Identification of Segments of French Urban Public Transport with a Latent Class Frontier Model

William Roy[a1] and Carlos Pestana Barros[b]

[a] Laboratoire d'Economie des Transports (LET), Université Lumière Lyon 2 - CNRS, Lyon, France.
[b] Instituto Superior de Economia e Gestão (ISEG), Technical University of Lisbon, Portugal.

Abstract: This paper analyses technical efficiency of French urban public transport from 1995 to 2002 with unbalanced panel data. The latent class frontier model is used allowing the identification of different segments in the production frontier. We find that there are three statistically significant segments in the sample. Therefore, we conclude that no common transport policy can reach all of the transportation companies analysed, thereby requiring transport policies by segments.

Key words: Urban public transport, stochastic production frontier, latent class model, technical efficiency, panel data.

JEL-code: C24, C81, D24, L92

---

[1] Correspondent author: william.roy@let.ish-lyon.cnrs.fr
Postal address: LET-ISH \\ 14 av. Berthelot \\ 69363 LYON Cedex 07 \\ France
Tel: +33 4 72 72 64 44
Fax: +33 4 72 72 64 48

1. Introduction

Efficiency in urban transport is a theme that has attracted research in the past (Quinet and Vickerman 2004). A review of the literature shows two main approaches, the DEA-Data envelopment Analysis, Chu et al. (1992), Chang and Kao (1992), Karlaftis and McCarthy (1997) among others and the stochastic frontier models, Jorgenson, Pedersen and Solvoll (1995) and Jorgensen, Pedersen and Volden (1997) among others.

The aim of this research goes beyond previous research by adopting the latent frontier model in line with Orea and Kumbhakar (2004). The motivation for the present research is to combine operational and financial variables in the evaluation of the transport segments efficiency. Operational efficiency is easily observed on the field; however, financial efficiency (reported in the company accounts) has no such transparency. The separation of these two realities - operational and finance - makes the merging of them a technical problem that needs the support of econometric techniques, such as frontier models. The motivation for the present research is based on the fact that companies operating in the French urban transport can be identified as belonging to different segments: the private, the public and the mixed. The aim of this paper is to investigate the statistical factors that characterize who belong to such segments.

The contribution of this paper within transport research is twofold. First, it adopts an innovative stochastic frontier model, which endogenously identifies segments of transport companies in a sample, overcoming the hypothesis of homogeneity of the variables adopted by the frontier models previously mentioned. Moreover, it applies the model to the French urban transport companies. This industry is strongly regulated (GART 2002), and smack in the middle of its regulation reform (European Commission

2005). So this paper use an innovative model that may improve our understanding of the best way to regulate the industry.

The paper is organized as follows: in second section the contextual setting is presented. The third and forth sections present a survey of the literature and the model respectively. Section five presents the data and the results. Finally, section six discusses the results while section seven provides some conclusions.

2. Contextual Setting

The French urban transport is decentralised to local authorities (Kerstens 1996, 1999). These urban authorities choose to provide transport services by its own operator ("régie"), or alternatively delegates the operation to a private company[1] or to a semi-public company ("société d'économie mixte"). The not 'in-house' operators are legally selected through tendering processes, but competition is not very pregnant when a semi-public company wants to succeed to itself (Roy and Yvrande-Billion 2007). But whenever an organising authority delegates the operation, it signs a contract with the operator. Contracts can be defined according to alternative risk sharing rules (Gagnepain 1998, Gagnepain and Ivaldi 2002; Gagnepain 1998):

- the net cost contract ("CFF: Compensation Financière Forfaitaire") which provides incentives on receipts and costs
- the gross cost contract ("GPF: Gestion à Prix Forfaitaire") which allocates risks on costs to the operator and risks on receipts to the public authority

[1] The private companies are almost three international groups: Kéolis, Connex (Véolia) and Transdev.

- the management contract ("Gérance"), according to which all risks are borne by the public authority.

INSERT Figure 1

Data are collected under the responsibility of the CERTU, a ministerial agency (CERTU 2003), and controlled by the GART, a nation-wide association that gathers most of the local authorities in charge of an urban transport network. The unbalanced panel is composed of 135 French urban transport units, comprising all forms characterised above.

