitan)	50,774	3.52	98,767	2.69	0.82
China (2019)	Shenzhen	20,325	1.00	12,544	1.00	1.14
	Beijing	42,554	1.32	17,457	0.53	0.78
	Changsha	8649	0.29	2501	0.21	0.78
	Tianjin	8779	0.66	2694	0.10	0.71
	Guangzhou	24,794	0.64	19,303	0.57	0.67
	Chengdu	25,077	1.37	9835	0.58	0.60
	Nanjing	17,409	0.32	6681	0.06	0.39
	Xiamen	4826	0.11	912	0.03	0.29
	Hangzhou	12,089	0.21	4957	0.15	0.23
South Korea (2019)	Gwangju	578	0.15	99	0.36	1.78
	Seoul	16,886	6.12	15,516	2.99	1.30
	Incheon	1721	0.82	1063	0.82	1.20
	Busan	4780	1.35	2140	1.78	1.17
	Daegu	2602	1.00	882	1.00	1.06
	Daejeon	849	0.14	174	0.24	0.85
India (2019)	Hyderabad	144	1.12	2575	0.06	1.98
	Lucknow	753	0.31	521	0.97	1.76
	Chennai	550	1.00	1531	1.00	1.44
	Bangalore	2351	1.53	5443	0.18	1.14
	Jaipur	400	0.18	59	0.88	1.03
	New Delhi	14,353	21.87	323,762	7.90	0.92
	Kochi	509	0.41	297	0.25	0.85
North America (2019)	Toronto	15,251	1.79	4466	0.06	1.29
	New York	31,833	3.98	16,835	2.16	1.26
	Los Angeles	728	0.09	334	0.19	1.20
	Chicago	4884	0.59	1386	0.82	1.04
	Miami-Dade	642	0.07	137	0.60	1.00
	Washington DC	7470	1.00	1321	1.00	0.78
	Boston	1732	0.28	920	0.11	0.77

Table 1b
Top winning cities and most improved cities.

Region	Top Efficient City	Most Improved City
Europe	Madrid	Lisbon (2012–2019)
China	Shenzhen	Changsha (2016–2019)
Japan	Kobe	Osaka (2016–2019)
Korea	Gwangju	Daejeon (2016–2019)
India	Hyderabad	N/A*
US/Canada	New York & Toronto**	Miami-Dade (2015–2019)

* Since many of the subway services in India began operating in 2018, we could not choose the most improved system. ** New York and Toronto were tied and share the title for the most efficient subway systems in North America.

4.1. Europe

The framework of urban rail services in Europe is highly fragmented, showing significant differences in urban rail service management and operations from city to city, even within the same Country. This makes the collection of homogenous data for benchmarking very difficult. As a matter of fact, the European case shows a lack of enough comprehensive data to conduct a detailed and complete analysis at the continental level, allowing the formulation of general conclusions valid for all metro systems in Europe.

Besides, data availability was the main criterion used for selecting the European cities to be included in this analysis.

Considering the aforementioned difficulties, we started by creating a preliminary dataset that include 25 European cities for which we collected data related to the metro service characteristics (e.g., network length, number of stations), supply (e.g., yearly produced vehicle-km) and demand (e.g., yearly transported passengers and passengers-km) for the years spanning 2000 to 2020.

However, in only 11 cases, income and expense statements were made available to the public. The city of Amsterdam (Netherlands) was successively added to this selection, increasing to 12 the cities for which we were finally able to retrieve audited reports (Amsterdam, Barcelona, Berlin, Bilbao, Bruxelles, Bucharest, Hamburg, Lisbon, London, Madrid, Prague and Rome).

Upon analyzing the income and expense statements along with other audited documents released by operators or public transport authorities, it emerged that these 12 cities could be distinguished into two fairly homogeneous groups within themselves.

- Group 1: Cities whose transit operator operates only metro services or operates different transport modes, providing data disaggregated by transport modes in its audited reports and balance sheet (Barcelona, Bilbao, Bucharest, Lisbon, London, Madrid).
- Group 2: cities whose transit operator operates multiple transport modes, providing only aggregated data in its audited reports and balance sheets (Amsterdam, Berlin, Bruxelles, Hamburg, Prague, Rome).

