International Scholarly Research Network ISRN Civil Engineering Volume 2012, Article ID 984758, 13 pages doi:10.5402/2012/984758

## Research Article

## Analysis of the Technical Efficiency of Urban Bus Services in Spain Based on SBM Models

#### Pablo Jordá,<sup>1</sup> Rocío Cascajo,<sup>2</sup> and Andrés Monzón<sup>2</sup>

<sup>1</sup> Universidad Politécnica de Madrid, Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos, 28040 Madrid, Spain <sup>2</sup> TRANSyT-Transport Research Centre, Universidad Politécnica de Madrid, 28040 Madrid, Spain

Correspondence should be addressed to Pablo Jordá, pjordan@caminos.upm.es

Received 29 February 2012; Accepted 8 May 2012

Academic Editor: J. D. Nelson

Copyright © 2012 Pablo Jordá et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

During the first decade of the new millennium, fueled by the economic development in Spain, urban bus services were extended. Since the years 2008 and 2009, the root of the economic crisis, the improvement of these services is at risk due to economic problems. In this paper, the technical efficiency of the main urban bus companies in Spain during the 2004–2009 period are studied using SBM (slack-based measures) models and by establishing the slacks in the services' production inputs. The influence of a series of exogenous variables on the operation of the different services is also analyzed. It is concluded that only the 24% of the case studies are efficient, and some urban form variables can explain part of the inefficiency. The methodology used allows studying the inefficiency in a disaggregated way that other DEA (data envelopment analysis) models do not.

#### 1. Introduction

The flexibility in the management offered by buses is the main cause of their success compared to other modes, such as rail. Urban public transport offered by buses allows for the completion of different activities, be it work, education, shopping, leisure, and so forth. The services are, then, a fundamental pillar of society. This leads, in the main Spanish cities, to services being offered by one or a few public companies, without any competition from other companies. This lack of competition can bring about a complacency in management, or the need to improve the efficiency of the services.

The evaluation of the operational efficiency of an urban bus service is a complicated task. One must identify those elements which are decisive in the operating of the system, in order to subsequently define a model which, on the base of these decisive elements, reflects its functions in a simplified manner. In the evaluation of a service, therefore, one should consider the consumption of resources (inputs) when obtaining the results or production (outputs).

Beginning with these ideas, and aided by the information supplied by the database of the Spanish Metropolitan Mobility Observatory (MMO) [1–8], this paper will discuss the problem of studying the technical efficiency of urban bus services, considering the greater number of evaluation criteria related to the socioeconomic context and the evolution of public transport networks [9].

# 2. Analysis of Technical Efficiency with the DEA Method

The Data Envelopment Analysis (DEA) is a linear programming method whose object is to calculate the relative efficiencies of a group of decision-making units (DMUs). These DMUs can be, as in the case of this investigation, urban bus companies or, more specifically, a bus company in a given year.

The DEA is one of the existing frontier methods. Frontier methods are those in which an efficiency frontier is used to classify the different DMUs. The efficiency frontier is based on real observations and only the cases of best practices belong to it. All DMUs that are not on the frontier are considered inefficient.

When assessing the efficiency of urban bus companies, in addition to the DEA method, other methodologies have been used. The stochastic parametric methodologies has been

TABLE 1: Frontier methodologies.

Functional		Measurement error
form	Deterministic	Stochastic
Parametric	Corrected OLS, and so forth	Frontiers with explicit assumptions (exponential, half normal, etc.) for the technical efficiency distributions
Nonparametric	FDH, DEA-type models, and so forth	Resampling, chance- constrained programming, and so forth

Source: [10].

widely used [9, 10], as well as other frontier methodologies (see Table 1).

CCR and BCC models were the first DEA models to be formulated and are explained below. All subsequent DEA models have been developed from them.

Given the group of DMUs K, the technical efficiency of the DMU  $k_0$  is defined as the ratio of the weighted sum of its n outputs  $y_{jk0}$  and the weighted sum m of its inputs  $x_{jk0}$ , all expressed as positive values. As outputs, in the case of buses companies, it can be used vehicle-km or pax-km. As inputs, the variables most commonly used are number of buses, number of workers, companies costs, and infrastructure variables The problem of the fractional programming, known as the CCR model [14], is expressed as follows:

Maximize 
$$h_0 = \frac{\sum_{j=1}^{n} w_j y_{jk0}}{\sum_{i=1}^{m} v_i x_{ik0}}$$

Subject to:

$$\frac{\sum_{j=1}^{n} w_j y_{jk0}}{\sum_{i=1}^{m} v_i x_{ik0}} \le 1 \quad w_j, v_i > 0,$$

$$i = 1, \dots, m; \ j = 1, \dots, n; \ k = 1, \dots, K,$$
(1)

where the subindex *i* numbers the inputs, the *j* outputs, and the *k* the DMUs.  $w_j$  and  $v_i$  are weights.

The first restriction of this problem indicates that the efficiency value of each unit k, in function of the given weights, should be, at most, one. If one observes the objective function, its end is to maximize the relationship between a virtual output, composed of all considered outputs, and a virtual input that is also a compilation of various inputs. The second condition requires that the weights of the outputs and inputs  $w_j$  and  $v_i$  be greater than zero. This restriction seeks to prevent the outputs or inputs of the DMUs from varying with total freedom. In the analysis, the unit  $k_0$  is inefficient with respect to the rest of the DMUs if it does not reach a ratio of outputs to inputs equal to one.

In order to resolve the problem expressed in (1), the model is transformed and expressed as an equivalent linear form [14, 15]. The resulting input-oriented model enables evaluations of DMUs in constant returns to scale (CRS)

situation. It is named as envelopment model [16] and expressed in the following form:

Minimize 
$$h_0 = \theta_0$$

Subject to:

$$\sum_{k=1}^{K} \lambda_k x_{ik} = \theta_0 x_{ik0}, \quad 0 \le \theta_0 \le 1,$$

$$\sum_{k=1}^{K} \lambda_k y_{jk} = y_{jk0},$$

$$\lambda_k \ge 0,$$

$$i = 1, \dots, m; \ j = 1, \dots, n; \ k = 1, \dots, K.$$
(2)

The first restriction of this problem indicates that the number of inputs used by the DMU  $k_0$ , multiplied by the efficiency factor  $\theta_0$ , should be equal to the number of inputs used by the reference unit of the DMU  $k_0$  at the frontier, which is composed of other DMUs of the *K* group. That is to say, its capacity to transform inputs into outputs will be equal to or less than the capacity of the reference unit. The third restriction indicates that the reference unit at the frontier should produce the same number of outputs as the  $k_0$  unit.