3. A survey of the literature

There has been relatively extensive research into the bus industry using a variety of methods. For an extensive and comprehensive survey, see De Borger, Kerstens and Costa (2002); for a meta-analysis, see Brons, Nijkamp, Pels and Rietveld (2005). Restricting our survey to those studies that have used frontier models, we observe the predominant use of the non-parametric DEA model, whilst the parametric econometric frontier is applied more scantily. Table 1 below comprises a list of some frontier-model papers, namely recently published papers and those not surveyed in De Borger, Kerstens and Costa (2002).

INSERT Table 1

What emerges from this research is that there are substantial technical inefficiencies among urban transport operators (De Borger et al., 2002), calling for the subsidy allocation to be based on a yardstick type of contract (Dalen and Gómez Lobo, 2003).

4 Latent Class Frontier Models

In this paper, we adopt the stochastic cost econometric frontier approach (kumbhakar and Lovell 2000). This approach, first proposed by Farrell (1957), came to prominence in the late 1970s as a result of the work of Aigner, Lovell and Schmidt (1977), Battese and Corra (1977) and Meeusen and Van den Broeck (1977).

The frontier is estimated econometrically and measures the difference between the inefficient units and the frontier through the residuals. This is an intuitive approach based on traditional econometrics. If we assume that the residuals have two components (noise and inefficiency), the stochastic frontier model emerges. Consequently, the main issue here is the decomposition of the error terms. Let us present the model more formally. The general frontier cost function proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977) is the following:

$$C_{it} = C(X_{it}).e^{v_{it}+u_{it}} \;;\; \mathrm{i}=1,2,\ldots \mathrm{N}, \mathrm{t}=1,2,\ldots T \qquad (1)$$

Where C_{it} represents a scalar cost of the decision-unit i under analysis in the t-th period; X_{it} is a vector of variables including the input prices and the output descriptors present in the cost function; and $\varepsilon_{it} = u_{it} + v_{it}$ is the error term. This term may be decomposed into two components:

(i) the error term v_{it} is the one that is traditional of the econometric models, assumed to be independently and identically distributed, that represents the effect of random shocks (noise $v_{it} \sim N(0, \sigma_v^2)$) and is independent of u_{it}

(ii) the inefficient term u_{it} represents the technical inefficiencies and is assumed to be positive and normally distributed with a zero mean and variance σ_u^2. The positive disturbance u_{it} is reflected in a half-normal independent distribution truncated at zero, signifying that each bus company's cost must lie on or above its cost frontier. This implies that any deviation from the frontier is caused by management factors controlled by the bus companies.

The total variance is defined as $\sigma^2 = \sigma_v^2 + \sigma_u^2$. The contribution of the error term to the total variation is as follows: $\sigma_v^2 = \sigma^2 /(1+\lambda^2)$. The contribution of the inefficient term is: $\sigma_u^2 = \sigma^2.\lambda^2 /(1+\lambda^2)$. Where λ is defined as $\lambda = \dfrac{\sigma_u}{\sigma_v}$, providing an indication of the relative contribution of u and v to $\varepsilon = u + v$.

Given that the estimation procedures of equation (1) yield merely the residual ε, rather than the inefficiency term u, this term in the model must be calculated indirectly (Greene, 2000). In the case of panel data, such as that used in this paper, Battese and Coelli (1988) use the conditional expectation of u_{it}, conditioned on the realized value of the error term $\varepsilon_{it} = (v_{it} + u_{it})$, as an estimator of u_{it}. In other words, $E[u_{it} / \varepsilon_{it}]$ is the mean productive inefficiency for the i-th bus company at any time "t". Following Orea and Kumbhakar (2004), we can write equation (1) as a latent class frontier model:

$$C_{it}\big|_j = C(X_{it})\big|_j .e^{v_{it}|_j + u_{it}|_j} \;;\quad \mathrm{i} = 1,2,\ldots \mathrm{N},\ \mathrm{t} = 1,2,\ldots T \qquad (2)$$

Where subscript "*i*" denotes the firm, "*t*" indicates time and "*j*" represents the different classes or groups. The vertical bar signifies that there is a different model for each class "*j*" and, therefore, each bus company belongs to the same group in all the periods. Assuming that *v* is normally distributed and *u* follows a half-normal distribution, the likelihood function (*LF*) for each bus company "*i*" at time "*t*" for group "*j*" is (Cf. Greene, 2004):

$$LF_{ij} = C(x_{it}, \beta_j, \sigma_j, \lambda_j) = \frac{\Phi(\lambda_j . \varepsilon_{it}|_j / \sigma_j)}{\Phi(0)} . \frac{1}{\sigma_j} . \phi\left(\frac{\varepsilon_{it}|_j}{\sigma_j}\right) \qquad (3)$$