The decision was finally made to focus the analysis only on the first group of cities. Consequently, from the initial selection of 25 European cities, we were ultimately able to conduct the complete analysis only for six cities (Barcelona, Bilbao, Bucharest, Lisbon, London and Madrid) during the analysis period between 2012 and 2019.

These cities were chosen because official documents released by operators or transit authorities provided data with a sufficient level of detail to conduct the subway efficiency analysis.

The results of the analysis made on the six above-mentioned European metro systems indicate that Madrid emerges as the city whose metro system reached the highest Net VIP score in the last year of the analysis (2019). Alongside Bilbao, it maintains a consistently high level of productivity throughout the analysis period. Several factors may have contributed to the success of the Madrid Metro. Firstly, it ranks among the largest metro systems in Europe, and it is part of an extensive public transportation system that serves the entire city's metropolitan area. This network includes several interchange hubs that enable and facilitate the transfer of passengers between the different public transport modes available within the city, such as subways, light railways, urban and suburban railways and bus lines. The presence of a well-integrated transport system may foster the use of public transport. Indeed, Madrid is a city that has a significant share of public transport modes. Specifically, the Madrid metro is one of the busiest in Europe, showing an increasing trend of passengers transported and passenger-kilometers produced during the analysis period from 2012 to 2019. During the same period, the Madrid metro managed to reduce and stabilize soft costs and increase revenues coming from complementary businesses.

Lisbon exhibited the highest rate of Net VIP index improvement between 2012 and 2019 and was consequently recognized as the city with the most improved net VIP growth. Factors that contribute to the increase in Lisbon's metro productivity performances may include a constant rise in passenger volumes during the analysis period, which positively affects both ticket revenues and other revenue streams. On the other hand, during the same period, soft costs remained relatively constant, except for one year, which was marked by a significant increase. The interplay of these factors likely facilitated the enhancement of Lisbon's metro productivity.

In conclusion, all the European subway firms included in the analysis achieved satisfactory results during the analysis period spanning from 2012 to 2019: the net VIP variations are all included in the range between 0.73 and 1.10.

The analysis period closes in 2019 (the last year not impacted by the Covid-19 outbreak in Europe). The years 2020 and 2021 were not included in the analysis since the respective observations would have resulted in outliers due to COVID-19 and different lockdown policies. As a remark for future work, it is important to resume this analysis, including years 2022 and 2023, in order to assess whether subway firms' performances have changed from the years before the pandemic and to check which systems were able to return to performance levels comparable to the pre-pandemic ones. In so doing, it would be possible to identify the best cases that can be set as benchmarks for all metro operators.

Moreover, the need to harmonize at the continental level data and audited reports released by subway firms to enable transport authorities and regulators, public authorities and stakeholders to benchmark the performances of transport operators could also be considered a result of the analysis. This would help in the identification of business areas that need improvement and in setting goals for improving the efficiency with which transit operators conduct their business.

4.2. Japan

The urban railway transit system in Japan is globally recognized as one of the most complex and advanced. Japan's analysis covers two types of subway/railway systems with different geographic scopes: the first type covers only subway systems in nine cities of Sapporo, Sendai, Tokyo, Yokohama, Nagoya, Osaka, Kobe, Kyoto, and Fukuoka; and the second type covers both the subway systems and other urban rails in three metropolitan areas of Tokyo, Nagoya, and Osaka.

While looking at the nine cities and three metropolitan areas, Kobe achieved the highest net VIP score (1.11) in 2019, outperforming the other cities and regions. Thus, Kobe City is our winner in Japan. Compared to other cities and regions, Kobe has very limited resources despite its top productivity. The Japan analysis team conducted an interview with the Kobe Transport Authorities, the operator of the Kobe subway, and identified the following factors that may have contributed to the Kobe city's success.

- Strategic outsourcing: Kobe has effectively outsourced several operations to private contractors, reducing overall costs. They include outsourcing the operations of some railway maintenance tasks and four out of seven automobile depots.