The intensity of the efficiency factor can, therefore, be used to determine the minimum quantity of use of inputs that must be proportionally reduced for the  $k_0$  unit to be efficient. The efficiency factor  $\theta_0$  has values of between zero and one, both included. The units that contribute to the construction of the reference unit at the frontier will have weight values  $\lambda_k$  different from zero.

CCR models (2) make it possible to evaluate a group of DMUs that have constant returns to scale. However, there are DMUs with increasing returns to scale, others decreasing, and others constant. Therefore, in [15] they reformulated the CCR model to allow for variable returns to scale, thus defining a new model, the BCC. This model is expressed in the following manner:

Minimize 
$$h_0 = \theta_0$$

Subject to:

$$\sum_{k=1}^{K} \lambda_k x_{ik} = \theta_0 x_{ik0}, \quad 0 \le \theta_0 \le 1,$$

$$\sum_{k=1}^{K} \lambda_k y_{jk} = y_{jk0},$$

$$\sum_{k=1}^{K} \lambda_k = 1,$$

$$\lambda_k \ge 0,$$

$$i = 1, \dots, m; \quad j = 1, \dots, n; \quad k = 1, \dots, K.$$
(3)

TABLE 2: Advantages and disadvantages of DEA.

Advantages	Disadvantages
Simultaneous analysis of outputs and inputs	Ignores the effect of exogenous variables on the operation
It is not necessary, a priori, to define the frontier form	Ignores statistical errors
Relative efficiency, compared to the best observation	Does not say how to improve efficiency
Need no information on prices	Difficult to perform statistical tests with the results

Source: [11-13].

It is an input-oriented BCC model [11]. The only difference from the CCR model (2) is that a new condition is included: the summation of the weights  $\lambda_k$  is equal to one. This makes it possible to evaluate the variable returns to scale (VRS) in the DMUs and guarantees that each DMU is compared only with those of a similar size [17] to keep them from being considered inefficient simply because of the differences of scale between DMUs.

Finally, in Table 2 are shown the main advantages and disadvantages of the DEA method over other methods, like the stochastic parametric methodologies.

The lack of information on prices in the database used is one of the reasons why in this study is used the DEA method instead of the stochastic parametric methodologies.

#### 3. Slacks and SBM Models

In some cases, and in order for a DMU  $k_0$  to reach efficiency, the proportional reduction of inputs achieved using the  $\theta$  factor is not enough, as has been indicated in models (2) and (3). An additional reduction of input is necessary, or an additional output increase, as it is not proportional (radial). These complementary adjustments to the inputs ( $s_i^-$ ) (excess of input) or to the outputs ( $s_j^+$ ) (output shortfall) are together known as *slacks* [18]. Figure 1 clarifies these concepts.

Thus, given a frontier, the efficiency of the DMU M is evaluated, projecting it toward the frontier in a radial manner in the direction OM, and does not produce a problematic situation, as the slacks of the projection M' are zero. No additional reduction of inputs is necessary to achieve efficiency. For the DMU N, projected radially toward the frontier, it can be observed that its projection N' must consume the same quantity of input 1 as the DMU B, but it consumes a greater quantity of input 2 (segment BN'), for which it cannot be considered equally efficient to the DMU B.

So, it is not acceptable to measure the efficiency of a group of DMUs only by means of the radial factor  $\theta$ . The slacks must also be kept in mind when evaluating the efficiency value of the DMUs.

Some modifications must be introduced to the linear problem of the CCR model (2) to be able to evaluate the slacks. For a case of orientation to the input, once the slacks have been included, the resulting model is that proposed

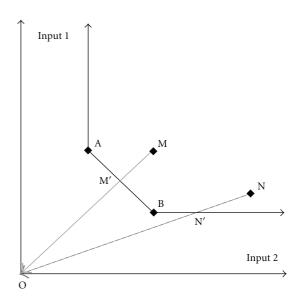


FIGURE 1: Radial projection and slacks. Input-input representation. Source: based on [19].

in [15]:

Minimize 
$$h_0 = \theta_0 - \varepsilon \left( \sum_{j=1}^n s_j^+ + \sum_{i=1}^m s_i^- \right),$$
  
Subject to:

$$\sum_{k=1}^{K} \lambda_k x_{ik} + s_i^- = \theta_0 x_{ik0}, \quad 0 \le \theta_0 \le 1,$$

$$\sum_{k=1}^{K} \lambda_k y_{jk} - s_j^+ = y_{jk0},$$

$$\lambda_k, s_i^-, s_j^+ \ge 0,$$

$$\varepsilon > 0, \text{ very small value}$$

$$i = 1, \dots, m; \quad j = 1, \dots, n; \quad k = 1, \dots, K.$$
(4)

Thus, the problem (4) also gives as results values of the efficiency factor  $\theta_0$  between zero and one. Additionally, and differing from the previous models, it provides the values of the vectors of slacks  $s_j^-$  and  $s_j^+$ , which can be equal to or greater than zero. In subsequent models, the parameter  $\varepsilon$  ceases to be utilized.

Based on several works [12, 18, 20–24], many models have been developed that include slacks when assessing the efficiency factor; these are known as *Slacks-Based Measures* or SBM.

If one maintains the original criterion that a DMU is efficient if it is on a frontier ( $\theta_0 = 1$ ), independently of the values of its slacks, there can be different types of efficiency in the DMUs. In [25] it is described up to three different categories: efficient DMUs, which fulfill the requirement that  $\theta$  is equal to one and its slacks are zero (DMU M', in Figure 1); extremely efficient DMUs, with the same characteristics as an efficient one, the part of the frontier that is asymptotic to the x and y axes originates with them (DMUs A and B, in Figure 1); weakly efficient DMUs, as, if the  $\theta$  value is equal to one, at least one of its inputs is susceptible to reduction to improve production (DMU N', Figure 1), as these are DMUs that belong to the part of the frontier that is asymptotic to the x and y axes (extensions).

Therefore, it seems more correct to use SBM models rather than the original radial models, as they enable a better and more adjusted calculation of efficiency.

