SESO

STUDIECENTRUM VOOR ECONOMISCH EN SOCIAAL ONDERZOEK

VAKGROEP PUBLIEKE ECONOMIE

Technical Efficiency Measurement and Explanation of French Urban Transit Companies

Kristiaan KERSTENS*

report 95/333

December 1995

* A preliminary version was presented at the "Ninth National Conference on Quantitative Methods for Decision Making (ORBEL 9)" (Brussels, ERM-KMS). It also benefitted from seminars at the Warwick Business School and the Université Catholique de Lille and from presentation at the "Fourth European Workshop on Efficiency and Productivity Measurement" (Louvain-la-Neuve, CORE). Detailed comments of B. De Borger, I. Janssens and S. Wouters, and two referees are gratefully acknowledged. We furthermore thank Peter Rousseeuw for providing us with the MINVOL software for outlier detection. The entire responsability remains with the author.

University of Antwerp, UFSIA-SESO, Prinstraat 13, B-2000 Antwerp, Belgium
Tel: + 32-3-220.40.60, Fax: + 32-3-220.40.26, E-mail: dse.kerstens.k@alpha.ufsia.ac.be.

Universitaire Faculteiten St.-Ignatius
Prinsstraat 13 - B 2000 Antwerpen

D/1995/1169/20

Abstract:

The performance of a sample of French urban transit companies is evaluated using a broad selection of nonparametric reference technologies for two specifications of the production process. In particular, the variable returns to scale Data Envelopment Analysis (DEA) models with either strong or weak disposability in both inputs and outputs, and the Free Disposal Hull (FDH) are applied.

An extensive comparison of the resulting radial output efficiency measures yields the following major methodological conclusions. First, the location of the efficiency distributions differs substantially depending on the methodology and especially on the output specification considered. The latter differences vanish if the impact of outliers is eliminated. Second, convexity has a stronger influence on the efficient-inefficient dichotomy than allowing for congestion by means of a weakly disposable DEA model.

For policy purposes, these efficiency distributions are explained using a Tobit model. The findings corroborate results reported elsewhere: the relevance of ownership, the use of risk-sharing incentives in contracting, the harmful impact of subsidies, etc. Furthermore, the the network structure seems to account for some differences in performance. Finally, a novelty in the urban transit context is the indirect monitoring effect of the French earmarked transportation tax.

Keywords: Technical efficiency; Urban transport; Censored regression.

JEL classification: D24, R41.

0. INTRODUCTION

Urban transit companies are a major part of the transportation network in an economy. They provide passenger services within the cities and their agglomerations mainly by buses, trams, and metros. In most countries the urban transit services are provided by public, private or mixed companies in a highly regulated environment. Moreover, important components of the transportation infrastructure are public goods. There are sound economic reasons for this large amount of state intervention based mainly on the recognition of a variety of market failures. First, there is lack of competition and one can even question whether transit markets are natural monopolies, which may be partly due to sunk costs and the decline in the modal share. Furthermore, there exist asymmetries in information, especially in the safety area. In addition, transportation is a major source of externalities imposed on the environment and it is deeply intertwined with locational choices. Finally, distributional issues have traditionally lead to consider transport as a merit good. Recently, however, concerns about regulatory failures have led to a reassessment of the traditional tools in transportation policy. Sources of regulatory failures include: at best guidelines focus on allocative efficiency but ignore technical efficiency; phenomenal increases in financial support and subsidies are needed to maintain existing transit operations; etc.[1]

An important aspect of transport policy is the supply of urban transport. In view of these recent policy debates, it is of interest to determine whether urban transit operators are working in a technically efficient way. Detecting technical inefficiencies (or X-inefficiencies) can serve a variety of purposes. For instance, technical efficiency measurement can contribute to the discussions on the relative merits of private versus public provision of these transportation services, on the role of regulation, on controlling subsidy levels, etc. (see Glaister, Starkie and Thompson (1990) and Button and Weyman-Jones (1994)). No doubt, these issues have been prominent arguments in the political and academic debates legitimating privatization and deregulation policies in many countries during the eighties. Transport policy reforms in the UK offer a well-known example.

It is the main purpose of this paper to investigate the performance of the French urban transit sector. The focus is on single mode companies operating buses. This has an immediate

---

[1] Banister, Berechman and De Rus (1992), Berechman (1993), and Glaister, Starkie and Thompson (1990) survey these issues in detail.

policy interest if one realizes that this sector has a rather low cost-covering ratio relative to other Western European countries (see Gwilliam and van de Velde (1990) or Pucher (1988)). Traditionally, research on the supply of transportation services has focused on representations of technology describing so-called average practice behaviour. Cost studies in this tradition and in particular the specifics of transport cost functions have been reviewed by Berechman and Giuliano (1985), Berechman (1993), and Jara Diaz (1982), among others. More recently, the development of efficiency measurement methodologies has led to new applications in the field of transportation (surveyed in Berechman (1993: 159-160)). This study concentrates on the evaluation of technical efficiency using deterministic nonparametric reference technologies. It employs different specifications of these frontier technologies and it uses two descriptions of the urban transit technology to control for the possible sensitivity of the resulting efficiency measures. In addition, the determinants of the performance of French urban transit companies are investigated using censored regression models.

The first of this paper's sections gives an accurate presentation of the various methodologies used to gauge the performance of the French companies. Section 2 describes the sample of French urban transit operators used in the empirical application. The results of technical efficiency measurement are presented in section 3. It also devotes attention to methodological issues regarding the precise effect of different technology assumptions on performance results and on the possible impact of outliers. In section 4 the focus is on policy matters. In particular, it attempts to determine the sources and causes of observed technically inefficient behaviour. Section 5 concludes.

1. DETERMINISTIC NONPARAMETRIC FRONTIERS

Applied production analysis has been characterized for years by the analysis of so-called average practice technologies, whereby researchers fit functions through the middle of the data. The inherent frontier nature of central theoretical concepts in economics, like the notion of a production function, was not recognized until the work of Koopmans (1951) and Debreu (1951) who provided respectively a definition and a measure of technical efficiency. Farrell (1957) added an allocative, or price, efficiency component and offered the first empirical frontier study. The acknowledgment of the possibility of technically inefficient behaviour and of the conceivable divergence between characteristics of average and best practice production technologies redirected attention to the development of frontier estimation techniques. Farrell

(1957) insisted on the relative nature of the frontier concept, i.e., relative to observed best practice in a reference set.

There are various approaches to reconstruct production frontiers, which can be usefully distinguished along the following lines: parametric versus nonparametric; and deterministic versus stochastic methods.[2] Parametric methods require the specification of a functional form for the production technology, and in some cases for the distribution of technical efficiency. Nonparametric methods impose no a priori functional form and number of parameters on the observations. Stochastic methodologies make explicit assumptions with respect to the stochastic nature of the data, while the deterministic do not. The analysis in this paper is restricted to a variety of deterministic nonparametric frontiers originating from the seminal contribution of Farrell (1957).

Deterministic nonparametric methods are based on piecewise linear frontiers estimated using mathematical programming techniques. These extremal methods envelop the data as tightly as possible subject to certain maintained assumptions --typically weaker than in the parametric approaches-- on the structure of the production technology. It is worth stressing that these technologies provide the tightest inner approximations of production technology consistent with a hypothesis (Färe, Grosskopf and Lovell (1994)). Consequently, technical efficiency measured relative to these inner bound reference technologies should be interpreted as an upper bound.

We first consider three types of reference technologies (see Färe, Grosskopf and Lovell (1985, 1994) and Tulkens (1993) for details), and then discuss how technical efficiency can be measured relative to these frontier specifications.

First, the output correspondence of the popular variable returns to scale (vrs) Data Envelopment Analysis (DEA) model with strong disposability both in inputs and outputs (sd) is defined as follows:

$$P(x)^{sd-vrs} = \{ y \mid Y^t z \geqq y,\ X^t z \leqq x,\ I_k^t z = 1 \},$$

where Y is the k x n matrix of observed outputs, X is the k x m matrix of observed inputs, z is a k x 1 vector of intensity or activity variables, y and x are n x 1 and m x 1 vectors of outputs respectively inputs, and I_k is a k x 1 unity vector.

[2] Detailed surveys are Lovell (1993) and Oum, Tretheway and Waters (1992), where the latter reference focuses on transportation.

This convex hull is represented by Figure 1. The inequality signs in the constraints of the linear programming formulation allow for strong disposability. For instance, an observed output vector can be smaller than the linear combination of observations D and E. All observations to the southwest of the line segment DE are therefore feasible. This explains the line originating in observation E and extending parallel to the second axis. Using a similar reasoning for all line segments yields the production possibility set bounded by the line segment BCDE and the lines parallel to both axes. Strong input disposability can be analogously illustrated.

< FIGURE 1 ABOUT HERE >

Second, another DEA reference technology --which has not been as widely applied-- relaxes the strong disposability assumptions of the previous model. Imposing weak instead of strong disposal of both input and output dimensions, its output correspondence is now defined in the following way:

$$P(x)^{wd-vrs} = \{ y \mid \gamma Y^t z = y,\ X^t z = \delta x,\ I_k^t z = 1,\ \gamma,\ \delta \in (0,1],\ z \in \mathbb{R}_+^k \},$$

where notation is analogous to the earlier formulation except for the scalars γ and δ. Figure 2 adds to the understanding of this weakly disposable technology which admits for congestion in production. The equality sign allows for linear combinations of activities in the non-economic region of the output correspondence, e.g., the line segment DE. For instance, starting from point D a reduction of output y_2 involves opportunity costs, either by using additional inputs and a given output y_1 (by moving away from D --direction origin-- parallel to the y_2 axis), or by requiring a reduction in output y_1 for a constant input vector (by moving from D towards E). Activities on the ray beyond E are technically inefficient. The basic idea of congestion is that wasting inputs or outputs involves opportunity costs.

< FIGURE 2 ABOUT HERE >

Third, the Free Disposal Hull (FDH) reference technology is obtained from the first model by restricting the intensity vector z to contain either zeros or ones. Since this excludes linear combinations of several observations, no convexity is imposed. Therefore, the FDH output correspondence can be defined as:

$$P(x)^{fdh} = \{y \mid Y^t z \geqq y,\ X^t z \leqq x,\ I_k^t z = 1,\ z_i \in \{0,1\} \}.$$

Each observed input-output combination spans one orthant, positive in the inputs and negative in the outputs, which simply reflects free disposal in inputs and outputs. The FDH reference technology is then the boundary to the union of all such orthants constituting the production possibility set. Its transformation curves typically have a staircase form, as illustrated in Figure 3.