- Personnel cost management: With an increasing number of retirements, Kobe has managed to rehire retired personnel at lower wages. This strategic move has helped in significantly cutting down the personnel costs. The comparison between Kobe City's payroll system and the outsourcing model reveals that outsourcing has led to considerable savings.
- Technological upgrades: The introduction of digital ticket only gates have led to a substantial reduction in maintenance costs, approximately halving them. Furthermore, the reduction in the number of ticket machines, due to decreased usage, has also contributed to cost savings of the machine maintenance.
- Infrastructure efficiency: The implementation of flexible tracks has extended their lifespan, ensuring long-term savings and reduced maintenance requirements. This investment in durable infrastructure has contributed to the overall efficiency and longevity of the subway system.
- Future-oriented initiatives: Kobe Transport Authorities are keen on further innovations, such as automating station operations and improving intercom responses. These future-oriented initiatives highlight Kobe's commitment to continuous improvement and adaptation to new technologies, ensuring sustained efficiency and service quality.

These strategies and innovations have improved the operational efficiency of the Kobe Municipal Subway while they also set a benchmark for other cities. The combination of strategic outsourcing, cost-effective personnel management, technological advancements, and efficient infrastructure utilization has positioned Kobe as a leader in urban railway transit systems in Japan. The city's ability to achieve high productivity with limited resources is a testament to its effective management and forward-thinking approach.

Based on the rate of net VIP index improvement for the Japanese sample period, Osaka was awarded as the city with the most improved Net VIP growth from 2016 to 2019. While the Metro Osaka system was owned and operated by the Osaka Transport Authority, a public body, until 2017, it was privatized in 2018. The privatization has provided the impetus for a more efficient and competitive operation. Under a privatization process, the company has implemented various measures to improve its financial sustainability and operational effectiveness. The Japan TF team also conducted an interview with the Osaka Metro, the operator of the Osaka subway, and identified the following factors as the factors that contributed to its success.

- Strategic cost management: Similar to Kobe, the Osaka Metro has actively pursued cost-saving measures, particularly through workforce optimization and streamlining operations. As preparing for privatization, the Osaka Metro reduced its workforce by about 700 employees since 2013. This reduction was facilitated by technological advancements and the elimination of redundant positions such as overlapping staff assignments at stations.
- Internalized operations: The Osaka Metro aims to internalize its operations as much as possible, preferring in-house management over outsourcing. For example, the maintenance of ticket vending machines and air conditioning are all performed by its own employees. Given the larger scale of its operations compared to Kobe subway, with a larger workforce, the company focuses on internalizing tasks to ensure better continuity and knowledge transfer within the organization.
- Population dynamics supporting revenue growth: Unlike many other regions in Japan, Osaka has not experienced a population decline. The increasing population along the railway lines has provided a steady stream of passengers, contributing to its sustained revenue growth. Efforts to align with the evolving transportation needs of the population have further enhanced the performance of Osaka Metro.
- Customer-centric strategies: The Osaka Metro's initiatives, such as coordinating last train schedules with other lines and lowering initial fares, demonstrate a commitment to meeting customer needs and enhancing the overall passenger experience. These customer-centric strategies have not only attracted more riders but also improved customer satisfaction levels.
- Focus on tourism: Osaka's significant increase in revenue from inbound tourism has played a crucial role in its improved performance. By catering to the needs of tourists and ensuring a seamless travel experience, the Osaka Metro has capitalized on this opportunity to boost ridership and revenue.
- Diversification of revenue streams: The privatized Osaka Metro began exploring new revenue streams, such as investment in real estate and restaurant business. While still in the early stages, these initiatives demonstrate the company's efforts to diversify its income sources and reduce dependence on traditional fare revenues.

Overall, the transition to privatization has been instrumental in driving the success of the Osaka Metro, allowing the company to implement strategic reforms, innovate, and adapt to a rapidly evolving urban transit landscape. The combination of strategic cost management, internalized operation, responsiveness to population dynamics, customer-centric strategies, a focus on tourism, and diversification of revenue streams has propelled Osaka Metro to achieve significant improvements in its net VIP index and emerge as a leader in urban railway transit in Japan.