#### 4. Case Studies

Spanish Law 7/1985, of the Basis of Local Government, considers the urban public transport as a public service that municipalities with more than 50,000 inhabitants have the obligation to provide. The law gives municipalities the power to the planning and management for the service within its territory. The most common management models are the direct management through the creation of a 100% public commercial company, typical of large cities, and the indirect management, through an administrative concession, modality widespread among medium-sized and small cities [26].

The sector of urban bus services in Spain comprises 189 companies [27], of which only 32 are public companies, usually in large cities. The remaining 157 companies are private companies operating in small towns through an administrative concession. All the companies invoice approximately 1,200 million  $\in$  a year (0.1% GDP, 2009), employing about 22,200 workers, 16,000 in public companies. The total fleet is about 7,200 vehicles, 4,800 of which belong to public companies.

The models described in the previous section will here be applied to obtain the technical efficiency of 18 urban bus companies in Spain. The data used in this study is drawn from the Spanish Metropolitan Mobility Observatory (MMO (the MMO is an initiative coordinated by the Ministry of the Environment and Rural and Marine Affairs, in collaboration with other national organizations in Spain, with the participation of various Spanish public transport authorities (PTA) as well as other entities and organizations related with metropolitan public transport)) and the period analyzed covers the years 2004 to 2009 [1–8]. Complete historical data series of the period 2004–2009 are available. The data from the MMO is provided by the urban bus companies and the National Statistics Institute.

Spanish metropolitan areas have seen significant population growth during the period analyzed. It has increased on average at the rate of 9.4% for 2004 to 2009, with much of this growth being absorbed by the outer metropolitan rings as opposed to the urban centers (only 3.1% of growth), thus reflecting ongoing dispersion trends. The consequence of this process favors the use of cars and also complicates the provision of public transport services.

Regarding the economic activity, the unemployment rate in the analyzed cities ranges from values of around 6– 15% (2005, according to the National Statistics Institute) before the crisis to values that in 2009 exceeded 20%. One consequence of this situation is a decline in the journey to work, thus, a decrease in the use of the different modes of public transport, including the urban bus. The lessened demand for bus trips, without a reduction in the amount of services offered, caused a major financial imbalance in the bus companies and as a result a greater necessity for subsidies on the part of the public administrations.

Despite the economic crisis, the GPD has increased in most of the areas during the years analyzed here. This is in line with the evolution of the motorization index, which reached values of 500 vehicles per 1,000 inhabitants, bringing it closer to that of other European nations.

Table 3 shows the variation of some indicators of the different bus services, both urban buses and metropolitan buses, in the analyzed cities in the period 2004–2009. With regard to the demand indicators, journeys and passenger-km, they suffer a decrease in most cases, due to the reduction in the work-related mobility because of the fall in the number of jobs during the crisis.

The supplies of the bus services are measured by vehiclekm and places-km. For urban buses, in most cases an increase in supply can be seen, though there is a certain number in which there is a decrease. For the metropolitan services, important increases can be observed in many cases. This increase in the supply is due to the increase in the population of the metropolitan ring. For the moment, the economic crisis does not appear to have affected the supply of services: even while demand decreased, the quantity of services offered increased.

Lastly, Table 3 shows the variation of infrastructure indicators, in terms of route length and number of stops, which, except for the rare exception, demonstrate increases for both urban and metropolitan buses. This fact reinforces the idea that the services have increased despite the crisis and the decrease in demand.

#### 5. Model Design for Determining the Efficiency Frontier

The free version of the software used in this study, MaxDEA 5.0 [28], allows one to use a radial or SBM approach and choose the desired orientation (output, input, or other mixed) as well as the returns to scale (variable or constant, or other alternatives).

As for the orientation, given that it is the production efficiency that is being evaluated (measured in vehicle-km), the logical choice is to maintain as constant the output and allow the model to vary the values of the inputs used in its production. Thus, an input-oriented approach is used. This approach is the most commonly used in other studies [11, 29–31].

Due to the difference in size of the bus companies analyzed here, it is necessary to use variable returns to scale (VRS). According to the literature, by using the constant returns to scale (CRS), one would obtain higher efficiency values than those obtained when using variable returns to scale (VRS).

Once the type of model to be used to obtain the efficiency frontier has been determined, the following step is the choice of the variables to include in the calculation (outputs and

Metropolitan	Metropolitan area Journeys*	Journeys*	Pax-km*	Journeys**	Veh-km*	Places-km*	Veh-km**	Places-km**	Routes	Number of	Routes	Number of
area population		(%)	(%)	(%)	(%)	(%)	(%)	(%)	length* (%)		length** (%)	stops** (%)
	Madrid	-10.0	-0.5	-10.4	3.8	0.1	20.2	9.3	20.4	17.8	11.5	24.1
	Barcelona	-5.6	10.1	3.5	1.9	-3.7	23.3	16.7	7.2	n.a.	40.3	1.6
>1 million	Valencia	-9.3	-25.8	-10.4	-5.0	-2.5	-8.5	-4.2	1.5	0.7	2.2	1.9
inhab.	Murcia***	-21.5	-10.0	n.a.	-2.6	-11.0	n.a.	n.a.	7.9	0.7	n.a.	n.a.
	Sevilla	-7.3	-6.9	-8.0	3.0	8.1	n.a.	18.5	-0.9	-9.2	33.4	19.4
	Asturias (Oviedo)	2.6	2.4	n.a.	16.6	21.4	n.a.	n.a.	52.2	16.6	n.a.	n.a.
	Málaga	8.9	-1.8	41.8	9.4	34.6	15.9	30.3	7.5	-1.3	26.8	87.5
Betrizeen 05	Mallorca***	0.0	0.0	-1.5	16.0	13.3	11.5	13.3	12.9	20.5	n.a.	n.a.
and 1 million	Gran Canaria***	-10.4	n.a.	-13.2	-4.9	0.5	-2.3	-3.5	14.1	9.5	16.3	-0.7
inhab.	Zaragoza	10.5	10.5	10.8	13.7	21.5	n.a.	n.a.	30.0	27.0	8.2	20.0
	oipuzkoa رومی	71	71	2 1	- 0	10	\$	2	د د	0 1	\$	\$
	Sebastián)***	1.0	1.0	1.1	1.0	1.0	11.4.	11.4.	C.7	1*0	11.4.	11.4.
	Granada	6.5	n.a.	8.4	9.4	n.a.	17.0	n.a.	5.3	n.a.	6.8	n.a.
	Alicante	-5.9	-3.8	-6.3	4.0	-4.4	5.7	-3.2	2.1	1.0	7.8	19.3
<0.5 million	Pamplona	10.6	12.2		22.8	25.8			12.7	-16.4		
inhab.	Vigo***	-4.9	-8.1		0.8	0.8			0.0	7.0		
	La Coruña	-1.7	-1.6		5.6	2.0			0.0	0.0		
Sevilla, Granada, and Alican *Only main city. **Whole metropolitan area. ***Changes using only avail Source: based on [1–8].	Sevilla, Granada, and Alicante: trips, not journeys. *Only main city. **Whole metropolitan area. ***Changes using only available data Source: based on [1–8].	journeys.										