< FIGURE 3 ABOUT HERE >

Existing studies on urban transit performance have employed a diversity of frontier methodologies. On the one hand, the studies of Chang and Kao (1992), Chu, Fielding and Lamar (1992), Obeng (1994), and Tone and Sawada (1990), for instance, are based on strongly disposable DEA models, while Gathon (1989) and Tulkens (1993) use the FDH. On the other hand, Fazioli, Filippini and Prioni (1993) and Viton (1986) apply parametric frontiers.

While this study neglects parametric approaches, it is good to point out that no urban transit sector study has used such a variety of nonparametric reference technologies on the same data set.[3] In particular, the use of the weakly disposable DEA model is novel in this context and needs some further explanation. Congestion is possible whenever the adjustment of (a subset of) inputs or outputs is constrained: by short run fixity, by lumpiness, by regulatory or other external constraints. Clearly, the outputs of companies providing transportation services in city centres may suffer from congestion, caused by private traffic and by the companies' own bus vehicles both creating congestion externalities. The former amount of traffic determines for a given capacity of the urban road network the operating environment for each bus operator. Given this urban operating environment, one example of inputs causing congestion in the outputs is the addition of busses on a given structure of the transit network. After a critical level the operator may well be affected by its own congestion externalities. The relative impact of congestion obviously depends, among others, on the exact nature of the transit network, e.g., whether bus lines share few or many routes. Therefore, for a given operating environment, the relative performance of any urban transit company can be affected by congestion among its inputs and outputs. The selected weakly

[3] An attempt to estimate a stochastic production frontier for this sample failed, because of a wrong third moment of the composed error. This failure to detect a frontier is well documented in the empirical literature. It is rather widespread in multi-industry or multi-country studies: see, e.g., Caves (1992: 8). Several plausible hypotheses for the breakdown of the composed error specification due to data problems are found in Caves and Barton (1990: 44-45). This possibility of a crucial misspecification test in the composed error model is, however, a protection against mistakes in inference.

disposable DEA model allows for all patterns of congestion.

Once these reference technologies have been specified, it is necessary to select an orientation of measurement to relate the observations to the boundary of the production possibility set. There are basically three orientations of measurement (see Färe, Grosskopf and Lovell (1985, 1994)). Technical efficiency can be measured in the inputs, in the outputs, or in both the inputs and the outputs. The choice of orientation depends on whether inputs, outputs, or both inputs and outputs are assumed to be freely adjustable. In the first case one is looking for a proportional reduction in the input usage which still allows to produce the same output vector. In the second case a proportional expansion of outputs which can be obtained from a given input vector is searched for. In the final case one seeks a maximum proportional change in all variables, i.e., a decrease in inputs and an increase in outputs.

From a policy perspective, the different parties involved in the organization of urban transit are probably interested in distinct orientations. On the one hand, the bus companies may favour an input-based orientation, as the outputs have been fixed in a contract with the public authority. On the other hand, the public authorities are eager to learn whether the output of public transport can be increased with given inputs. This would be helpful information when contracts with the companies are renewed. In line with Gathon (1989) and Tulkens (1993), this paper opts for a public policy point of view and, consequently, selects an output orientation. The radial output efficiency measures are illustrated on Figures 1 to 3.

The radial output efficiency measure is formally defined as:

$$DF_o(x,y) = \max\{ \mu \mid \mu \geq 1, \ \mu y \in P(x) \}.$$

It measures the maximum proportional increase in all outputs producible from given inputs. In the empirical application below this output-based measure is redefined so as to be situated between zero and one, with unity indicating efficiency. This is quite common in the empirical literature, as it allows for comparison with other measurement orientations. The Debreu-Farrell output measure becomes:

$$DF_o^*(x,y) = \min\{ \mu' \mid 0 < \mu' \leq 1, \ y/\mu' \in P(x) \}.$$

The computation of this radial efficiency measure involves solving one linear program for each observation in the case of the two DEA-type technologies and a mixed integer programming problem in the case of FDH.[4]

It is well-documented that the shape of the production possibility set determines the degree of technical efficiency detected. For instance, it is well-known that technical efficiency evaluated at a strongly disposable technology is no smaller than on a weakly disposable technology (i.e., $\mu^{sd\text{-}vrs} \geq \mu^{wd\text{-}vrs}$), and no smaller than on the FDH (i.e., $\mu^{sd\text{-}vrs} \geq \mu^{fdh}$). Furthermore, the technologies selected allow to answer the hitherto neglected methodological question whether the convexity or the weak disposability assumptions have a larger impact on technical efficiency measurement (i.e., the relation between $\mu^{wd\text{-}vrs}$ and μ^{fdh} is a priori unclear).

2. ORGANIZATION OF FRENCH URBAN TRANSIT AND DESCRIPTION OF THE SAMPLE

2.1 *Organization of French Urban Transit*

The institutional context in which the French companies operate can be concisely portrayed as follows.[5] Urban transport services are supplied by a single urban transport operator during a certain period and within a transport perimeter defined in an agreement with a public organizing authority. The latter can be any commune or association of communes --various legal forms of association coexist-- and owns most often the infrastructure, the equipment and the rolling stock. The transport perimeter need not coincide with territorial limits, but is only functional in dividing urban from interurban transport links. The single urban transport operator can be a private, public or a mixed company which offers a series of prespecified transportation services. In the latter mixed or semi-public companies (Sociétés d'Economie Mixte (SEM)) the majority of the capital stock is under public control. France is one among several Western European countries (UK, Scandinavian counties, ...) where the private sector is playing a substantial role in the urban transit industry.

The contractual agreements between operator and public organizing authority are governed

---

[4] For details on the mathematical programming formulations and the algorithms used, the reader is referred to Färe, Grosskopf and Lovell (1994) and Tulkens (1993); or to the appendix.

[5] The description of the institutional organization of French urban transport is based on Ministère de l'Equipement, du Logement, des Transport et de l'Espace (1990) and on Mitric (1988). Comparisons with transit sectors in other European and North American countries are made in Banister, Berechman and De Rus (1992), Berechman (1993) and in Pucher (1988), among others.

by a series of laws defining a set of minimal rules.[6] There are two main methods to manage the transit services. Either the organising authority sets up a partially or fully independent public administration (régie) to operate the transit services, or it delegates the operation to a company within the framework of an agreement. In the latter case, the agreements between operator (franchisee) and public organizing authority (franchisor) must minimally specify rules with respect to a prespecified range of topics. These laws allow in principle for four types of contracts differing mainly in their degree of risk-sharing. One can distinguish between a commercial risk, relating to the possible variations in revenues, and an operating cost or industrial risk, relating to the possible variations in costs. The risks of any divergencies between actual and targeted revenues and costs can or cannot be part of a sharing agreement between both contracting parties. To be precise, the parties have a choice between four possible contractual arrangements: (i) a full-risk contract; (ii) a contract with a guarantee of revenue; (iii) a fixed price contract; and (iv) a management contract. The first three contract types involve different degrees of risk-sharing, the last category is risk-free for the operator. All contracts are in principle limited to a 5 years period, but this period can be prolonged if the operator is financing a large amount of the capital investments.

As in most countries subsidies are an important element in French urban transit. The receipts from fares cover generally about 45% of total outlays, which is relatively low compared to other Western European countries. The main additional sources of financing are: operating subsidies provided by the local authorities; selective state subsidies for improving transportation productivity and service quality; and the income of police fines for parking and driving offences. Furthermore, there is a unique French locally levied transportation tax based on the salary bill of companies with more than 9 employees operating within the transport perimeter (le Versement Transport). It is earmarked for supporting public transport. If it is introduced by the public organizing authority, the tax rate can be set by these local authorities within the limits of a ceiling determined by the State. This transportation tax has been a main source of finance for public transport and represents nowadays almost 35% of the budget of the public organizing authorities.

Furthermore, the majority of the transport operators are part of a larger group or are members of a management association. For instance, the three largest groups, two of which

[6] Especially the 1982 law on internal transport (LOTI) provides the legal framework for these contractual agreements. See Ministère de l'Equipement, du Logement, des Transport et de l'Espace (1990: 35-41).

are private and one is semi-public, actually control about 65% of all transport systems in the urban areas of over 30.000 inhabitants. Regulatory policies include the specification of the network and level of service, fare regulation, and entry and exit.

In the last two decades France is one of the few countries which has experienced a growth in transit ridership through a strategy of expanding both the range and quality of its supply of transit services. This success required only a moderate increase in subsidy levels relative to costs. In short, in a European perspective the French urban transit industry combines a mix of private and public sector involvement --the latter especially at the local level-- in a highly regulated environment.

2.2 *Description of the Sample*

This subsection offers more details on the sample and the definition of inputs and outputs. To facilitate the analysis, the empirical application focuses on single mode urban transport companies. The sample contains 114 companies operating services outside the Paris region in 1990. About 60 other observations were discarded, either because they operated multi-mode transport systems (8), or because some data were missing or were suspected to be wrong after a careful scanning of the data (52).[7] It is assumed that these suspect observations were randomly distributed, such that the sample is still representative.

In the empirical analysis the following traditional outputs and inputs are used to model the production technology. There are two alternative outputs: the number of vehicle kilometres (y_1); and the number of seat kilometres (y_2). The inputs are: the average number of vehicles in use over the year (x_1); the average number of employees over the year (x_2); and the total fuel consumption over the year (x_3). The outputs are classical units times distance per unit time concepts: vehicle kilometres are the number of kilometres each vehicle has driven during the year; seat kilometres correct the above output definition for the differences in the number of seats and places on each vehicle. Observe that both outputs are pure supply indicators. The definition of the inputs closely follows the tradition in the transportation

[7] The scanning is mainly based on comparing the same variables in different years when data are available. The data have been published by the French ministry of transport (see Ministère de l'Equipement, du Logement, des Transport et de l'Espace (1991)). This ministry has been contacted in order to complete the missing data and to get more information on the suspect observations. But they were unable to provide us with any further details since the data are collected via a system of regional departments which --seemingly-- are hard to control. Furthermore, in practice it is unfeasible to force the individual operators to provide accurate information. Therefore, the analysis is based on the data which are available assuming that their quality meets some minimal standards.

literature. To experiment with the model specification both alternative output indicators are combined with the three traditional inputs. This yields two models: the three inputs and the number of vehicle kilometres (y_1); and the three inputs and the number of seat kilometres (y_2) (for convenience referred to as Model I respectively Model II).