4.3. China

Among the nine cities studied in China, based on the Net VIP (2019) results, the top three cities in terms of the efficiency of subway systems are Shenzhen, Beijing, and Changsha. The top three cities in terms of Net VIP growth are Shenzhen, Changsha, and Nanjing. Shenzhen outperforms other sample Chinese cities in Net VIP mainly because of its much larger "Other revenues output". Although Beijing and Guangzhou have higher passenger traffic volume (passenger-km), Shenzhen's subway system generates more revenues from retail, real estate and other derivative businesses. This is attributed to the innovative investment and financing mode adopted by Shenzhen, namely the TOD (Transit-Oriented Development) mode of comprehensive development. Taking Shenzhen's Metro Line 4 BOT project as an example, resource management, property management, engineering design and consulting, and the trading of building materials and supplies have been developed to finance the line construction and generate continuous revenues. Beijing has the largest-scale subway network in China. However, it also has very high soft costs and a large number of employees, driving down its net VIP.

4.4. South Korea

Because of the data availability issue, South Korean team has focused on the cities with subways served by the public sector. They also excluded the cities served mainly by the KORAIL national railway system because KORAIL does not provide separate accounting data separately for each subway system they serve.

As reported in Table 1a, out of the six cities included in South Korea, Gwangju achieved the highest net VIP score (1.78) in 2019 as compared to Seoul (1.30). Thus, Gwangju is the winning city in South Korea.

During the 2016–2019 period, Daejeon achieved the highest annual growth of the Net VIP score (9.07%), and thus, become the city with the most improved operating efficiency.

The South Korean team identified the following factors that contributed to the success of Gwangju and Daejeon.

- Rapid growth of traffic: Since Gwangju's relatively new subway route, which connects new towns and old cities, there has been high

traffic growth, which helps improve the operating efficiency of the system;
- Revenue diversification: Gwangju subway generates proportionally higher non-ticket revenue by increasing advertising and leasing station spaces. This helps improve revenue which increases outputs beyond traffic revenue (outputs);
- Continuous management improvement efforts: Gwangju transit authority made a conscious effort to improve efficiency by implementing the Gwangju urban railway-type ESG management system consistent with the city government policy.

The S. Korean TF team also identified the following factors that contributed to the City of Daejeon's success as the most improved Net VIP scores during the sample period (2016–2019).

- Technological upgrades: Maintenance and labor costs were reduced by establishing a scientific safety management system by introducing new technologies of the 4th Industrial Revolution. In addition, the introduction of new technologies has contributed to quickly improving and achieving management efficiency by creating various revenue sources.
- Customer-centric strategies: Daejeon made an effort to attract passengers to subways by using special customer-tailored marketing and by collaborating with the local communities in order to improve customer satisfaction.
- Management innovation: By adopting a consumer-centered communication budget system, the city was able to improve passenger opinions on subway services related to safety and service level. In addition, the city's effort on management innovation contributed to improving operating efficiency by strengthening job-related training for employees.

4.5. India

For the benchmarking purpose the Indian TF team have analyzed subway systems in seven cities: New Delhi, Bangalore, Hyderabad, Lucknow, Chennai, Bangalore, Jaipur, and Kochi. A few cities such as Calcutta and Mumbai where subway lines began operating from 2019 were not considered. The future benchmarking efforts will include more cities including these cities.

Among the seven cities from India, Hyderabad achieved the highest net VIP score (1.98) in 2019 followed by Lucknow (1.76). Thus, Hyderabad is our winning city from India for year 2019. It is noted that Hyderabad Metro was also chosen as the Best Urban Mass Transit Project by the Government of India in November 2018.

Although suburban rail systems in major Indian cities have been operating since the British period, the subway systems are very recent additions. Only Calcutta and Delhi had subway systems before 2010. Out of the seven cities considered, Lucknow, Hyderabad and Kochi started operation in 2018, Jaipur in 2017, and Chennai in 2016. Some of these systems had only one line of 10–15 km long when at the time of inauguration. Therefore, we have not selected the city with the most improvement from India for 2019.

4.6. Canada – USA

Among the six cities included from USA and Canada, Toronto (1.29) and New York (1.26) have achieved the highest net VIP scores in 2019, as compared to Miami-Dade (base 1.0). Therefore, we chose these two cities subway systems as the top efficient subways in North America. Based on the improvement rate of net VIP index for the sample period, Miami-Dade was chosen as the city with the most improved Net VIP growth.