TABLE 3: Changes in supply and demand indicators of bus services (%). Period 2004–2009.

### ISRN Civil Engineering

inputs). In the literature considered here, two variables are used as outputs: vehicle-km, which is the supply indicator, and pax-km, which reflects the service demand [29, 32, 33]. However, the first of the two indicators is the most widely used, though authors that have the second at their disposal use it as well and compare the results [30, 32–36].

As for the inputs that are commonly employed, variables are selected that can be used in a Cobb-Douglas [37] formula: work, capital, and consumption/cost variables. As the variable reflecting the work factor, the number of workers is commonly used, but this variable is not available in the database used here. As the variable for the capital factor, the size of the fleet is used. And finally, the operating costs of the urban bus companies are used as the variable to represent consumption.

In this study, the desired approach will differ from that proposed by others, including variables that reflect the infrastructure used in the provision of services. Some authors use variables related to infrastructure, such as routes length, distance between stops, or network-km [29, 32, 38] but include them in the models as exogenous variables. In [39] it is the first time that routes length is introduced as an input in the models. Subsequently, [36, 40] also use infrastructure variables (routes length and number of stops) as input.

Finally, an infrastructure variable and some of the classic variables from the Cobb-Douglas models are used for the final formulation of the efficiency:

$$\theta = \frac{f(uVK)}{f(v_1RL, v_2NV, v_3CC)},$$
(5)

where VK is the vehicle-km, RL is the routes length, NV is the number of vehicles, CC is the company costs, and u,  $v_1$ ,  $v_2$ , and  $v_3$  are model coefficients.

In Table 4 the different variables used in (5) and in the rest of the study are defined as well as their minimum and maximum values and their units. It is indicated also the type of variable: output, input, endogenous, or exogenous. There is only one output (vehicle-km), three inputs (routes length, number of vehicles in the fleet, and company costs), and 15 variables exogenous to the operation. The efficiency values are determined with the first four, as shown in (5). The 15 remaining variables will allow for studying the influence of exogenous variables on the efficiency of the companies.

# 6. Calculation and Analysis of the Technical Efficiency and Slacks

The efficiency of the urban bus companies has been calculated from 2004 to 2009 (see Table 5), according to the model (5). When studying the results, it should be noted the different sizes of the companies and the metropolitan areas. In order to avoid problems of scale in the calculations, it is used a variable returns to scale (VRS) model.

Some conclusions can be gleaned from the analysis of the results. Madrid, Oviedo, Zaragoza, Pamplona, and La Coruña are the cities whose companies provide the best results, being efficient (a value of 1) every or almost every year of the study. Almeria and Lérida are efficient in the only year in which they were studied, 2009. On the contrary, Malaga, Granada, and Las Palmas de Gran Canaria have a negative tendency every year, with their bus companies achieving efficiency values in the 0.5 to 0.7 interval every year of the study. Large cities such as Barcelona and Valencia demonstrate modest results of between 0.65 and 0.75.

Due to the peculiarity of the data available, a single frontier would be chosen for all the years, as opposed to a different frontier for each. This fact must be kept in mind when analyzing the results. The consequences of this type of evaluation are as follows: if frontiers were established for each year, some DMUs from that year would have to be used to define it and would, therefore, have an efficiency value of 1.

By using a single frontier to join all the years, there is no requirement for efficient DMUs every year, such as in the case of the year 2008. This has two advantages: the first is that all the DMUs can be compared by using a single frontier, and doubts do not arise as to the comparability of two efficient DMUs in relation to two different efficiency frontiers. The second advantage is that one can discover if the technical efficiency in one year is less than that of another. Thus, according to the results, in the year 2008, the efficiency of all the DMUs is worse when compared to the other years. If a frontier had been established for 2008, the DMUs would have had better results. It must also be mentioned that, as the data series is discontinuous and has different DMUs for each year, the comparison of the frontiers would also present problems.

In [29], the efficiency values obtained are stable throughout the years. In that case, the same DMUs are always used and the variable values, both inputs and outputs, do not display important variations from one year to another. However, in [41], every year the same cities are used and there are important variations from one to another. Thus, the variability of the results of certain cities seems normal.

One interesting finding is that of the 81 DMUs studied, 19 (23.5% of the total) belong to the efficiency frontier. These 19 companies per year belong to nine different cities.

Once the efficiency values have been analyzed, the results of the slack values obtained for each of the inputs used to calculate the frontier are studied. For this end, it has been calculated the ratio of the slack divided by the value of the input. In [42] they do something similar, though in that case the value of the slacks is monetized. The results are shown in Table 6.

The relative slacks of the routes length variable present a range that goes from 0% efficient companies to 75% in the case of Murcia in 2008, though the majority of the inefficient companies show slacks of around 40–60%. The slack values are repeated, in general, throughout the years. In some cases, interesting situations can be observed. For example, Madrid is efficient every year except for 2007 and 2008, in which slacks of approximately 15% can be seen. Another noteworthy case is that of Vigo, which has no slacks in 2006 and 75% in 2008. Oviedo reaches 20% slacks in 2007 while every other year it is efficient. Mallorca and Pamplona are similar cases.