Where $\varepsilon_{it}|_j = \ln C_{it}|_j - \beta_j^{'} x_{it}$, $\sigma_j = \left[\sigma_{uj}^2 + \sigma_{vj}^2\right]^{1/2}$, $\lambda_j = \sigma_{uj} / \sigma_{vj}$, while ϕ denote the standard normal density and Φ the cumulative distribution function. The likelihood function for bus company "*i*" in group "*j*" is obtained as the product of the likelihood functions in each period.

$$LF_{ij} = \prod_{t=1}^{T_i} LF_{ijt} \qquad (4)$$

The likelihood function for each bus company is obtained as a weighted average of its likelihood function for each group *j*, using as weights the prior probabilities of class *j* membership.

$$LF_i = \sum_{j=1}^{J} P_{ij} LF_{ij} \qquad (5)$$

The previous probabilities must be in the unit interval: $0 \leq P_{ij} \leq 1$. Furthermore, the sum of these probabilities for each group must be one: $\sum_j P_{ij} = 1$. In order to satisfy these two conditions we parameterized these probabilities as a multinomial logit. That is:

$$P_{ij} = \frac{\exp(\delta_j q_i)}{\sum_{j=1}^{J} \exp(\delta_j q_i)} \qquad (6)$$

where the q_i is a vector of variables which are used to split the sample, and δ_j is the vector of parameters to be estimated. One group is chosen as the reference in the multinomial logit. The overall log-likelihood function is obtained as the sum of the individual log-likelihood functions:

$$\ln LF = \sum_{i=1}^{N} \ln LF_i = \sum_{i=1}^{N} \ln\left(\sum_{j=1}^{J} P_{ij} \prod_{t=1}^{T} LF_{ijt} \right) \qquad (7)$$

The log-likelihood function can be maximised with respect to the parameter set $\theta_j = (\beta_j, \sigma_j, \lambda_j, \delta_j)$ using conventional methods (Greene, 2004). Furthermore, the estimated parameters can be used to estimate the posterior probabilities of class membership using Bayes Theorem:

$$P(j|i) = \frac{P_{ij} LF_{ij}}{\sum_{j=1}^{J} P_{ij} LF_{ij}} \qquad (8)$$

One important issue in these models is the way in which the number of classes has to be determined. The usual procedure entails estimating several models with different numbers of groups and then applying statistical tests to choose the preferred model. Greene (2005) proposes a procedure in which, beginning from a J^* known to be at least as large as the true J, one can test down (given that the $J-1$ class model is nested with the J class model imposing $\theta_j = \theta_{j-1}$, and based on the likelihood ratio tests). An alternative method is to use Information Criteria such as the Akaike Information

Criteria (AIC) or the Schwarz Bayesian Information criteria. Moutinho, Machado and Silva (2003) avail of these in a latent class framework using the following expressions:

$$SBIC = -2.\ln LF(J) + \ln(n).m \qquad (9)$$

$$AIC = -2.\ln LF(J) + 2m \qquad (10)$$

where *LF(J)* is the value that the likelihood function takes for *J* groups, *m* is the number of parameters used in the model and *n* is the number of observations ($n = N.T$, where *N* is the number of bus companies and *T* denotes the number of seasons). The preferred model is that for which the value of the statistic is the lowest. Based on the panel data, presents the maximum likelihood estimators of model (1) as found in other authors' recent studies Greene (2000, 2004, 2005).

5. Hypothesis

Consider a French bus company operating in a city. Based in previous research, Roy and Yvrande-Billion (2007), Gagnepain and Ivaldi (2002), the frontier model allows the definition of the following null hypotheses:

Hypothesis 1 (public vs. private): Public companies are less efficient than private companies. This hypothesis is based on the traditional hypothesis related to private versus public property in markets (Williamson 1979). This is a traditional hypothesis in transportation markets (Davis 1971, Chang and Kao 1992, Roy and Yvrande-Billon 2007). Reasons to support this hypothesis is that with private ownership the rewards and costs of the activity are more directly concentrated in the stockholders restricting the principal-agent relationship (Jensen and Meckling 1976). However, there are some research that found no evidence that private operators are more efficient than public operators (Caves and Christensen 1980) justifying the present hypothesis.