Descriptive statistics are provided in Table 1. It is clear that transit operators vary considerably in size. It is exactly to avoid the confusion of scale and technical efficiencies that all reference technologies have been specified with a flexible returns to scale assumption. In the three production models discussed in the previous section, the variable returns to scale hypothesis is embodied in the constraint on the activity vector (i.e., $I_k^t z = 1$).

< TABLE 1 ABOUT HERE >

The choice of the output variables warrants some discussion, as there are at least two more output specifications commonly used in the transportation literature (see Jara Diaz (1982)).

The first is the use of demand-related output measures, such as the number of passenger-miles or passenger-trips. This output specification reflects the economic motive for providing the services, namely the transportation of passengers. But, there are at least four reasons not to pursue an output definition reflecting the consumption of transit services by users.[8] First, inputs do not vary very systematically with these demand-related output measures. Second, supply-related output indicators are to a larger extent under the control of operators than demand-related output specifications. Third, supplying the desired services for the least amount of resources is a condition for achieving broader social goals such as effectiveness, i.e., objectives defined in amounts of transit services actually consumed. Finally, it is unlikely that societal goals can be agreed upon by all parties involved. For instance, subsidy suppliers are interested in the efficient use of their funds, while operators are inclined to stress the effectiveness in terms of ridership, spatial availability of services, etc. These are probably also the reasons why the majority of the technical efficiency studies in urban transit uses pure supply indicators.[9]

[8] These four arguments are based on Berechman and Giuliano (1985) and on Berechman (1993), pages 97 and 152-153.

[9] Technical versus demand related output measures may yield different results: see Berechman and Giuliano (1985) for a classical production analysis and especially Chu, Fielding and Lamar (1992) in a frontier context. As pointed out by a referee, it could be useful to follow the latter study by analyzing any eventual differences between an efficiency versus an effectiveness approach. This is, however, outside the scope of this paper, but it is certainly a promising topic for future research.

The second output specification takes account of the spatial, temporal and quality characteristics of urban transit services.[10] In a parametric context the seminal article of Spady and Friedlaender (1978) has led to the specification of hedonic output composites which link the generic units times distance per unit time output concepts with the spatial, temporal and quality characteristics of the network.

In the non-parametric deterministic reference technologies, however, there is a serious problem in defining similar hedonic composite outputs. If in addition to the inputs and the generic outputs a large number of dimensions representing the characteristics are included, then this almost automatically leads to an increase in the efficiency scores and a larger number of efficient observations (see Kerstens and Vanden Eeckaut (1995 a,b) for details). Since there are no test procedures offering guidance in the selection of eventual additional dimensions, this could ultimately undermine the discriminatory power of the analysis. Therefore, in this study the limitations of the traditional aggregate output indicators are remedied by controlling for the spatial, temporal and quality characteristics in an explanatory analysis of the technical efficiency scores in an additional, second stage. The assumptions underlying this common practice in the efficiency literature are explained in section 4.1.

3. TECHNICAL EFFICIENCY IN FRENCH URBAN TRANSIT

3.1 Analysis based on the Complete Sample

The performance of this sample of French urban transport operators is gauged using the deterministic nonparametric deterministic reference technologies presented earlier. Radial efficiency measures in the outputs are calculated for the variable returns to scale DEA model with strong disposability in both inputs and outputs (denoted DEA-sd), the same model but assuming weak disposability in inputs and outputs (DEA-wd), and the FDH model (FDH).

The descriptive statistics of the resulting output measures of technical efficiency are found in Table 2.[11] The efficiency distributions for Model I and II are represented respectively in Figures 4 and 5. Before discussing these distributions in detail, the more elementary

[10] On the problem of accounting for these output characteristics: see Jara Diaz (1982) and Oum, Tretheway and Waters (1992).

[11] The problem of slacks in the measurement of technical efficiency is neglected. But this is unlikely to be important as the radial output efficiency measure leaves no slacks for the single output dimension in both models. A detailed discussion of slacks in DEA and FDH models is found in Kerstens and Vanden Eeckaut (1995).

classification into technically efficient and inefficient observations is briefly mentioned. The number of efficient observations generated by each production model is reported in the last column of Table 2. These numbers are largely determined by the theoretical relations holding between the efficiency measures evaluated relative to the different nonparametric models. It is evident that the number of efficient observations on a strongly disposable DEA model is smaller than the same number on respectively a weakly disposable DEA model or on FDH. While the relation between the latter two models is unclear a priori, it turns out that FDH yields by far the largest number of efficient observations.

< TABLE 2 ABOUT HERE >

< FIGURES 4 AND 5 ABOUT HERE >

On the one hand, the similarities common to the efficiency distributions of Model I and II can be summarized as follows. For a large part the distributions have central tendencies reflecting the a priori theoretical relations. In the nonparametric models, the mean technical efficiency is larger on the FDH than on the weakly disposable variable returns to scale DEA reference technology (DEA-wd), which itself has a larger mean than its strongly disposable variant (DEA-sd). The range is most narrow for FDH. Furthermore, all efficiency distributions are positively skewed and have a positive kurtosis, indicating long tails to the right and fat tails relative to the normal distribution. One exception is the efficiency measure on the FDH technology for Model I which is negatively skewed, i.e., it has a long and fat tail to the left.

On the other hand, the differences between Model I and II can be characterized as follows. Most importantly, Model II has a much lower central tendency and a larger range. Furthermore, Model II has fatter tails, with the exception of the FDH reference technology. Finally, the number of efficient observations is lower in Model II.

A comparison with related studies in the field of urban transport may prove instructive at this point. The results from the vehicle kilometres (Model I) output specification are certainly in line with the distribution of DEA efficiency scores reported in Chang and Kao (1992), Chu, Fielding and Lamar (1992), Obeng (1994), and Tone and Sawada (1990), and with the typically high average FDH-based efficiency measures as in Gathon (1989). The second output specification, however, yields a rather unusual low efficiency distribution. One plausible reason explaining the different results from both model specifications is mentioned by Cancalon and Gargaillo (1991). In their opinion, the standards used to convert vehicle into seat kilometres are inconsistent, leaving the companies with too much discretion for window dressing.

This large variation in the measured technical efficiency undoubtedly poses a serious problem if one takes the economic interpretation of the radial efficiency measure in terms of possible output expansions literally. It seems that one must be extremely careful in selecting a reference technology in terms of the plausibility of its underlying assumptions.

Of course, one can also interpret the efficiency measure in a more limited way as a device to rank organisations according to their performance. In this case it is important to assess the similarities in the rankings. These similarities are evaluated by looking at the Pearson product moment correlation coefficients, which are presented for both models in Table 3. The highest correlation is between the two DEA models. Observe that Model II yields higher similarities than Model I. This analysis of correlations reveals again that the choice of the reference technology must be given due attention as the implied rankings may differ markedly.

< TABLE 3 ABOUT HERE >

A more detailed look at the elementary classification of the observations as either technically efficient or inefficient allows to determine the impact of the reference technologies and their underlying assumptions. This classification is analysed using Tables 2 and 4.[12] The effect of imposing weak instead of strong disposability in DEA models is important. From the last column in Table 2 it is clear that the number of efficient observations more or less doubles. This result, as well as the difference in the average efficiency scores, indicates that there is a problem of congestion for these urban transit companies, as some of the observations are located on backward bending parts of the boundary.[13] If, relative to a sample of similar operators, a company's relative performance is hampered by congestion, it may, for instance, be induced to reconsider its labour organization, its network structure, etc. Dropping the convexity assumption is even more important, as the number of efficient observations in Table 2 at least quadruples if one compares FDH and the strongly disposable DEA model.

< TABLE 4 ABOUT HERE >

[12] The first two lines in Table 4 are simply the differences between the numbers of efficient observations reported in Table 2.

[13] Comparing technical efficiency measures evaluated on weakly and strongly disposable DEA models provides a basis for testing for specific factors contributing to congestion (see Färe, Grosskopf and Lovell (1994) for details).

A hiherto open issue is the relative importance of imposing either convexity and weak disposability, as in the weakly disposable DEA model, or non-convexity and strong disposability, as in FDH. From a theoretical viewpoint both reference technologies are equally entitled to claim that they incorporate the weakest assumptions. Otherwise formulated, one may wonder whether the economic phenomenon of congestion, which can be observed by means of a weakly disposable technology, is more or less important than the occurrence of nonconvexities, which are allowed for by the FDH technology. It turns out that imposing convexity affects the efficiency status of more observations than imposing weak disposability. Indeed, the last two lines in Table 4 show that in this sample more inefficient observations become efficient by imposing convexity than by imposing weak disposability. Although this difference is pronounced, more empirical evidence is needed to warrant the conclusion that convexity is more important than weak disposability from a practical point of view.

The results presented in this subsection indicate that from a methodological viewpoint the popularity of the strongly disposable reference technology relative to both its weakly disposable variant and FDH may be misleading, since the effect of both the disposability and the convexity assumptions is considerable. Therefore, in empirical work it is desirable to compute the impact of these assumptions.

3.2 *Analysis based on the Sample without Outliers*

This subsection first outlines the theoretical notion of an outlier in the context of production frontiers. Then, it evaluates empirically the impact of outliers on the measurement of technical efficiency for the nonparametric reference technologies used to evaluate the performance of French urban transit firms.