Considering the situations described above, and knowing that the efficiencies seen here are not absolute values, but instead relative efficiencies between the different companies,

Variable name	Type	Definition	Minimum value	Maximum value	Units
Vehicles-km per year (VK)	Endogenous. Output	This figure is calculated taking into account existing vehicles in the fleet and the distance they have traveled in operation throughout the year.	1.9	100.4	Million vehicle-km/year
Routes length (RL)	Endogenous. Input	and end) of each line. If several lines share the same stretch of the route, the route is counted as many times as existing lines. The circular lines are counted twice. Only daytime network is accounted	147	3,870	km
Number of vehicles (NV)	Endogenous. Input	Number of buses in the fleet	42	2,092	$n^{\circ}$ of vehicles
Company costs (CC)	Endogenous. Input	It Includes: personnel costs (including social security and pensions), energy costs, purchase of goods and external services (including outsourcing), maintenance costs (except those included in investment expenses), financial expenses, depreciation costs, taxes, and fees. In all cases, it includes only the actual expenditure incurred, not spending approvals made	5.69	413.75	Million euro/year
Main city population (MCP)	Exogenous. Socio economic	Population of the main city. Calculated at the first of January of the analyzed year	131,731	3,255,944	Inhabitants
Metropolitan area population (MAP)	Exogenous. Socio economic	Population of the metropolitan area. Calculated at the first of January of the analyzed year	246,056	6,386,932	Inhabitants
Main city surface (MCS)	Exogenous. Socio economic	The main city's surface is the municipality surface	25	974	$\mathrm{km}^2$
Metropolitan area surface (MAS)	Exogenous. Socio economic	The metropolitan area's surface is the zone where public transport authority offers services	37	11,313	$\mathrm{km}^2$
Regional GDP per capita (RGDP)	Exogenous. Socio economic	Gross domestic product in the metropolitan area's region. It is divided by the total number of inhabitants of the region	15,021	31,496	Euro/inhabitant
Labor force rate of the province (LFRP)	Exogenous. Socio economic	The labor force includes people aged 16 and over who supply labor for the production of economic goods and services or are in a position to do	49	65	%
Unemployment rate of the province (URP)	Exogenous. Socio economic	Rate of labor force that is unemployed	5	28	%
Car ownership at the province (COP)	Exogenous. Socio economic	Total number of cars and motorbikes at the province. Taxis and vans are not included. It is divided by the total number of inhabitants of the province	413	209	Vehicles/1,000 inhabitants
Main city pop. density (MCPD)	Exogenous. Urban form	Main city's total population divided by main city's total surface	477	8,091	Inhabitant/km <sup>2</sup>
Metropolitan area pop. density (MAPD)	Exogenous. Urban form	Metropolitan area's total population divided by metropolitan area's total surface	65	6,650	Inhabitant/km <sup>2</sup>
Population ratio (PR)	Exogenous. Urban form	Main city's population divided by metropolitan area's population	0	1	Adimensional. Values between 0 and 1

		TABLE 4: Continued.			
Variable name	Type	Definition	Minimum Maximum value value	Maximum value	Units
Surface ratio (SR)	Exogenous. Urban form	Main city's surface divided by metropolitan area's surface	0	1	Adimensional. Values between 0 and 1
Management (PRIV)	Exogenous. Management and alternative modes	Dummy variable. It takes value 0 if public management, value 1 if private management.	0	1	Adimensional. Values between 0 and 1
Main city alternative rail modes (MCARMs)	Exogenous. Management and alternative modes	It takes value 0 if there are no alternative rail modes at the main city, value 1 if there are	0	1	Adimensional. Values between 0 and 1
Metropolitan area alternative rail modes (MAARMs)	Exogenous. Management and alternative modes	It takes value 0 if there are no alternative rail modes at the metropolitan area, value 1 if there are	و 0	1	Adimensional. Values between 0 and 1
Data refers to period 2004–2009.	.60				

_
· U
<ul> <li>1</li> </ul>
_
. =
_
-
~~
()
$\sim$
নি
(T)
_

ISRN Civil Engineering

Matropolitan area population	Main city	Year							
Metropolitan area population	Main city	2004	2005	2006	2007	2008	2009		
	Madrid	1.000	1.000	1.000	0.901	0.847	1.000		
	Barcelona	0.741	0.694	0.677	0.682	0.626	0.651		
>1 million inhab.	Valencia	0.750	0.695	0.714	0.672	0.617	0.654		
	Murcia	_	_	—	0.602	0.644	0.670		
	Sevilla	0.741	0.741	0.669	000         0.901         0.847           777         0.682         0.626           714         0.672         0.617           -         0.602         0.644           669         0.696         0.686           000         0.919         0.895           635         0.516         0.514           -         1.000         0.841           613         0.615         0.573           950         0.970            -         -         0.742           665         0.652         0.618           -         -            759         0.718         0.729           -         -            767         0.758         0.756           000         0.892         0.743	0.691			
	Oviedo	1.000	1.000	1.000	0.919	0.682       0.626         0.672       0.617         0.602       0.644         0.696       0.686         0.919       0.895         0.516       0.514         1.000       0.841         0.615       0.573         0.970           0.742         0.652       0.618             0.718       0.729             0.758       0.756         0.892       0.743	1.000		
	Málaga	0.546	0.532	0.535	0.516	0.514	0.515		
	Mallorca	_	_	—	1.000	0.841	0.846		
Between 0.5 and 1 million inhab.	Gran Canaria			0.613	0.615	0.573	0.591		
between 0.5 and 1 minion minab.	Zaragoza	1.000	1.000	0.950	0.970	_	1.000		
	San Sebastián	_	_	0.669         0.696         0.686           1.000         0.919         0.895           0.535         0.516         0.514            1.000         0.841           0.613         0.615         0.573           0.950         0.970              0.742           0.665         0.652         0.618                0.759         0.718         0.729                0.767         0.758         0.756           1.000         0.892         0.743	0.746				
	Granada	0.693	0.680	0.665	0.652	0.618	0.589		
	Almería						1.000		
	Alicante	0.808	0.794	0.759	0.718	0.729	0.718		
<0.5 million inhab.	Lérida						1.000		
<0.5 minon miao.	Pamplona	1.000	0.964	0.767	0.758	0.756	0.766		
	Vigo	_	_	1.000	0.892	0.743	_		
	La Coruña	1.000	0.999	1.000	—	0.966	1.000		

one must consider the following: for some cities with these intervals, the maximum value probably better reflects the inefficiency, even while in other years it is efficient. On the other hand, in the extreme case of Vigo, the result is possibly motivated by the problem of not always using the same DMUs each year, meaning that the efficiency references change.