Since the price index for computing the soft input quantity index for New York (USA) and Toronto (Canada) did not have the same base year, it was impossible to fully account for the difference in the soft input price index between New York and Toronto. Our task force went back to start calculating and updating the soft cost input price index of Toronto and New York beginning with the year 1990 price index (hoping that assuming Toronto and New York had the same soft input prices in 1990 and updating the respective price levels to 2019 would ameliorate the situation of incomparable price index levels for New York and Toronto). This is not a perfect solution, but in the absence of bilaterally linkable Canadian and US city price indices, this is probably the best solution for us. For this reason, using an imperfect soft cost input price index, we chose to make both New York and Toronto the winners (although our Net VIP index scores show New York is 0.03 percentage points lower (1.26) than Toronto (1.29).

Toronto's TTC system is the 3rd largest in North America after New York and Mexico City. With a farebox recovery ratio of almost 70% during pre-pandemic years, Toronto's TTC system is the least subsidized system in Canada and one of the least subsidized systems in North America. A total of 75 stations supports 76.5 km of rail, resulting in a high density of stations at just over an average of 1 km distance between stations. In turn, the ease of accessibility attracts a high demand of transit users. The combination of high station density servicing a stable and mature transit demand pattern allows the system to be operated at a relatively high efficiency.

New York is an outlier in terms of transit use in the car-dominated US society. Compared to any other city in the US, New York has a very robust subway system that covers a large segment of New York City. New York's heavy rail system has 472 stations, whereas the combined total of other heavy rail stations in the US is a little over 500, which indicates the extent of the transit service in New York (MTA, 2020). Individual use patterns are also important, as 56% of the population of New York City uses public transit to commute to work (New York Public Transit Association), whereas the national average is 5%. Thus, it is expected that New York would achieve the highest net VIP score compared to other systems.

The most improved status of Miami is likely due to increased funding. Miami's metro rail system is an elevated 25-mile-long system that has not gone through much expansion since its inception in 1980. Compared to all the other heavy rail systems of North America in this report, the Miami system is very small. The system is going through a decline in ridership during the analysis period and, consequently, a decline in fare revenue. However, we do see an enormous increase in revenue from other sources (e.g., subsidy on transit operation) as well as a small share of soft cost. These two factors may have contributed to the higher productivity index score for this agency during the analysis period and given it the most improved status compared to the other analyzed systems in North America.

5. Significance and policy implications of this WCTRS subway TF efficiency benchmarking project

5.1. Significance of the TF efficiency benchmarking project results

The WCTRS Task Force's inaugural subway efficiency benchmarking work represents a pivotal advancement in the field of subway and urban rail efficiency analysis. This initiative is significant for several reasons.

- Comprehensive Analysis: By employing the net Variable Input Productivity (Net VIP) index computed by the translog multilateral indexing method, this study provides a robust, transitive comparison of the subway operating efficiency of the cities included in our inaugural benchmarking study. This transitive productivity (efficiency) index ensures that the benchmarking results are comparable across cities as well as over time.
- Global Scope: Covering subway systems in Asia, Europe, and North America, this benchmarking exercise offers a truly international perspective. It computes an aggregate index summarizing the Net Variable Input Productivity which indicate the operating efficiency

of each city included in our sample, and also allows to compare changing operating efficiency over time. The results help us to attempt to identify best practices and the areas needing improvement across different operational contexts.

- Focused Metrics: By concentrating on the Net Variable Input Productivity (Net VIP), particularly labor and soft input productivity, the study zeroes in on the critical aspects of operational efficiency. This focus helps transit authorities understand how effectively they are utilizing their variable resources and identify specific areas for optimization.
- By focusing on the Net VIP, the aggregate indicator of the operating efficiency it allows us to make comments on success or failure of the current management team of the subway firms or authorities.
- Actionable Insights: The derived Net VIP scores, adjusted for the factors beyond the transit management's control, and thus, provide actionable insights for urban transit authorities. These refined metrics facilitate targeted interventions aimed at improving efficiency and overall system performance.