Any specification of the reference technology may be susceptible for outliers. Vaguely stated outliers are observations which are unlike the remainder of the data, or equivalently which do not follow the pattern of the majority of the data. It has been noted that frontier methodologies are especially vulnerable to outliers and leverage points since they concentrate on the extremes of the performance distribution, not on its central tendency (Sexton, Silkman and Hogan (1986: 81)). More specifically, among the nonparametric models, it is expected that efficient outliers have a larger effect in DEA than in FDH models (see Tulkens (1993)). In the latter each observation only spans an orthant, negative in the outputs and positive in the inputs. Consequently, the subset of observations which can be potentially influenced by the error is relatively small. In the former the possible impact is no smaller.[14]

The analysis of multivariate outliers in the nonparametric analysis of technical efficiency received previous attention from Burgess and Wilson (1993), Dusansky and Wilson (1994), Seaver and Triantis (1992), and Wilson (1993, 1995) among others. The outlier detection methods applied by these authors, however, are all sensitive to the masking effect by which a group of outliers in the same direction remains undetected. In a frontier context Janssens and Van Den Broeck (1993) and especially Seaver and Triantis (1995) have convincingly argued for the use of estimators with a high breakdown point, i.e., estimators for which a high fraction of the sample can be contaminated by arbitrary values before the finite sample bias becomes infinitely large. These robust estimators, developed by Rousseeuw and Leroy (1987) and Rousseeuw and Van Zomeren (1990), are least sensitive to the masking effect. Before using them in the empirical application, the basic intuition behind the estimators with a high breakdown point is briefly indicated.[15]

In the context of multivariate analysis the classical Mahalanobis distance, which measures how far a random vector is from the middle of its distribution taking into account the shape of the multivariate cloud, can be made more robust. For each observation this traditional Mahalanobis distance (MD) is defined as:

$$MD(z_i,Z) = \sqrt{(z_i - T(Z))C(Z)^{-1}(z_i - T(Z))^t},$$

where Z = (Y X) is a k x (n + m) matrix containing all variables, and T(Z) and C(Z) are respectively the row vector of arithmetic sample means and the sample covariance matrix. These first two moments, however, are vulnerable to the masking effect: any group of outliers attracts T(Z) and magnifies C(Z). Therefore, the minimum volume ellipsoid (MVE) estimator is proposed which has a breakdown point of 50%. This means that slightly less than half of the data points can be replaced by arbitrary values without affecting the estimates. T(Z) is redefined as the centre of the MVE covering half of the observations and C(Z) is adjusted accordingly. Inserting these new estimates of the first two moments in the Mahalanobis distance yields a robust Mahalanobis distance (RD). Because the squared MD

---

[14]It depends partly on the returns to scale assumptions, with constant returns to scale having the largest impact.

[15] Details are in Rousseeuw and Leroy (1987) and Rousseeuw and Van Zomeren (1990).

follows a chi-squared distribution with degrees of freedom equal to the number of variables analysed, it is possible to construct a confidence ellipse. The same holds true for its more robust variant. Observations outside this ellipse are potential outliers.[16]

In line with the above analysis, outliers in the French urban transit data are determined as follows. A computation of RD for all observations and for both model specifications yields the following results: for Model I and II respectively 43 and 42 observations are outside the 97.5 confidence ellipse. Once outliers are detected various possibilities remain open. One radical response is to discard all outliers. Other reactions, however, are equally plausible: data can be examined and corrected, models can be adjusted, etc. In this empirical analysis all outliers are removed, since there is no independent source yielding sufficient information to pursue any of the other possibilities. However, since all reference technologies are treated alike, there is no risk for a systematic bias in the results.

All of these dubious observations have been discarded, instead of only eliminating the efficient outliers. The reason is that the number of outliers which is efficient depends on the specification of the reference technology. For instance, among the outliers respectively 37 and 9 observations are efficient on a FDH technology while respectively 6 and 2 of these outliers are efficient on a strongly disposable DEA model. Eliminating all outliers allows one to compare the impact of outliers for the different technologies on a sample of equal size.

For the samples without outliers, descriptive statistics for input and output variables in Model I and II are found in the second part of Table 1. The distributions have shifted downwards, indicating that the outliers are especially found among the larger operators. Furthermore, when comparing the ranges of the variables it becomes apparent that size differences have become less pronounced.

Descriptive statistics for the output efficiency measures computed on the three nonparametric frontier methodologies on a sample without outliers are presented in Table 2. The elementary classification between efficient and inefficient observations is similar to the one obtained before and merits no further discussion. Comparing the descriptive statistics

[16] A referee suggested that there is a dilemma between on the one hand using deterministic frontier techniques and on the other hand discussing potential outliers based on making assumptions regarding distributions. I agree, but would like to make two remarks. First, some recent research efforts concentrate on adding stochastic assumptions to the non-parametric frontier techniques (see Lovell (1993)), but so far no consensus has emerged on the best way to proceed. If this debate comes to an end, then the referee's contention would disappear. Second, while there are some outlier detection methods where assumptions regarding distributions of the data need not be made, this is not (yet) the case for the robust methods opted for in this study.

with the complete sample reveals that for Model I the distributions resulting from the DEA models have shifted upwards, while the FDH distribution remains almost the same. The effect of outliers on Model II is more pronounced: for all models the central tendency and the minimum almost double. In addition, the sharp increase in the number of efficient observations indicates that in the initial data set some observations have a tremendous influence on the overall performance evaluation. Clearly, the results of both model specifications have become very similar, which is also evident at a glance from Figures 6 and 7. Moreover, the remarkable difference with the typical efficiency results reported in the urban transit literature has disappeared.

< FIGURES 6 AND 7 ABOUT HERE >

Comparing results on weakly and strongly disposable DEA models confirms that there is a problem of congestion for at least some of these urban transit companies. But dropping convexity is again most important in terms of both the efficiency distribution and the number of efficient observations. Consequently, in applied work computing the effect of these different assumptions is strongly recommended.

As is evident from Table 3, the effect of eliminating potential outliers on the rankings is minor, especially for Model II. Therefore, once outliers have been accounted for vehicle and seat kilometres seem to provide very similar output specifications of technology. The conversion of vehicle to seat kilometres is only a problem for outlying observations.

Bearing in mind that frontier technologies only allow for comparisons relative to observed best practice, one can conclude that technical efficiency distributions in France do not seem to differ from those observed elsewhere. Although the difficulties to compare transit operators of different countries are extremely high, it is useful to add that there are indications that French companies are performing on or below the average in a European context (see Gathon (1989)).

This subsection has provided additional evidence that the effect of outliers is potentially important when using deterministic frontier methods for performance gauging. From a methodological perspective, it seems advisable that the use of appropriate outliers diagnostics becomes standard practice when calculating efficiency on deterministic nonparametric production frontiers.

4. EXPLANATION OF TECHNICAL EFFICIENCY PATTERNS

4.1 Methodological Introduction

While the theoretical basis for explaining technical efficiency is weak relative to the refined methodologies for measuring technical efficiency, it is necessary to go beyond the pure measurement of performance (as stressed in Button and Weyman-Jones (1994)). Investigating the determinants of technical efficiency is particularly useful for formulating policies aimed at improving performance.

Following common practice, this section is the second stage of a two stage procedure to the statistical explanation of technical efficiency. In the first stage, i.e., section 3, technical efficiency is evaluated on a certain reference technology; in the second stage these first stage results are explained by a series of relevant variables using appropriate statistical techniques. The crucial assumption underlying this second stage is that the explanatory variables only influence technical efficiency, but not the transformation process from inputs into outputs analysed in the first stage (see Lovell (1993: 53-54) for a detailed discussion).

Explaining technical efficiency measures requires the specification of an appropriate multivariate statistical model. Due to the use of nonparametric production frontiers, the dependent variable always has a nonnegligible proportion of observations with an efficiency score of unity. Consequently, the statistical analysis by means of a series of observed characteristics of the urban transit operators is most often based on censored regression models.[17] This censoring process reflects the relative nature of the frontier technologies and the fact that technical efficiency measured relative to these inner bound technologies only yields an upper bound. Since in our analysis the output efficiency measures have been defined to be no larger than unity, a standard Tobit model with upper censoring at unity is employed.

4.2 Determinants of Performance

It is possible to distinguish between five main categories of determinants of technical efficiency (see Caves and Barton (1990), Caves (1992) and Pestieau and Tulkens (1993)). First, competitive conditions are thought to foster technical efficiency. Second, a series of

[17] See Berechman (1993: 180) or Lovell (1993: 53).

organizational factors may affect the productive efficiency of activities. These factors include the size and diversification of organizations, its labour organization, the ownership structure, among others. Third, the heterogeneity between organizations being evaluated can cause structural differences in observed patterns of technical efficiency. This includes heterogeneity in production processes, in market structures, etc. Fourth, technical efficiency is affected by dynamic factors. This includes, among others, capital-vintage effects, market growth, and innovations. Finally, government regulation can constrain the choices of producers in a variety of ways, which can have an adverse impact on observed technical efficiency.

Right from the start it must be stressed that there has been little research based on frontier techniques aiming to explain the technical efficiency of urban transit operations. The scattered evidence will be discussed below in connection with the findings for the French urban transit sector. There does, however, exist a large non-frontier literature on urban transit performance. Attention is limited to the findings for which a reasonable degree of consensus seems to exist.

First, the survey by Perry, Babitsky and Gregersen (1988) on the effect of ownership and management systems on the performance of urban transit firms reports mixed results. First, variations in ownership and management have little predictable associations with operating efficiency. Although, it must be admitted that one can discern a trend that in the more recent studies private ownership performs better. Probably, this is because worldwide there has been a tendency in the last decades for unprofitable private suppliers to become publicly owned. This selection process may bias performance comparisons. Also the use of outside expertise under the form of contract management is no guarantee for improved performance. Second, the tendency of revenues to fall short of costs has increased over time. This trend was probably strengthened by the increasing level of subsidies. Third, service levels are higher under public ownership.

Second, recent deregulation discussions and experiences, which counter the historical trend towards increasing public sector involvement in urban transit, can shed some light on the importance of competition and regulation in urban transit. The deregulation experience of local bus services in the U.K. has led to substantial cost reductions among all operators, mainly due to productivity enhancing working practices, the increased use of minibuses, a reduction in the wage rates and a deterioration of the service or output levels offered (see Heseltine and Silcock (1990)). Also the requirement that subsidized bus services should be subjected to competitive tendering, i.e., a bidding process for the monopoly right to supply a predefined service at a particular spatial level during a particular period, has lowered subsidy costs by about 20% (Glaister, Starkie and Thompson (1990)).[18]

4.3 Potential Determinants of French Urban Transit Performance

The potential explanatory elements in our sample are classified according to the previously presented scheme. This range of explanatory factors is partly determined by data availablility.

First, there is little or no competition in the sector. The French laws allow to use either a competitive tendering procedure, or negotiations to choose among the potential operators. Unfortunately, no data are available to determine which contractual agreements between operators and organizing authorities are based on either procedure. In general, the perception is that competitive tendering procedures are little used (see Gwilliam and van de Velde (1990: 339)). Berechman (1993: 272) conjectures that this lack of competition despite an important number of private operators is due to the perception of transit services as merit goods, which are deemed indispensible for the functioning of urban areas.