Finally, the values must not be taken in a literal way by, for example, assuming that 75% of Murcia's network is superfluous. In [42] the slack values (that have been monetized) also reach 75% in some cases, which means that this does not seem to be an anomaly. The relative values of the slacks instead highlight tendencies. Thus, Murcia clearly has a problem with the design of its lines, while in other cases a 30–40% slack would indicate the need to study or reconfigure the lines. The same can be said for cases of efficient companies which present slacks one year: the value does not indicate that they are not perfectly efficiency, but that they could have some problems with some lines.

With regard to the relative slacks of the number of vehicles variable, it can be observed that they do not reach values as extreme as those of route length, with maximum values of around 40%. Indeed, the most extreme interval is found in the case of Pamplona, which varies from 0% to 15%. It seems, then, that the slacks values of this input exhibit the greatest stability.

Finally, with regard to the relative slacks of the company costs variable, the values are not as stable as in the case of the number of vehicles. In many cases, they have a variation of 25–30% (Madrid, Barcelona, Oviedo, etc.). The maximum values are those of Barcelona in 2008 and 2009, exceeding 50%. Other cities such as Valencia, Sevilla, or Málaga also exhibit high slack values of around 40%. For this variable,

the high slack values seem more reasonable. The companies cover their operating deficits with public subsidies, meaning that they have less incentive to drastically reduce their costs and operate in a more efficient manner.

### 7. The Influence of Exogenous Variables on Slacks

A process based on the *two-stage* method mentioned in [19] is followed for the study of the effect of the exogenous variables on the operational efficiency. Thus, the outputs and inputs chosen to define the frontier are introduced into the model (5), which produces the efficiency value and the slack value for each of the inputs. Below, the relationship of the slack value to a series of chosen exogenous variables will be determined by means of a multiple linear regression. The objective of the regression is to determine if a correlation exists between the slack value and these exogenous variables.

The analysis consists of an iterative process. In the first iteration, a linear regression is performed with all the variables of the considered group. Once the regression results are obtained, those variables whose probability of error is greater than 5% are excluded, those for which there exists a 95% probability of an error when included in the regression. In this way, after various iterations, a result can be reached in which all the variables have a probability of error that is less than 5%.

The iterative evaluation process of the slacks for each one of the input variables is repeated twice, once for each of the groups of exogenous variables. In Alternative 1, socioeconomic variables are used, such as management and alternative mode variables. In Alternative 2, population

TABLE 6: Ratio slacks/input value (%), per year and input.

			Alternati	ve 1					
Input slacks			Exog	enous var	iables sign	s in the regressio	n		$R^2$ coefficient
input slacks	MCP	MAP	MCS	MAS	LFRP	RGDP	MACARM	MAARM	K coefficient
Routes length	_	_	+	+	0,30				
Number of vehicles		_		+		+			0,46
Company costs	_	+	+	_	0,56				
				Alt	ternative 2	2			
Input slacks		E	xogenous	variables si	igns in the	regression		$R^2$ co	pefficient
	MCPD	MAPD	PR	SR	PRIV	MACARM	MAARM		
Routes length	_	+	+	_	+	0,	15		
Number of vehicles	_	+				+	_		0,70
Company costs	_	+	_	_		0,57			

TABLE 7: Summary of regression results of the inputs slacks versus the exogenous variables.

See Table 4 for the definition of variables.

density, management, and alternative mode variables are used. The socioeconomic variables of population and surface area are not used at the same time as the density variables, as the latter are derived from the first, making their use redundant and, therefore, incorrect.

After completing the calculations, some of the variables do not appear in the regressions as they are not significant.

Table 7 shows the results of the regressions. For Alternative 1, it can be observed that there is no clear tendency in the behavior of the exogenous variables with relation to the slacks. Only in the case of the metropolitan area population (MAP) variable and regional GDP (RGDP) are the sign repeated in two inputs. However, in the case of the company costs input, the results of the regression with regard to population are contradictory to those of the other two inputs. Most of the variables are significant in some regressions and not in others. Finally, it bears mentioning that the regression coefficients  $R^2$  are not very high and that only for operating costs is the  $R^2$  value greater than 0.5. The Alternative 1 results, then, are hardly conclusive.

With regard to the results of Alternative 2, it can be observed that certain tendencies exist. Thus, for the exogenous variables of the main city population density (MCPD) and metropolitan area population density (MAPD) variables, the three inputs all produce the same sign. In the case of the surface ratio (SR), there is also the same tendency for two of the three inputs. Finally, the Metropolitan area alternative rail modes (MAARMs) variable also presents a similar tendency in two of the three inputs. On the other hand, the regression coefficients  $R^2$  of two of the regressions are greater than 0.5, even reaching 0.7 in the case of the number of vehicles variable. However, the routes length variable has a very low  $R^2$  (0.15), while the signs of the exogenous variables have the same tendency for the other two inputs. Due to the fact that the regression coefficients are not higher, one can say that the exogenous variables used here are not sufficient to explain completely the behavior of the slacks, but it is unknown what other variables would do so.

As for the signs of the exogenous variables used in Alternative 2, it can be seen that for the main city population density (MCPD) variable, the greater the density, the higher the slacks. This seems illogical, as a greater population density allows for concentration of supply and, thus, for greater efficiency. In contrast, the greater the population density in the metropolitan areas, the lower the values of the slacks. This could be explained by the fact that denser and more populated metropolitan areas tend to have close relationships with the main city, facilitating the use of the urban bus services by inhabitants of other municipalities. These services, with the increase in users, can achieve economies of scale. When it comes to the surface area ratio, the higher the ratio, the higher the slacks. The ratio provides the relationship between the surface area of the main city and that of the rest of the metropolitan area, and thus, a higher ratio value means that the main city represents an important part of the total surface area of the metropolitan area, for which the "direction" of the sign will be the opposite. Finally, for the metropolitan area alternative rail modes (MAARMs) variable, which reflects the rail modes of the metropolitan area but not of the capital city, their very existence increases the value of the slacks, which is logical as these modes act as competitors to the bus services. The slacks have a negative sign, meaning that if the variable has a negative coefficient, the variable will increase the value of the slacks. The variables with positive coefficients reduce the negative value of the slacks.

Some findings can be drawn from the comparison of the results of exogenous variables with other authors.

(i) Population variables: the main city population density (MCPD) variable has a negative sign, meaning that it increases the slack and, therefore, the inefficiency, whereas the metropolitan area population density (MAPD) variable acts in the opposite way. Of all the studies considered here, only in two do population variables appear as inputs or outputs [39, 43], and in a third study, the variable is used as exogenous, but without a significant result [32]. Finally, in two other studies, the variable acted in favor of efficiency in one [44] and against efficiency in the other [36]. Thus, clear conclusions cannot be drawn with respect to this type of variable.