Hypothesis 2 (Private companies managed by cost plus contract): Private companies with cost plus contract perform efficiently. This hypothesis is based in the theory of transaction costs and property rights, Klein, Crawford and Alchian (1978), Williamson (1981, 1985) and Grossman and Hart (1986). This theory is based in two critical assumptions: First, the firms cannot write complete contracts concerning their transport allocation based in the cost-plus rule. Second, investments are specific to firms' assets so that the same investment is less valuable with different assets. When both assumptions hold, the theory predicts that firms under-invest because they are afraid that their relationship with the other firm may end at same point. To minimise under-investment, firms allocate dedicated asset specificity (Williamson, 1981), which refers to investment which take place with the prospects of selling a significant amount of product to a particular customer. Therefore, assuming the asset specificity strategy, private companies managed by cost plus contract that are assumed to be efficient.

Hypothesis 3 (Private companies with Net cost Contract): Private companies with net cost contract perform efficiently searching for profits. This hypothesis is based on previous research on transportation and on the strategic-group theory (Caves and Porter, 1977) which justifies differences in efficiency scores as being due to differences in the structural characteristics of units within an industry. In the case of bus companies, units with similar asset configurations pursue similar strategies with similar results in terms of performance (Porter, 1979). Although there are different strategic options to be found among the different sectors of an industry, because of mobility impediments, not all options are available to each bus companies, causing a spread in the efficiency scores of the industry. Therefore it is assumed that French private bus companies with net cost contract, adopt this type of contract because it corresponds to a strategy inherent to an efficient drive.

Hypothesis 4 (private companies with gross cost contract) Private companies with gross cost contract perform efficiently searching for profits. This hypothesis is based on previous research the resource-based theory (Barney, 1991; Rumelt, 1991; Wernerfelt, 1984), which justifies different efficiency scores on the grounds of heterogeneity in relation to the resources and capabilities on which the bus companies base their strategies. These resources and capabilities may not be perfectly mobile across the industry, resulting in a competitive advantage for the best-performing bus companies.

Purchasable assets cannot be considered to represent sources of sustainable profits. Indeed, critical resources are not available in the market. Rather, they are built up and accumulated on the bus's premises, their non-imitability and non-substitutability being dependent on the specific traits of their accumulation process. The difference in resources thus results in barriers to imitation (Rumelt, 1991) and in the bus managers' inability to alter their accumulated stock of resources over time. In this context, unique assets are seen as exhibiting inherently differentiated levels of efficiency; sustainable profits are ultimately a return on the unique assets owned and controlled by the bus companies (Teece et al., 1997).

These hypotheses will be tested with the random frontier model.

6. Data and Results

To estimate the production frontier, we used panel data on French urban transport companies for the years 1995 to 2002 (8 years, 135 units resulting in an unbalanced panel data of 981 observations). Frontier models require the identification of inputs (resources) and outputs (transformation of resources). Several criteria can be used in their selection. One empirical criterion is availability. For the applicability of the model's results and its management, it is important to "buy in" to the process that the measures of inputs and outputs are relevant and adequately measurable, and that appropriate archival data are available. Usually this latter criterion is used, since it encompasses the other two already mentioned criteria. Secondly, the literature survey is a way of ensuring the validity of the research and therefore constitutes another criterion that needs to be taken into account. The final criterion for measurement selection is the professional opinion of managers. In this paper, we follow the first two criteria.

Table 2 presents the characteristics of the variables used in the analysis. We transformed the variables according to the description column. We adopted the traditional log-log specification (Translog) to allow for the possible non-linearity of the frontier.

INSERT Table 2

6.1 Results

In this study, we estimate a stochastic translog production[2] function with output descriptors (*Y*), input descriptors (*X*), additional business conditions (*Z*) and a trend (*T*).

$$\ln(VehicleKm_{it}) = \alpha_0 + \alpha_t T + \alpha_{tt} T^2 + \sum_j \alpha_j \ln X_{j,it} + \sum_k \alpha_k \ln Y_{k,it} + \frac{1}{2}\left[\sum_j \sum_k \gamma_{jk} \ln X_{j,it} \ln X_{k,it} + \sum_k \sum_l \gamma_{kl} \ln Y_{k,it} \ln Y_{l,it}\right] + \sum_j \sum_k \gamma_{jk} \ln X_{j,it} \ln Y_{k,it} + \sum_l \alpha_l \ln Z_l + (v_{it} - u_{it}) \qquad (11)$$

The sole output descriptor is (journey). The input descriptors are (labour1, labour2, energy, vehicles, network length), business conditions (population, public, spublic, private, Managed, Net, Gross).