Second, organizational differences may account for part of the observed patterns of technical efficiency. A variety of elements, on which information is available, come to mind. Firstly, the ownership status of the urban transit companies is available for most of the companies in the sample. The ownership variable (OWNER) is defined as a dummy taking the value of unity only when the bus company is privately owned. Public and semi-public operators are represented by a zero value. Secondly, since many transport operators are member of a group or a management association, a dummy variable (named GROUP) is defined taking a value of unity if the operator is a group member and zero otherwise. The impact of the membership of a group on performance is a priori unclear.[19] On the one hand, there may exist strategic advantages in operating costs unrelated to driving (e.g., maintenance). On the other hand, this membership may enhance technical inefficiency by increasing the regulatory capture of various government tiers (e.g., in renegotiations to prolong monopoly contracts).

---

[18] Banister, Berechman and De Rus (1992) discuss the contestability of transit markets in view of the U.K. experience.

[19] This is also the interpretation of Berechman (1993: 273).

Third, the heterogeneity in the production process and in the operating environment should be considered whenever possible. In the case of urban transit it seems useful to complement the traditionally used output aggregates with information on the spatial, temporal and quality characteristics of the services. Unfortunately, no details are available concerning temporal and quality aspects, such as base versus peak levels or timeliness of services. There is some information on spatial characteristics of the network: its total length, the number of lines, and the total number of stops. There is some evidence that the network length and the number of stops served in the network have a negative effect on performance (see Fazioli, Filippini and Prioni (1993)). But since this information is still insufficient to get a detailed picture of size and structure of an urban transit network, it is safer to assume that it is a priori unclear how these variables affect technical efficiency.[20] The available information is incorporated by defining two variables proxying for the spatial structure of the network: the mean length of a line (LINELENGTH), i.e., the ratio of network length to the number of lines; and the average distance between stops (STOPLENGTH), defined as the ratio of network length to the number of stops.

Furthermore, the heterogeneity of the environment can be accounted for by a limited number of variables. First, for some companies the average commercial speed with which they are able to supply their transportation services is available (SPEED). Following Gathon (1989), this average speed is postulated to have a positive effect on technical efficiency. Although this speed is largely determined exogenously, especially by the extent of the congestion in the urban area, it is an important qualitative characteristic of the services. Second, since the population and the area within the transport perimeter of the companies are known, it is possible to define the density of the urban environment (POPDENS). On the one hand, population density can be interpreted as a mere proxy for the intensity of traffic and congestion in the urban area, in which case it provides a similar function as the average speed variable. On the other hand, population density may have cost saving effects in the design and the structure of the transportation network, in that, e.g., fewer lines or stops are needed (Berechman (1993: 19)). Concluding, a priori the effect of population density is ambiguous.

Fourth, information on dynamic factors is hardly available. Only the effect of

[20] This is emphasized by Fazioli, Filippini and Prioni (1993) themselves. They suggest to use dispersion indexes from the theory of graphs, but this requires information for each operator that is unavailable in published statistics.

capital-vintages on perceived technical efficiency is proxied by the average age of bus vehicles in years (VEHAGE). Age may adversely affect the maintenance of the vehicles as well as their fuel consumption. While the latter effect is already accounted for by the fuel input in the production models, this explanatory variable captures the fact that vehicles become less operational over the years.

Finally, the impact of government regulation must be carefully assessed. In the French urban transit case, a number of aspects can be distinguished.

Firstly, the organising authority either sets up a partially or fully independent public administration (régie), or it delegates the urban transit operation to a company within the framework of an agreement. As discussed earlier, these agreements are classified according to four major contract types, which differ mainly in the degree of risk-sharing agreed on. A variable representing the risk-sharing properties of the contracts (CTYPE) is defined as follows. As there are no companies with a full-risk contract in the sample, it is convenient to aggregate the remaining contract types into a simple dummy variable. Accordingly, a distinction is made between on the one hand public administrations and management contracts, which assume no risk at all, and on the other hand fixed price contracts and contracts with a guarantee of revenue, which at least involve some risk-bearing on behalf of the operator. Only for the latter the dummy takes the value of unity. Obviously, one would expect that shifting the risks partly to the operators stimulates performance.

Another aspect of these contracts is their fixed duration. Data on the duration of the contracts in years are available in a variable named CTERM. While the early literature on competitive tendering stressed the risk that the incumbent could benefit from a too long term, the more recent literature on the contestability and sustainability of natural monopolies emphasizes the importance of sunk-costs.[21] Thus, a priori the effect of the term of the contract on technical efficiency is unclear.

Secondly, the financing of the urban transit system may affect its performance. There are two aspects which can be studied given the available information. The first aspect is the relative amount of the subsidy covering the difference between operational revenues and costs. This leads to the definition of a variable containing the share of subsidies in total operating costs (SSUB). There is, of course, already a variable describing the eventual risk-

---

[21] See Banister, Berechman and De Rus (1992) for a detailed discussion in the context of urban transit.

sharing agreed on in the contract and its implications on the subsidy formula. In view of the received literature mentioned earlier, it is a priori unclear how the relative amount of the subsidies affects the performance when the risk-sharing properties of the contract are already controlled for. Note that only the aggregate subsidy is known, not the variety of its sources.

The second aspect is the rate of the local transportation tax on the wage bill, which is earmarked for supporting public transport (TAX). One can conjecture that in general high tax rates increase the monitoring effort of citizens and, indirectly, regulators.[22] This argument is even strengthened in the case of an earmarked tax levied at the local level.

To recapitulate, this section has reviewed at length a number of variables which potentially determine the performance of French urban transit companies. From this examination of available data, it is also clear that a series of other potential determinants are lacking. For example, notwithstanding the fact that about 60% of the workforce is unionised (Gwilliam and van de Velde (1990: 338)), there is no information available on the degree of unionization among employees for each operator.

4.4 *Technical Efficiency of French Urban Transit Companies: Tobit Results*

As it turned out that the efficiency measures calculated on the nonparametric models are partly sensitive to outliers, especially when seat kilometres are used as output, there are good reasons to give a major weight to the results for the sample without outliers. The bottom part of Table 1 reports sample statistics on the potential explanatory variables outlined in the previous section. Most of the variables show a lot of variation. This holds even true for the tax rate which is bound to vary within certain legal limits.

Unfortunately, various companies fail to report data for some explanatory variables. An analysis of the complete cases only, i.e., the operators for which all variables are present, reduces the sample size drastically. These missing data do not imply a problem for estimation purposes if the underlying mechanism is of the "ignorable" kind (see, e.g., Greene (1993)). Loosely stated, the gaps in the data should not be related to the phenomenon being modelled. A simple nonparametric test statistic does not reject the hypothesis that the missing data

[22] In a different context, i.e., local public good provision, Davis and Hayes (1993) found evidence of a positive relation between tax rates and monitoring effort.

mechanism in the sample is of the "ignorable" kind.[23] A preliminary analysis also reveals that there is multicollinearity between average commercial speed (SPEED) and population density (POPDENS).[24] Given the interpretation of both variables, this is taken as evidence that they are essentially measuring the same phenomenon. This close collinearity is exploited to fill up some of the missing cases in the commercial speed variable.

The results for the Tobit estimation are summarized in Table 5 for the samples without outliers, where the latter are further adjusted for the availability of explanatory variables. The largest set of restrictions common to all dependent variables sets the group, the commercial speed, and the average vehicle age variables equal to zero. The fact that commercial speed (or population density) does not influence technical efficiency may indicate that congestion in the urban operating environment does not affect the performance of operators. If this interpretation is valid, then differences between strongly and weakly disposable DEA models can be solely attributed to congestion within companies. The absence of vintage effects could indicate that state subsidies are in effect spread evenly among operators.

< TABLE 5 ABOUT HERE >

As expected, not all of the coefficients can be estimated with precision. This is especially true for Model I, and in particular for its weakly disposable DEA model. Furthermore, since this Tobit analysis is based on a limited subsample the results require careful interpretation. However, despite the quantitative differences among the production models and according to the reference technology used, the signs of the variables are quite stable. Given these caveats, the conclusions to be inferred from this table can be summarized as follows.

First, private ownership has a positive effect on urban transit performance, although the statistical significance of the coefficients is not very pronounced. Second, heterogeneity in the network of services clearly justifies some of the differences in technical efficiency. The mean length of a line (LINELENGTH) and the average distance between stops (STOPLENGTH) have respectively positive and negative effects. This cannot be taken

[23] To be specific, a Mann-Whitney U test is computed for the various efficiency measures comparing their distributions conditional on the availability or not of all the explanatory variables. The null hypotheses that the dependent variables share a common distribution cannot be rejected.

[24] This analysis is performed using the condition indexes defined in Belsley, Kuh and Welsch (1980). Details are suppressed due to space limitations.

literally to imply that, for instance, the number of stops should be reduced to enhance performance. These results, however, do indicate that network structure affects operating efficiency. Third, the risk-sharing dummy and the duration of the contract both are positively related to technical efficiency. The former effect, in line with expectations, confirms the positive role of risk-sharing agreements in regulated industries. Finally, while subsidies subvert technical efficiency, the transportation tax promotes the performance of companies, most plausibly by increased monitoring. The former result indeed indicates a negative effect of the relative share of subsidies independent of the positive, disciplining impact of risk-sharing arrangements. Since adding an interaction term between the risk-sharing dummy and the share of subsidies does not add to the explanation, this effect seems to be valid for all companies. One can wonder to which extent the lack of competition (or yardstick competition) in French urban transit is responsable for this perverse side-effect of regulatory policy.

For policy purposes it is of particular interest to have a knowledge of the partial effects implied by these estimates. Since it can be assumed that policy makers have an obvious interest in improving the performance of the inefficient observations, it is useful to concentrate on the effect of the independent variables on the observed variable below the unity censoring level. Therefore, for each equation a multiplicative correction factor is calculated, evaluated at the sample means, transforming the estimates into partial effects for the truncated sample. These correction factors are reported in the last line of the table.

5. CONCLUSIONS

The empirical analysis of French urban transit companies yields interesting answers to both methodological and policy-oriented issues.

To begin with the former. The study confirmed the significance of the choice between deterministic nonparametric reference technologies for technical efficiency measurement. The wide dispersion in the radial output efficiency measures complicates their economic interpretation, as they all claim to evaluate the technical efficiency of the same sample using the same efficiency measure. It was evidenced that some of the underlying assumptions have more important consequences than others. Especially dropping the convexity assumption turns out to be more important than weakening the disposability assumptions. This finding questions the popularity of the strongly disposable DEA models and indicates that the underlying production assumptions merit closer attention in empirical applications. The assessment of the impact of outliers relied on a recent methodology based on robust estimators with a high breakdown point. It confirms that nonparametric frontier methodologies can indeed be vulnerable to outliers.