- (ii) Surface area variables: none of the studies directly used the main city (MCS) or metropolitan area surface (MAP) variables. In other studies, these appear as population density variables.
- (iii) Variables similar to wealth variables (RGDP) and alternative modes (COP, MCARM, and MAARM) are only employed by [36] and his study differs from this one in that there they appear as insignificant.
- (iv) In [32], a dummy owner variable is included which is only valid if the business is private. This variable acts in favor of efficiency. In the regression of the route length variable slacks in Alternative 2, the PRIV variable represents this idea and acts in the same way.

Thus, after comparing various studies, it can be stated that it is not possible to come to conclusions with respect to the exogenous variables in order to contextualize the efficiency of the sector. The results are not homogenous.

### 8. Conclusions

The purpose of this paper is to analyze the technical efficiency of the main urban bus companies in Spain. In order to do this, an SBM model has been used because it enables a more adjusted calculation of efficiency, due to the presence of the slacks. In the end, the final model obtained for calculating the efficiency included one variable indicating transit supply, one related to infrastructure, and two of the classic variables from the Cobb-Douglas models (fleet size and operating costs).

The efficiency values fluctuate between 0.514 and 1.000. 19 DMUs reach the maximum value (1.000), representing 23.5% of all the DMUs analyzed. Half of the cities present the maximum efficiency in, at least, one year. In these cities the bus companies are of various sizes. The companies of Madrid, Oviedo, Zaragoza, and La Coruña are efficient in the majority of the years, while the companies of Mallorca and Pamplona are efficient only in one year.

From the analysis of the slacks, the results cannot be taken literally; instead, they serve to highlight tendencies with regard to the inefficiency value compared with a determined variable. That is the case of Murcia, which has 75% of routes length slacks. It does not mean that this % of the lines is remaining, but it is clear that Murcia may have a problem of design of the network. The same can be said for efficient companies: they do not have a perfect service, but their surpluses are not very high.

The use of the slacks to perform the regression of the exogenous variables enables one to easily and comfortably study the inefficiency produced by the operating conditions. Analyzing slack by slack, one can study in a disaggregated way the inefficiency of each input value used. In this way one can act in an individualized way with regard to each production factor in order to improve efficiency.

As for the exogenous variables analyzed here, although some of the socioeconomic variables appear in the regressions, their presence is not repeated through all the regressions performed and they do not, therefore, display any tendency. Additionally, the regression coefficients  $R^2$  of these regressions are not high. The urban population variables are indeed repeated in the different regressions and, as with their signs, display tendencies. The regression coefficients  $R^2$  in these cases show higher values and thus are more reliable.

With this methodology, it is possible to study the efficiency of each variable used as an input: the study of the slacks allows for an understanding of the over sizing or under sizing that exists for each one of the inputs. Then, it is possible to make recommendations to the companies in such a way that they improve their efficiency. The analysis of the exogenous variables also allows to understand part of the inefficiency, but in a way that companies cannot do so much, because they have no power to influence these variables.

#### References

- A. Aparicio, *Informe 2002 del Observatorio de la Movilidad Metropolitana*, Ministerio de Medio Ambiente y Ministerio de Fomento, Madrid, Spain, 2004.
- [2] A. Monzón, A. M. Pardeiro, and P. Pérez, *Informe 2003 del Observatorio de la Movilidad Metropolitana*, Ministerio de Medio Ambiente, Madrid, Spain, 2005.
- [3] A. Monzón, A. M. Pardeiro, and P. Pérez, *Informe 2004 del Observatorio de la Movilidad Metropolitana*, Ministerio de Medio Ambiente y Ministerio de Fomento, Madrid, Spain, 2006.
- [4] A. Monzón, R. Cascajo, A. M. Pardeiro, P. Jordá, P. Pérez, and M. A. Delgado, *Informe 2005 del Observatorio de la Movilidad Metropolitana*, Ministerio de Medio Ambiente y Ministerio de Fomento, Madrid, Spain, 2007.
- [5] A. Monzón, R. Cascajo, P. Jordá, P. Pérez, and I. Rojo, *Informe 2006 del Observatorio de la Movilidad Metropolitana*, Ministerio de Medio Ambiente y Medio Rural y Marino y Ministerio de Fomento, Madrid, Spain, 2008.
- [6] A. Monzón, R. Cascajo, P. Jordá, P. Pérez, M. Fernández, and J. del Álamo, *Informe 2007 del Observatorio de la Movilidad Metropolitana*, Ministerio de Medio Ambiente y Medio Rural y Marino y Ministerio de Fomento, Madrid, Spain, 2009.
- [7] A. Monzón, R. Cascajo, and P. Jordá, *Informe 2008 del Observatorio de la Movilidad Metropolitana*, Ministerio de Medio Ambiente y Medio Rural y Marino y Ministerio de Fomento, Madrid, Spain, 2010.
- [8] A. Monzón, R. Cascajo, P. Jordá, B. Muñoz, and L. Delgado, *Informe 2009 del Observatorio de la Movilidad Metropolitana*, Ministerio de Medio Ambiente y Medio Rural y Marino y Ministerio de Fomento, Madrid, Spain, 2011.
- [9] P. Jordá, Metodología de evaluación de la eficiencia de los servicios de autobús urbano. Aplicación a las grandes ciudades españolas en el periodo 2004–2009, Tesis doctoral, Departamento de Ingeniería Civil: Transporte, Universidad Politécnica de Madrid, Madrid, Spain, 2012.
- [10] B. De Borger, K. Kerstens, and A. Costas, "Public transit performance: what does one learn from frontier studies?" *Transport Reviews*, vol. 22, no. 1, pp. 1–38, 2002.
- [11] J. Odeck and A. Alkadi, "Evaluating efficiency in the Norwegian bus industry using data envelopment analysis," *Transportation*, vol. 28, no. 3, pp. 211–232, 2001.
- [12] N. K. Avkiran and T. Rowlands, "How to better identify the true managerial performance: state of the art using DEA," *Omega*, vol. 36, no. 2, pp. 317–324, 2008.
- [13] C. Pestana and N. Peypoch, "Productivity changes in Portuguese bus companies," *Transport Policy*, vol. 17, no. 5, pp. 295–302, 2010.