This is the production frontier model, known as the error components model in Coelli, Rao and Battese (1998). The variables have been defined and characterized in Table 2. Table 3 presents the results obtained for the stochastic production frontier, using a Gauss program and using a half-normal distribution specification. For comparative purposes a non random frontier model alongside a traditional production function are estimated.

INSERT Table 3

[2] Compared to the standard cost function model presented previously, a negative sign precede *u*

The results of the latent frontier model are presented, which reports the parameter estimation and their significance for all the variables. The log-likelihood value of the estimated latent mixed logit model is 1252.132. The overall fit of the model is reasonably good with Chi-square statistic square value of 205.123 with 10 degrees of freedom and level of significance of 0.00052.

What do these results mean for the hypotheses proposed? Primarily, one can conclude that there are three segments in the data sample, which are statistically significant. The first one is the more representative, since the probability of a company to be allocated to this segment of 0.527 whereas the second segment has a probability of 0.317 and the third 0.156.

What are the characteristics of these segments? This first segment is composed by companies characterised by a positive relationship between the inputs and outputs contributing to production, Varian (1987), but with a negative relationship with Trend square, Managed, Net and Gross. This segment can be characterised as generic bus companies, Roy and Yvrande-Billion (2007). The second segment is composed of companies with a positive relationship with inputs and outputs contributing to production and a negative relationship with trend-square, managed, Net and Gross. This segment can be characterised as public enterprises model, based in the significant result of the public dummy and the insignificant parameters related with private bus companies and its attributes (public, managed, Net and Gross). The third segment has also a positive relationship with inputs and outputs but a negative relationship with square trend, managed, Net and Gross. These results combined with the statistical insignificant sign for public allow the definition of the segment as private enterprises.

Finally, the production increases at decreasing rate according the trend square variable, signifying that technological improvement increase at decreasing rate.

6.2 Discussion

The interpretation of the previous results is as follows. First, they allow us to conclude that latent frontier models describe the French urban public transport well, when allowing for heterogeneity and defining segments in the sample. This is possibly the main result of the present paper. The implication of this result is that a common transport policy would be inappropriate for being applied to all transport companies, since the three segments identified advocate for the existence of heterogeneity. Therefore any transport policy which involves some of these heterogeneous variables has to be tailored by segments. Second, our analysis has led us to identify three segments. Does this mean that there are only two segments in the French public urban transport companies? No, it signifies that given the small number of observations available, the model cannot differentiate more segments. The third finding is that, for companies in the first segment, are defined as generic French bus companies, a picture that emerges with traditional homogenous frontier models (Roy and Yvrande-Billion 2007). The results for the first segment are intuitive and it is in accordance with the economic theory (Cf. Varian, 1987).

In the second segment, the companies are identified as public companies. How can the signs of the variables, for the second segment, be explained? All the statistical significant parameters have signs equal to the first segment. At European level public companies are recognised as tending to contract more labour than need and therefore this result supports previous research in the field, Davis (1971), Chang and Kao (1992), Roy and Yvrande-Billon (2007). The third segment is defined as private enterprises.

What is the overall rationality of these results? They are in fact quite intuitive since bus companies are not homogenous. Notice, for instance that there are small and large bus companies;. These visible characteristics result in different performance levels, dissimilar transportation flux and discrepancy in their financial debts, which provokes different segments in the sample. In the case of the French bus companies there are two different segments that can be distinguished on the grounds of the values and signs of the exogenous variables: Generic, private and public.

What is the implication of the results for the hypotheses? First, we do not reject the first hypothesis since there is evidence that private enterprises are more efficient than public enterprises. This result is based in the negative sign of the attributes of private managed companies (Manage, Net and Gross) when compared with the sign of the coefficient of public companies. This result is also supported in previous research in this field, Roy and Yvrande-Billon (2007), Gagnepain and Ivaldi (2002) with alternative models. This result serves as a cross-validation of previous results in this field.

Second, we do not reject hypothesis two, based in the sign of managed in all segments. Therefore we conclude private companies with cost plus contract perform efficiently searching for profits, Williamson (1979, 1985).

Third, we do not reject hypothesis three, based in the sign of Net variable in all segments. Therefore private companies with net cost contract perform efficiently searching for profits (Caves and Porter, 1977).