Next, a Tobit model has been discussed explaining the technical efficiency scores. Using a general framework for explaining and interpreting the technical efficiency of organizations, some of the specific determinants of urban transit performance in general and of the French transit companies in particular have been extensively reviewed. In terms of the policy debates mentioned in the introduction, the empirical findings confirm the importance of appropriate incentives in contracting for monopoly. Risk-sharing agreements and, to a lesser extent, private ownership both seem to spur the performance of organizations. Also the detrimental effects of subsidies on the one hand and the indirect disciplining impact of the unique French transportation tax are worth recalling. Finally, the network structure may partly account for differences in performance. Of course, since information was lacking on some potential determinants, e.g., the limited use of competition in France could not be controlled for, these conclusions should be treated with some caution.

While most of these findings are in line with the urban transit literature, this contribution has systematically explored these relationships on a single data set using several modern frontier methods. Obviously, the analysis reported in this paper is only a first attempt to evaluate the French urban transit performance. There remains ample room for additional studies corroborating or falsifying the detected determinants.

APPENDIX

The radial efficiency measure in the outputs ($DF_o(x,y)$) evaluated relative to DEA-sd is calculated by solving the following linear program for each observation (x°, y°):

$$\begin{aligned} &\max_{\mu, z} \ \mu \\ \text{s.t. } & Y^t z \geqq y^\circ \mu \\ & X^t z \leqq x^\circ \\ & I_k^t z = 1 \\ & \mu \geqq 0,\ z \geqq 0 \end{aligned}$$

To obtain a radial efficiency measure in the outputs which is no larger than unity, this LP must be modified in two ways. First, the objective function becomes:

$$\min_{\mu', z} \ \mu'$$

Second, the constraints on the outputs now are:

$$Y^t z \geqq y^\circ / \mu'.$$

Evidently, solving the first LP problem and calculating $\mu' = 1/\mu$ is identical to solving the second problem (see Färe, Grosskopf and Lovell (1985: 198)).

Relative to DEA-wd $DF_o(x,y)$ is determined by solving for each observation (x°, y°) the following nonlinear program:

$$\begin{aligned} &\max_{\mu, z, \gamma, \delta} \ \mu \\ \text{s.t. } & \gamma Y^t z = y^\circ \mu \\ & X^t z = \delta x^\circ \\ & I_k^t z = 1 \\ & \mu \geqq 0,\ z \geqq 0 \\ & 0 < \gamma \leqq 1 \\ & 0 < \delta \leqq 1 \end{aligned}$$

This problem can be reparametrized into a simpler LP problem (see Färe, Grosskopf and Lovell (1985)). A radial efficiency measure in the outputs which is bounded above by unity can be obtained in a way similar as above.

On the FDH model $DF_o(x,y)$ is calculated by solving for each observation (x°,y°) the same programming problem as for DEA-sd, except that one constraint is added:

$$z_i \in \{0,1\} \quad \text{for } i = 1,\dots,k.$$

This mixed integer programming problem can be solved by a simple vector dominance algorithm in two steps (see Tulkens (1993)):

(i) Define for each observation (x°,y°) to be evaluated an index set $DO(x^\circ,y^\circ)$ containing the observations which weakly dominate (x°,y°) in that they produce at least as much of each output with no more of any input:

$$DO(x^\circ,y^\circ) = \{(x_i,y_i) \mid x_i \leq x^\circ,\ y_i \geq y^\circ\}.$$

(ii) Calculate $DF_o(x,y)$ by applying the following algorithm:

$$DF_o(x,y) = \max_{(x_i,y_i)\in DO(x^\circ,y^\circ)} \ \min_{j=1,\dots,n} \left(\frac{y_{ij}}{y_{\circ j}}\right).$$

To obtain a radial efficiency measure in the outputs which is no larger than unity, this second step needs a straightforward modification.

REFERENCES

Banister, D., J. Berechman, G. De Rus (1992) Competitive Regimes within the European Bus Industry: Theory and Practice, *Transportation Research, Part A: General*, 26(2), 167-178.

Belsley, D., E. Kuh, R. Welsch (1980) *Regression Diagnostics: Identifying Influential Observations and Sources of Collinearity*, New York, Wiley.

Berechman, J. (1993) *Public Transit Economics and Deregulation Policy*, Amsterdam, North-Holland.

Berechman, J., G. Giuliano (1985) Economies of Scale in Bus Transit: A Review of Concepts and Evidence, *Transportation*, 12(4), 313-332.

Burgess, J., P. Wilson (1993) Technical Efficiency in Veterans Administration Hospitals, in: H. Fried, C.A.K. Lovell, S. Schmidt (eds) *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford, Oxford University Press, 335-351.

Button, K., T. Weyman-Jones (1994) X-efficiency and Technical Efficiency, *Public Choice*, 80(1-2), 83-104.

Cancalon, F., L. Gargaillo (1991) *Les Transports Collectifs Urbains: Quelles Méthodes pour Quelle Stratégie?*, Paris, Celse.

Caves, R., D. Barton (1990) *Efficiency in U.S. Manufacturing Industries*, Cambridge (Mass.), MIT Press.

Caves, R. (1992) Introduction and Summary, in: R. Caves (ed) *Industrial Efficiency in Six Nations*, Cambridge (Mass.), MIT Press, 1-27.

Chang, K-P., P-H. Kao (1992) The Relative Efficiency of Public versus Private Municipal Bus Firms: An Application of Data Envelopment Analysis, *Journal of Productivity Analysis*, 3(1-2), 67-84.

Chu, X, G. Fielding, B. Lamar (1992) Measuring Transit Performance Using Data Envelopment Analysis, *Transportation Research, Part A: General*, 26(3), 223-230.

Davis, M., K. Hayes (1993) The Demand for Good Government, *Review of Economics and Statistics*, 75(1), 148-152.

Debreu, G. (1951) The Coefficient of Resource Utilization, *Econometrica*, 19(3), 273-292.

Dusansky, R., P. Wilson (1994) Technical Efficiency in the Decentralized Care of the Developmentally Disabled, *Review of Economics and Statistics*, 76(2), 340-345.

Färe, R., S. Grosskopf, C.A.K. Lovell (1985) *The Measurement of Efficiency of Production*, Boston, Kluwer.

Färe, R., S. Grosskopf, C.A.K. Lovell (1994) *Production Frontiers*, Cambridge, Cambridge University Press.

Farrell, M. (1957) The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society Series A: General*, 120(3), 253-281.

Fazioli, R., M. Filippini, P. Prioni (1993) Cost-Structure and Efficiency of Local Public Transport: The Case of Emilia Romagna Bus Companies, *International Journal of Transport Economics*, 20(3), 305-324.

Gathon, H-J. (1989) Indicators of Partial Productivity and Technical Efficiency in the European Urban Transit Sector, *Annals of Public and Cooperative Economics*, 60(1), 43-59.

Glaister, S., D. Starkie, D. Thompson (1990) The Asssessment: Economic Policy for Transport, *Oxford Review of Economic Policy*, 6(2), 1-21.

Greene, W. (1993) *Econometric Analysis*, Second Edition, New York, Macmillan.

Gwilliam, K., D. van de Velde (1990) The Potential for Regulatory Change in European Bus Markets, *Journal of Transport Economics and Policy*, 24(3), 333-350.

Heseltine, P., D. Silcock (1990) The Effects of Bus Deregulation on Costs, *Journal of Transport Economics and Policy*, 24(3), 239-254.

Janssens, G., J. Van Den Broeck (1993) Outliers, Sample Size and Robust Estimation of Stochastic Frontier Production Models, *Journal of Information & Optimization Sciences*, 14(3), 257-274.

Jara Diaz, S. R. (1982) The Estimation of Transport Cost Functions: A Methodological Review, *Transport Reviews*, 2(3), 257-278.

Kerstens, K., P. Vanden Eeckaut (1995a) Technical Efficiency Measures on DEA and FDH: A Reconsideration of the Axiomatic Literature, Louvain-la-Neuve, UCL (CORE Discussion Paper 9513).

Kerstens, K., P. Vanden Eeckaut (1995b) The Economic Cost of Subsidy-Induced Technical Inefficiency: A Methodological Postscript, *International Journal of Transport Economics*, 22(2), 225-229.

Koopmans, T., (1951) Analysis of Production as an Efficient Combination of Activities, in: T. Koopmans (ed) *Activity Analysis of Production and Allocation*, New Haven, Yale Univ. Press, 33-97.

Lovell, C.A.K. (1993) Production Frontiers and Productive Efficiency, in: H. Fried, C.A.K. Lovell, S. Schmidt (eds) *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford, Oxford University Press, 3-67.

Ministère de l'Equipement, du Logement, des Transport et de l'Espace (1990) *Transports Publics Urbains en France: Organisation Institutionnelle / Urban Public Transport in France: The Institutional Organisation*, Paris, Direction des Transport Terrestres.

Ministère de l'Equipement, du Logement, des Transport et de l'Espace (1991) *Données et Analyses sur les Transports Collectifs Urbains*, Paris, Direction des Transport Terrestres.

Mitric, S. (1988) Organization of Urban Public Transport in France: Lessons for Developing Countries, *Transportation Research Record*, 1202, 113-120.

Obeng, K. (1994) The Economic Cost of Subsidy-induced Technical Efficiency, *International Journal of Transport Economics*, 21(1), 3-20.

Oum, T., M. Tretheway, W. Waters (1992) Concepts, Methods and Purposes of Productivity

Measurement in Transportation, *Transportation Research, Part A: General*, 26(6), 493-505.

Perry, J., T. Babitsky, H. Gregersen (1988) Organizational Form and Performance in Urban Mass Transit, *Transport Reviews*, 8(2), 125-143.

Pestieau, P., H. Tulkens (1993) Assessing and Explaining the Performance of Public Enterprises, *Finanzarchiv*, 50(3), 293-323.

Pucher, J. (1988) Urban Public Transport Subsidies in Western Europe and North America, *Transportation Quarterly*, 42(3), 377-402.

Rousseeuw, P., A. Leroy (1987) *Robust Regression and Outlier Detection*, New York, Wiley.

Rousseeuw, P., B. van Zomeren (1990) Unmasking Multivariate Outliers and Leverage Points, *Journal of the American Statistical Association*, 85(411) 633-639.

Seaver, B., K. Triantis (1992) A Fuzzy Clustering Approach Used in Evaluating Technical Efficiency Measures in Manufacturing, *Journal of Productivity Analysis*, 3(4), 337-363.