5.2. Urban transport management and policy implications

The findings from this benchmarking study have profound implications for urban transport policy.

- Policy Formulation and Resource Allocation: The benchmarking results offer empirical evidence that can guide the policy and management strategy formulation and resource allocation. Policymakers can identify which systems are underperforming, and management can prioritize areas of investment to enhance efficiency most effectively.
- Adoption of Best Practices: The identification of top-performing subway systems provides a benchmark for others to aspire to. Policymakers and transit authorities can study these exemplars, adopting and adapting their best practices to local contexts to improve efficiency.
- The winning cities in each country/continent should be closely examined in order to find the reasons for the best operating efficiency results.
- Strategic Planning: Insights from the benchmarking can inform strategic planning efforts. By understanding the factors contributing to the higher efficiency of the winning cities' subway management and operations, urban transit authorities can develop long-term strategies that align with broader goals such as sustainability, economic viability, and enhanced public service.
- Technology and Innovation: The study underscores the importance of integrating advanced technologies and innovative practices in subway operations, such as smart preventive maintenance systems and data-driven decision-making.
- Performance Monitoring and Accountability: Regular benchmarking, as advocated by the WCTRS Task Force, promotes continuous performance monitoring and accountability. It encourages transit systems to maintain high standards of efficiency and fosters a culture of continuous improvement.
- Above all, the WCTRS Task Force members who volunteered their time and effort hope to raise strong awareness of efficiency and productivity as an important goal for managing and operating the subways and other city transit systems.

This inaugural benchmarking work by the WCTRS Task Force not only sets a new standard for evaluating subway efficiency but also provides a critical tool for policymakers and transit authorities worldwide. By leveraging these insights, urban transit systems can enhance their operational performance, ultimately leading to more sustainable, efficient, and effective public transportation networks.

Special acknowledgement

The Task Force also gratefully acknowledge the following TF Advisory Committee members for their contribution: Prof. Ennio Casetta (Italy), Prof. Hirono Kato (Tokyo University), Prof. Meng Li (Tsinghua University, Beijing), Prof. Daniel Sperling and Prof. Kari Watson (both University of California – Davis).

Acknowledgement for research assistants

Tae Hoon Oum also would like to acknowledge the research assistant work for the Task Force conducted by the following WCTRS president's assistants for since the initiation of the Task Force in January 2020: 2020–2021: Unseo (Erin) Lee; Anatasiia Nosach; Niraj Hajani; 2021–2022; Jinyun Su; Carl Chen; Xianyi Liu; Runqi Zhu; 2022–2023; Minkyeong Baik (grad student); Sejin Ahn (grad student); Siran Li 2023–2024; Ziyuan Wang (grad student); Elaine (Hanji) Wei; Hoa Ton That Summer 2024; Subhan Chiluwal (she is currently helping me to deal with our TP paper).

CRediT authorship contribution statement

Tae Hoon Oum: Writing – original draft, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization. **Kun Wang:** Formal analysis, Investigation, Validation, Writing – review & editing. **Chunyan Yu:** Conceptualization, Investigation, Validation, Writing – review & editing. **Pierlugi Coppola:** Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft. **Luigi Castagna:** Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. **Hironori Kato:** Data curation, Formal analysis, Investigation, Validation, Writing – original draft. **Yiping Le:** Data curation, Formal analysis, Investigation, Validation, Writing – original draft. **Xiao Luo:** Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. **Suhui Gan:** Data curation, Investigation, Validation, Writing – review & editing. **Kyungtaek Kim:** Data curation, Formal analysis, Investigation, Writing – original draft. **Woojin Kim:** Data curation, Formal analysis, Investigation, Writing – review & editing. **Gopal Patil:** Formal analysis, Investigation, Validation, Writing – original draft. **Vinayak Gaur:** Data curation, Formal analysis, Investigation, Writing – review & editing. **Giovanni Circella:** Formal analysis, Investigation, Supervision, Validation, Writing – original draft. **Hossain Mohiuddin:** Data curation, Formal analysis, Investigation, Writing – review & editing. **Maria Carolina Lecompte:** Data curation, Formal analysis, Investigation, Writing – review & editing. **Clark Lim:** Data curation, Formal analysis, Investigation, Writing – review & editing.