- [14] A. Charnes, W. W. Cooper, and E. Rhodes, "Measuring the efficiency of decision making units," *European Journal of Operational Research*, vol. 2, no. 6, pp. 429–444, 1978.
- [15] R. D. Banker, A. Charnes, and W. W. Cooper, "Some models for estimating technical and scale inefficiencies in data envelopment analysis," *Management Science*, vol. 30, no. 9, pp. 1078–1092, 1984.
- [16] W. W. Cooper, L. M. Seiford, K. Tone, and J. Zhu, "Some models and measures for evaluating performances with DEA: past accomplishments and future prospects," *Journal of Productivity Analysis*, vol. 28, no. 3, pp. 151–163, 2007.
- [17] L. R. Murillo-Zamorano, "Economic efficiency and frontier techniques," *Journal of Economic Surveys*, vol. 18, no. 1, pp. 33– 45, 2004.
- [18] K. Tone, "Slacks-based measure of efficiency in data envelopment analysis," *European Journal of Operational Research*, vol. 130, no. 3, pp. 498–509, 2001.
- [19] H. O. Fried, S. S. Schmidt, and S. Yaisawarng, "Incorporating the operating environment into a nonparametric measure of technical efficiency," *Journal of Productivity Analysis*, vol. 12, no. 3, pp. 249–267, 1999.
- [20] P. Bogetoft and J. L. Hougaard, "Efficiency Evaluations Based on Potential (Non-Proportional) Improvements," *Journal of Productivity Analysis*, vol. 12, no. 3, pp. 233–247, 1999.
- [21] K. Tone, "A slacks-based measure of super-efficiency in data envelopment analysis," *European Journal of Operational Research*, vol. 143, no. 1, pp. 32–41, 2002.
- [22] H. Morita, K. Hirokawa, and J. Zhu, "A slack-based measure of efficiency in context-dependent data envelopment analysis," *Omega*, vol. 33, no. 4, pp. 357–362, 2005.
- [23] J. Liu and K. Tone, "A multistage method to measure efficiency and its application to Japanese banking industry," *Socio-Economic Planning Sciences*, vol. 42, no. 2, pp. 75–91, 2008.
- [24] N. K. Avkiran, "Removing the impact of environment with units-invariant efficient frontier analysis: an illustrative case study with intertemporal panel data," *Omega*, vol. 37, no. 3, pp. 535–544, 2009.
- [25] J. Zhu, "Super-efficiency and DEA sensitivity analysis," European Journal of Operational Research, vol. 129, no. 2, pp. 443– 455, 2001.
- [26] M. L. Delgado, M. A. Sánchez, and L. Mora, *Reflexiones sobre la organización del transporte urbano colectivo en las grandes ciudades*, Cátedra de Ecotransporte, Tecnología y Movilidad, Universidad Rey Juan Carlos, Madrid, Spain, 2009.
- [27] CEOE, Memorándum: El sector del transporte en España, Confederación Española de Organizaciones Empresariales, Madrid, Spain, 2009.
- [28] G. Cheng and Z. Qian, "MaxDEA 5.0.," 2011, http://www .maxdea.cn.
- [29] J. F. Nolan, "Determinants of productive efficiency in urban transit," *Logistics and Transportation Review*, vol. 32, no. 3, pp. 319–342, 1996.
- [30] P. A. Viton, "Technical efficiency in multi-mode bus transit: a production fronteir analysis," *Transportation Research B*, vol. 31, no. 1, pp. 23–39, 1997.
- [31] J. Cowie, "The technical efficiency of public and private ownership in the rail industry: the case of Swiss private railways," *Journal of Transport Economics and Policy*, vol. 33, no. 3, pp. 241–252, 1999.
- [32] K. Kerstens, "Technical efficiency measurement and explanation of french urban transit companies," *Transportation Research A*, vol. 30, no. 6, pp. 431–452, 1996.

- [33] M. G. Karlaftis, "A DEA approach for evaluating the efficiency and effectiveness of urban transit systems," *European Journal* of Operational Research, vol. 152, no. 2, pp. 354–364, 2004.
- [34] P. A. Viton, "Changes in multi-mode bus transit efficiency, 1988–1992," *Transportation*, vol. 25, no. 1, pp. 1–21, 1998.
- [35] K. Kerstens, "Decomposing technical efficiency and effectiveness of French urban transport," Annales d'économie et de statistique, vol. 54, pp. 129–155, 1999.
- [36] I. M. García-Sánchez, "Technical and scale efficiency in Spanish urban transport: estimating with data envelopment analysis," *Advances in Operations Research*, vol. 2009, Article ID 721279, 15 pages, 2009.
- [37] C. W. Cobb and P. H. Douglas, "A theory of production," *The American Economic Review*, vol. 18, no. 1, pp. 139–165, 1928.
- [38] P. Cantos, J. M. Pastor, and L. Serrano, "Productivity, efficiency and technical change in the European railways: a nonparametric approach," *Transportation*, vol. 26, no. 4, pp. 337– 357, 1999.
- [39] A. Novaes, "Rapid transit efficiency analysis with the assurance-region DEA method," *Pesquisa Operacional*, vol. 21, no. 2, pp. 179–197, 2001.
- [40] M. Asmild, T. Holvad, J. L. Hougaard, and D. Kronborg, "Railway reforms: do they influence operating efficiency?" *Transportation*, vol. 36, no. 5, pp. 617–638, 2009.
- [41] J. F. Nolan, P. C. Ritchie, and J. E. Rowcroft, "Identifying and measuring public policy goals: ISTEA and the US bus transit industry," *Journal of Economic Behavior and Organization*, vol. 48, no. 2, pp. 291–304, 2002.
- [42] T. Holvad, J. L. Hougaard, D. Kronborg, and H. K. Kvist, "Measuring inefficiency in the Norwegian bus industry using multi-directional efficiency analysis," *Transportation*, vol. 31, no. 3, pp. 349–369, 2004.
- [43] V. Pina and L. Torres, "Analysis of the efficiency of local government services delivery. An application to urban public transport," *Transportation Research A*, vol. 35, no. 10, pp. 929– 944, 2001.
- [44] D. T. Barnum, S. Tandon, and S. McNeil, "Comparing the performance of bus routes after adjusting for the environment using data envelopment analysis," *Journal of Transportation Engineering*, vol. 134, no. 2, pp. 77–85, 2008.