Seaver, B., K. Triantis (1995) The Impact of Outliers and Leverage Points for Efficiency Measurement and Evaluation Using High Breakdown Procedures, *Management Science*, 41(6), 937-956.

Sexton, T. R. Silkman, A. Hogan (1986) Data Envelopment Analysis: Critique and Extensions, in: R. Silkman (ed) *Measuring Efficiency: An Assessment of Data Envelopment Analysis*, San Francisco, Jossey-Bass, 73-105.

Spady, R., A. Friedlaender (1978) Hedonic Cost Functions for the Regulated Trucking Industry, *Bell Journal of Economics*, 9(1), 159-179.

Tone, K., T. Sawada (1990) An Efficiency Analysis of Public vs. Private Bus Transportation Enterprises, in: H. Bradley (ed) *Operational Research '90*, New York, Pergamon Press, 357-365.

Tulkens, H. (1993) On FDH Efficiency Analysis: Some Methodological Issues and Applications to Retail Banking, Courts, and Urban Transit, *Journal of Productivity Analysis*, 4(1-2), 183-210.

Viton, P. (1986) The Question of Efficiency in Urban Bus Transportation, *Journal of Regional Science*, 26(3), 499-513.

Wilson, P. (1993) Detecting Outliers in Deterministic Nonparametric Frontier Models with Multiple Outputs, *Journal of Business & Economic Statistics*, 11(3), 319-323.

Wilson, P. (1995) Detecting Influential Observations in Data Envelopment Analysis, *Journal of Productivity Analysis*, 6(1), 27-45.

Figure 1: Output correspondence of a strongly disposable DEA technology

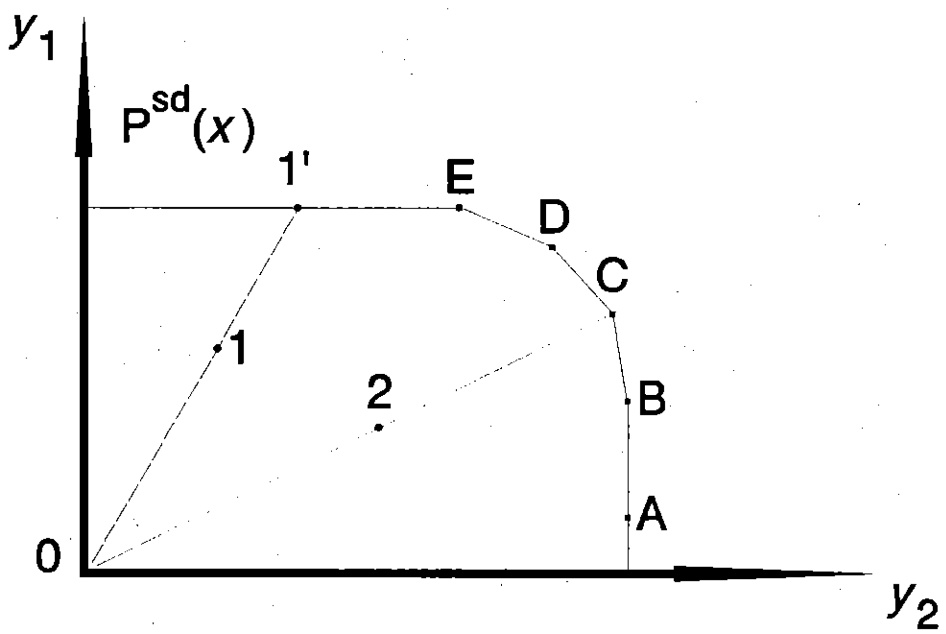

Figure 2: Output correspondence of a weakly disposable DEA technology

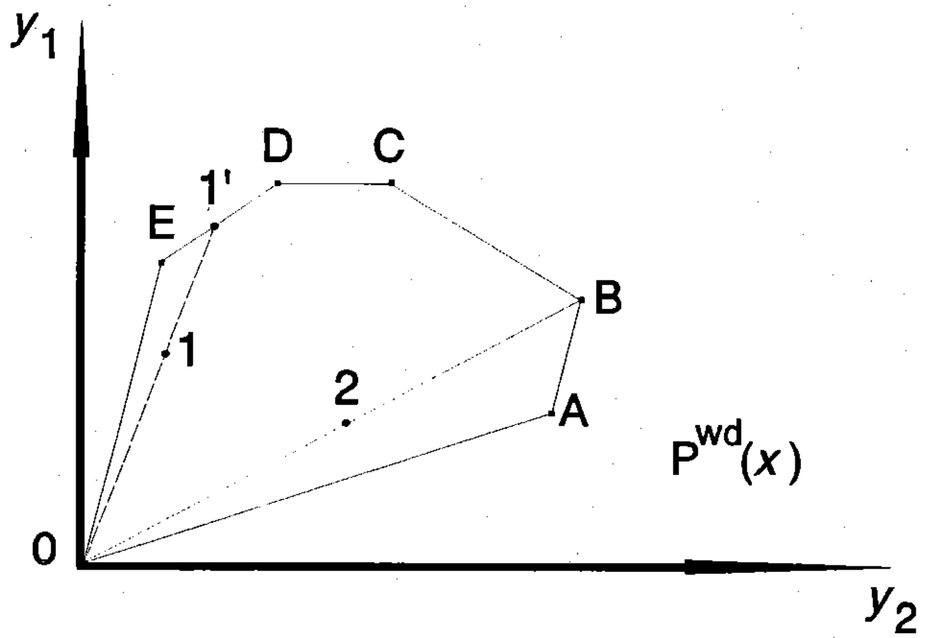

Figure 3: Output correspondence of an FDH technology

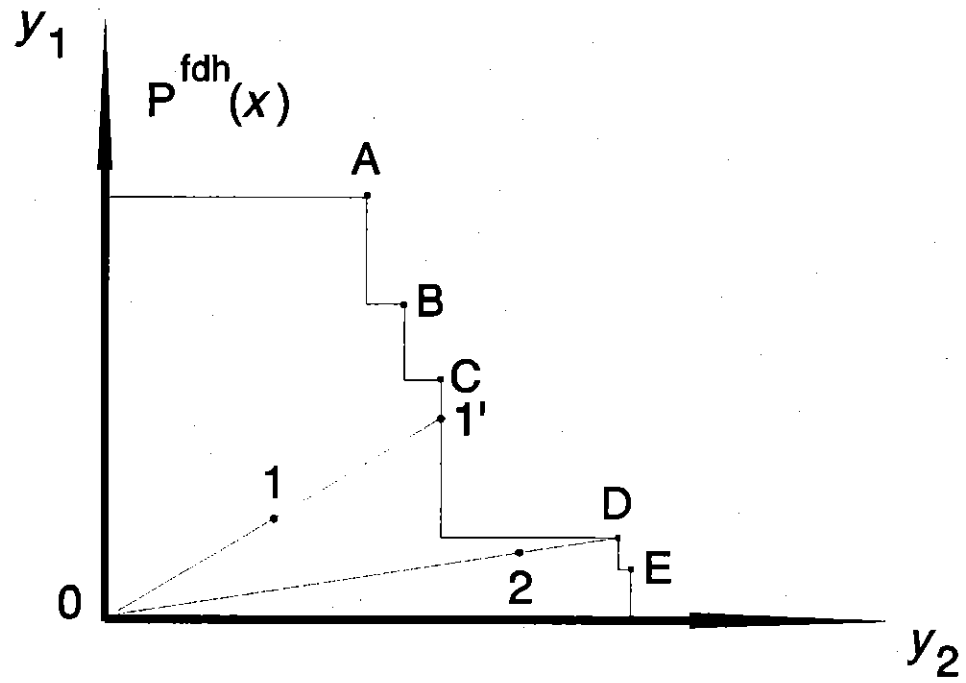

Figure 4: Densities of output efficiency measures (Model I)

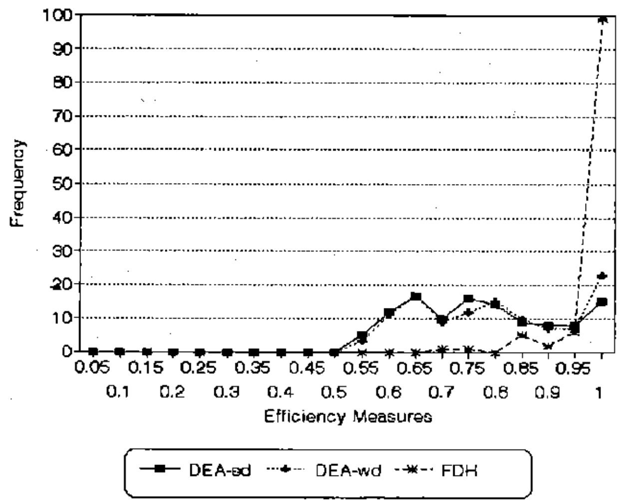

Figure 5: Densities of output efficiency measures (Model II)

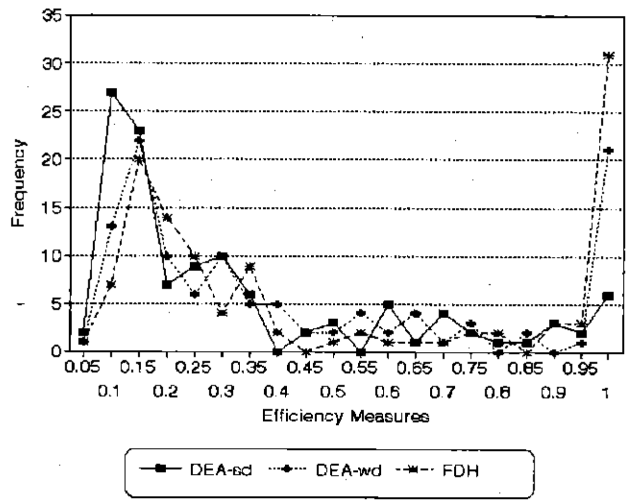

Figure 6: Densities of output efficiency measures without outliers (Model I)