Acknowledgements

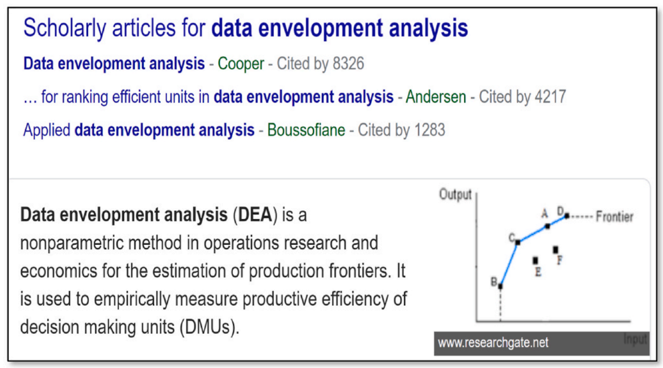
Tae Hoon Oum (President of the WCTR Society and the Task Force chair) acknowledges gratefully the insight grant support of the Social Science and Humanities Research Council of Canada. Prof. Oum also acknowledges the support of the WCTR Society's Steering Committee for USD10,000 initiation budget for this Task Force in order to pay honorarium to the TF Secretary General (Clark Lim) and tiny appreciation stipends to the graduate students who volunteered their time for several years to complete this inaugural TF report.

The Japan TF team would like to give a special gratitude to Ryoma Sugai (Shibaura Institute of Technology) for helping to collect the data of Japan's urban subways/railways.

Appendix A. Step by Step illustration of how DEA may produce intransitive efficiency index

In order to clearly explain this, why we cannot use the DEA based methods for our efficiency ranking of the cities, we have produced an illustrative (not real) example for ranking U.S. cities efficiency ranking by a simple DEA method. For example, the relative efficiency ranking of New York, Chicago, Miami, San Francisco may depend on whether we include Houston or not in our data set. This is illustrated graphically below (note: Cooper who wrote this Wikipedia article is one of the three original mathematicians who introduced the DEA method in 1978 article (Charnes, Cooper and Rhodes, “Measuring the efficiency of decision-making units,”, European J. of Operations Research 2, 429–444)(Cooper et al., 2006).

To start, the DEA description in Wikipedia with data points (DMUs or cities) A to F.



Let us use similar as the above Cooper’s exposition of how the DEA measures the efficiency scores in order to explain why DEA may produce different ranking of cities which are not on the efficiency frontier. The rankings depend on which cities form the efficiency frontier.

Let us say, in the first set of US cities, we included:

A: New York, B: Seattle, C: Atlanta, D: Chicago, E: Miami, F: San Francisco.

And, let us also say that we found the following DEA efficiency results as shown in the graph below. The results shows that Seattle, Atlanta, New York, Chicago are on efficiency frontier (meaning 100% efficient); but then, E (Miami) and F (San Francisco) are below the efficiency frontier. But if we draw perpendicular line from E(Miami) and F (San Francisco) to the efficiency frontier: i.e., Point E to E′ and Point F to F’, Miami and San Francisco are almost equal distance to the efficiency frontier, meaning that Miami and San Francisco are almost similarly inefficient (Figure A).

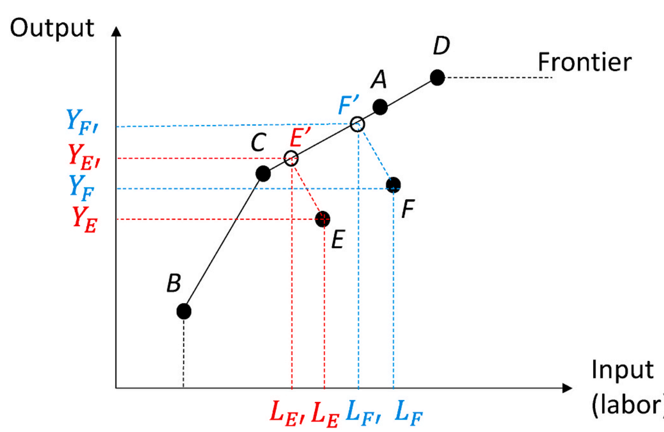


Fig. A. Efficiency Frontier Plot Based on a Set of DMUs (Cities).

Now, let us add another city, G (Houston) into our data set (Figure B). Then, it turns out the added Houston (point G) changes the efficiency frontier as shown the graph below.

The addition of G (Houston) in our data set makes the following very significant changes in the relative efficiency ranking of the cities as can be seen in the graph below.

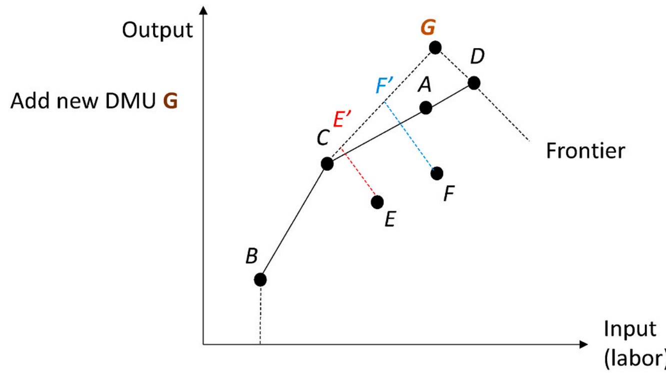


Fig. B. Efficiency Frontier Plot Altered by the Introduction of a new DMU (City).

First, the addition of G (Houston) removes New York(A) and Chicago(D) from the efficiency frontier, and thus make them inefficient cities. Furthermore, Chicago (D) become further away from the new efficiency frontier made up of now BCGD than New York (A), making Chicago less efficient than New York. Second, E(Miami) is now closer to the new efficiency frontier than F (San Francisco), making San Francisco less efficient than Miami. **Therefore, the DEA efficiency ranking is not 'transitive', which makes it impossible to use.**

Data availability

The authors do not have permission to share data.

References

Andrikopoulos, A.A., Loizides, J., Prodromidis, P., 1993. Fiscal policy and political instability: the long-term correlation between debt and election results. Publ. Choice 75 (3), 263–279.

Boame, A.K., 2004. The technical efficiency of Canadian urban transit systems. Transport. Res. E Logist. Transport. Rev. 40 (5), 401–416.

Castagna, L., Lobo, A., Coppola, P., Couto, A., 2024. Benchmarking the efficiency of European metros from a production perspective. Research in Transportation Business & Management 53, 101102. https://doi.org/10.1016/j.rtbm.2024.101102.

Caves, D.W., Christensen, L.R., Diewert, W.E., 1982. Multilateral comparisons of output, input, and productivity using superlative index numbers. Econ. J. 92 (365), 73–86.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision-making units. Eur. J. Oper. Res. 2 (6), 429–444.

Coelli, T., Rao, D.S.P., O'Donnell, C.J., Battese, G.E., 2005. An Introduction to Efficiency and Productivity Analysis. Springer.

Cooper, W.W., Seiford, L.M., Tone, K., 2006. Introduction to Data Envelopment Analysis and its Uses: with DEA-Solver Software and References. Springer.

Graham, D.J., 2008. Productivity and efficiency in urban railways: parametric and non-parametric estimates. Transport. Res. E Logist. Transport. Rev. 44 (1), 84–99.

Lan, L.W., Lin, E.T., 2002. Measuring technical and scale efficiency in rail industry: a comparison of 85 railways using DEA and SFA. Traffic Transport. 21, 75–88.

MTA, 2020. MTA subway and bus ridership for 2019. https://new.mta.info/agency/new-york-city-transit/subway-bus-ridership-2019.

New York Public Transit Association. (n.d.). Public Transit Facts. Retrieved May 25, 2024, from https://nytransit.org/resources/public-transit-facts.

Oum, T.H., Waters, W.G., Yu, C., 1999. A survey of productivity and efficiency measurement in the rail transport industry. J. Transport Econ. Pol. 33 (1), 9–42.

Viton, P.A., 1997. Technical efficiency in multi-mode bus transit: a production frontier analysis. Transp. Res. Part B Methodol. 31 (1), 23–39.

Zhu, J., 2003. Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets and DEA Excel Solver. Springer.