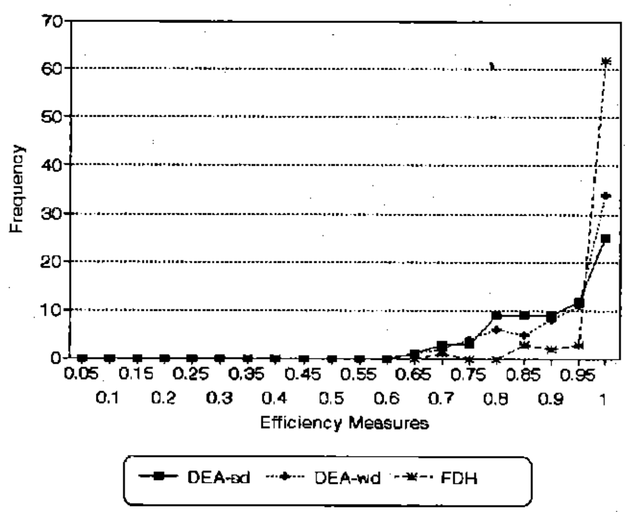

Figure 7: Densities of output efficiency measures without outliers (Model II)

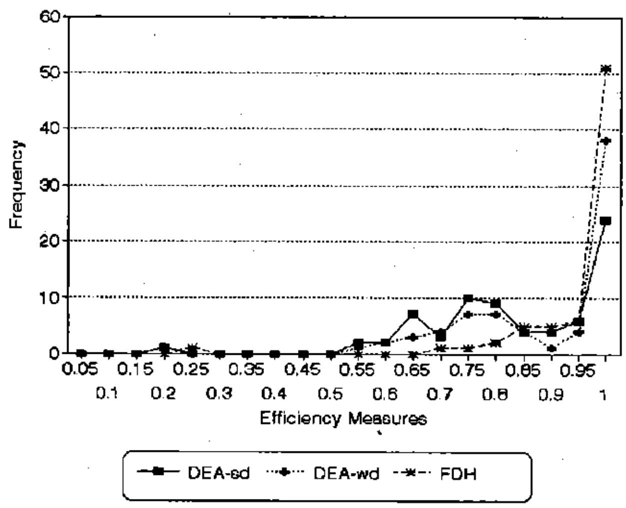

Table 1: Summary statistics on French urban transport

	Mean	Standard Deviation	Minimum Value	Maximum Value
All Observations (N=114)				
Outputs:				
VEHICLE KMS (y_1)	2416.1	3290.7	110	213.5E+02
SEAT KMS (y_2)	237360	315110	5164	200.0E+04
Inputs:				
VEHICLES (x_1)	63.44	85.46	4	521
EMPLOYEES (x_2)	149.82	248.44	5	1662
FUEL (x_3)	972.37	1433.3	39	8986
Observations without Outliers (N=71 & N=72)				
Outputs:				
VEHICLE KMS (y_1)	728.70	527.82	110	2570
SEAT KMS (y_2)	75586	63570	5164	253E+03
Inputs:*				
VEHICLES (x_1)	21.11	14.40	4	70
	24.00	18.71	4	92
EMPLOYEES (x_2)	41.02	31.84	5	135
	45.74	37.50	5	169
FUEL (x_3)	283.77	224.47	39	1020
	319.35	267.28	39	1073
All Observations (N=114)				
Explanatory Variables:				
OWNER	.72	.45	0	1.00
GROUP	.64	.48	0	1.00
LINELENGTH	10.89	4.56	4.50	35.56
STOPLENGTH	.56	.27	.19	1.74
SPEED	17.60	3.06	11.30	26.30
POPDENS	9.70	6.84	.69	48.06
VEHAGE	6.95	1.99	1.00	13.40
CTYPE	.55	.50	0	1.00
CTERM	8.12	4.45	1.00	30.00
SSUB	.49	.18	.24E-03	.83
TAX	.72	.30	.10	1.50

* Numbers on the first and the second line refer to Model I respectively Model II.

Source: Ministère de l'Equipement, du Logement, des Transport et de l'Espace (1991) *Annuaire Statistique sur les Transports Collectifs Urbains: Statistiques 1983-1990*, Paris, Direction des Transport Terrestres.

Table 2: Descriptive statistics of radial output efficiency measures calculated on a variety of reference technologies

	Mean	Standard Deviation	Skewness	Kurtosis	Minimum Value	Maximum Value	# Efficient Observations
Model I: Vehicle Kilometres							
All Observations (N=114)							
DEA-sd	.755	.141	.275	1.927	.519	1.000	11
DEA-wd	.780	.148	.168	1.753	.524	1.000	20
FDH	.977	.058	-3.332	14.830	.653	1.000	83
Observations without Outliers (N=71)							
DEA-sd	.886	.102	-.493	2.171	.626	1.000	18
DEA-wd	.911	.102	-.928	2.715	.626	1.000	29
FDH	.978	.059	-3.498	16.478	.653	1.000	51
Model II: Seat Kilometres							
All Observations (N=114)							
DEA-sd	.305	.283	1.262	3.312	.019	1.000	6
DEA-wd	.411	.338	.784	2.058	.019	1.000	19
FDH	.487	.373	.423	1.384	.050	1.000	27
Observations without Outliers (N=72)							
DEA-sd	.819	.163	-.954	4.626	.167	1.000	14
DEA-wd	.870	.166	-1.412	5.545	.167	1.000	33
FDH	.948	.112	-3.918	23.015	.246	1.000	45

Table 3: Correlation matrix between radial output efficiency measures calculated on a variety of reference technologies

	DEA-sd	DEA-wd	FDH
Model I: Vehicle kilometres			
All Observations (N=114)			
DEA-sd	1.000		
DEA-wd	.944	1.000	
FDH	.380	.387	1.000
Observations without Outliers (N=71)			
DEA-sd	1.000		
DEA-wd	.842	1.000	
FDH	.576	.497	1.000
Model II: Seat kilometres			
All Observations (N=114)			
DEA-sd	1.000		
DEA-wd	.808	1.000	
FDH	.651	.564	1.000
Observations without Outliers (N=72)			
DEA-sd	1.000		
DEA-wd	.861	1.000	
FDH	.696	.670	1.000

Table 4: Efficient observations affected by the assumptions underlying different reference technologies

Change in # of Efficient Observations	**Model I: Vehicle kilometres**	**Model II: Seat kilometres**
DEA-sd $\neq$ 1 $\wedge$ DEA-wd = 1	9	13
DEA-sd $\neq$ 1 $\wedge$ FDH = 1	72	21
DEA-wd $\neq$ 1 $\wedge$ FDH = 1	63	15
DEA-wd = 1 $\wedge$ FDH $\neq$ 1	0	7

Table 5: Tobit estimates for the output efficiency measures (N=33)

	Model I: Vehicle kilometres			Model II: Seat kilometres		
Dependent Variables	**DEA-sd**	**DEA-wd**	**FDH**	**DEA-sd**	**DEA-wd**	**FDH**
CONSTANT	.715	.722	.676	.496	.575	.600
	(.103)***	(.124)***	(.092)***	(.116)***	(.188)***	(.242)**
OWNER	.051	.077	.031	.142	.155	.143
	(.048)	(.057)	(.035)	(.053)***	(.082)*	(.109)
LINELENGTH	.009	.007	.013	.018	.020	.011
	(.005)*	(.006)	(.006)**	(.006)***	(.009)**	(.011)
STOPLENGTH	-.158	-.143	-.143	-.251	-.345	-.003
	(.086)*	(.102)	(.091)	(.101)**	(.155)**	(.222)
CTYPE	.073	.054	.110	.122	.189	.268
	(.040)*	(.047)	(.033)***	(.045)***	(.070)***	(.097)***
CTERM	.016	.014	.014	.010	.031	.032
	(.007)**	(.009)	(.007)*	(.007)	(.014)**	(.017)*
SSUB	-.219	-.167	.121	-.479	-.836	-.766
	(.127)*	(.152)	(.102)	(.142)***	(.226)***	(.310)**
TAX	.229	.231	.119	.477	.569	.516
	(.139)*	(.172)	(.092)	(.136)***	(.239)**	(.355)
Log-Likelihood	19.80	10.44	9.42	21.43	2.44	-3.29
Correctionfactor	.645	.535	.167	.882	.555	.284

Standard errors are between brackets

* Statistically significantly different from zero at the 90 % level.

** Statistically significantly different from zero at the 95 % level.

*** Statistically significantly different from zero at the 99 % level.

LIJST VAN RECENTE SESO-RAPPORTEN

VANHOUDT P., *Het belang van investeringen in mensen, materieel en know-how voor economische groei : empirische evidentie,* april 1995, 29 p. **(95/318)**

NONNEMAN W. en P. VANHOUDT, *A further augmentation of the Solow model and the empirics of economic growth for OECD countries,* April 1995, 12 p. **(95/319)**

MEERSMAN H. en D. CASSIMON, *Investment in the EC-countries : does reducing uncertainty matter ?,* April 1995, 35 p. **(95/320)**

VANDENBUSSCHE H., *Is European antidumping protection against Central Europe too high ?,* June 1995, 33 p. **(95/321)**

CORTENS I. en W. NONNEMAN, *Parental education and the demand for higher education : some evidence for Belgium,* June 1995, 22 p. **(95/322)**

SCHROYEN F., *Preference characterisation and indirect Allais coefficients,* June 1995, 40 blz. **(95/323)**

DOUVEN R. en J. PLASMANS, *S.L.I.M. - A small linear interdependent model of eight EU-member states, the USA and Japan,* June 1995, 35 blz. **(95/324)**

DOUVEN R. en J. PLASMANS, *Convergence and international policy coordination in the EU : a dynamic games approach,* June 1995, 39 blz. **(95/325)**

VAN REETH D., *How the Belgian treasure should auction its debt,* July 1995, 32 blz. **(95/326)**

DUCHESNE I. en W. NONNEMAN, *A neo-classical demand model for higher education in Belgium,* October 1995, 52 blz. **(95/327)**

DE BORGER B. en S. WOUTERS, *Transport externalities and optimal pricing and supply decisions in urban transportation : a simulation analysis for Belgium,* October 1995, 52 blz. **(95/328)**

VAN HOOYDONK C., *A TSP-procedure to test for the order of integration of a time series by means of (A)DF, PP and KPSS tests,* October 1995, 44 blz. **(95/329)**

VERKOOIJEN W., PLASMANS J. en H. DANIELS, *Long-run exchange rate determination : a neural network study,* November 1995, 44 blz. **(95/330)**

BRUGGEMAN A., *Disaggregating annual real GDP data into quarterly figures,* November 1995, 48 blz. **(95/331)**

VERHETSEL A., *Ruimtelijke patronen van de socio-economische ontwikkeling 1991-2010, Gewestplanscenario versus structuurplanscenario,* december 1995, 69 blz. **(95/332)**