Finally, we do not reject hypothesis four, based in the sign of Gross variable in all segments. Therefore private companies with net cost contract perform efficiently searching for profits (Barney, 1991; Rumelt, 1991)

Therefore, it can be concluded that, in this context, private bus companies are more efficient, than public bus companies. This is supported in the theory of transaction costs and property rights, Klein, Crawford and Alchian (1978), Williamson (1979, 1985), signifying that French bus companies have a clear relationship with the regulator, with the firms allocating dedicated asset specificity, Williamson (1985) according to the population density and network length. This means that private French bus companies search for profits in a regulated environment relatively clear. This results in private companies operating either under management cost-plus contract, or net cost contract or even gross costs contract being efficient. What are the most efficient private forms of management? The model does not clearly separates the different forms of management, but the value and significance of the parameters can serve as a clue that cross-validates previous research in the field, Roy and Yvrande-Billion (2007).

Therefore, assuming the asset specificity strategy, private companies managed by whichever contract are assumed to be efficient, displaying unique specific management assets are seen as exhibiting inherently differentiated levels of efficiency and sustainable production are ultimately a return on the unique assets owned and controlled by the transport companies (Teece et al., 1997). The property resource is the main factor segmenting the sample. In addition, the strategic-groups theory (Caves and Porter, 1977), which justifies different efficiency scores on the grounds of differences in the structural characteristics of units within an industry, could explain part of the efficiency differences observed in the French public urban transport. All these dynamics are derived from a clear allocative contract that minimises transaction costs and property rights, Klein, Crawford and Alchian (1978), Williamson (1979, 1985)

7. Conclusion

This article has proposed a simple framework for the comparative evaluation French urban public transport companies and the rationalisation of their operational activities. The analysis is carried out through implementation of the latent frontier model, which allows for the incorporation of multiple inputs and outputs in determining the relative efficiencies and the inclusion of heterogeneity observed in the data. Several managerial insights and interesting implications of the results have been discussed. The main conclusion is that, on average, the latent frontier model captures the dynamics observed in the data better than the precedent techniques applied in this context. In particular, our results suggest that labour1, energy, vehicles, network length and population are major homogenous among different factor. Property status and labour 2 differentiates the segments. Nevertheless, more investigation is needed to confirm the results of the present research.

The main limitation of this paper is related to the data set, since the data span is relatively short, restricting the estimation of latent classes to only three classes. Having access to a larger data span would allow more latent variables to emerge. In any case, the scope of the present paper is satisfactorily achieved, given the exploratory character of it. In specific, the intention of this study is to draw the attention of the public transport companies towards identifying segments among them, and defining business strategies for each transport segment in order to adapt strategies to their characteristics. To achieve more conclusive policy prescriptions, a larger data set would be required. Precisely, the limitations of the present paper suggest directions for new research. Firstly, additional research is needed to confirm the results of this paper, as well as to clarify the issues described above. Secondly, research taking into account the presence of heterogeneity must be expanded to consider transport companies in other countries.

References

Aigner, D.J., Lovell, C.A.K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6, 21-37.

Barney, J., 1991. Firm resources and sustained competitive advantage. *Journal of Management* 17, 99-120.

Battese, G.E., Corra, G.S., 1977. Estimation of a Production Frontier Model: With application to the pastoral zone of Eastern Australia. *Australian Journal of Agricultural Economics* 21, 169-179.

Battese, G.E., Coelli, T.J., 1988. Prediction of Firm-level technical efficiencies with a generalised frontier production function and panel data. *Journal of Econometrics* 38, 387-399.

Boame, A. K., 2004. The Technical efficiency of Canadian Urban Transit Systems. *Transportation Research Part E* 40, 401-416.

Brons, M., Nijkamp, P., Pels, E., Rietveld, P., 2005. Efficiency of urban public transit: A meta analysis. *Transportation* 32, 1-21.

Caves R., Porter M.E., 1977. From entry barriers to mobility barriers: Conjectural decisions and contrived deterrence to new competition. *Quarterly Journal of Economics* 91, 241-261.

Caves, D.W., Christensen, L.R., 1980. The Relative Efficiency of Public and Private Firms in a Competitive Environment: The Case of Canadian Railroads. *The Journal of Political Economy* 88(5), 958-976.

CERTU (Ed.), 2003. *Transport Collectifs Urbains- Evolution 1997-2002*. Lyon.

Chang, K.P., Kao, P.H., 1992. The relative efficiency of public versus private municipal bus firms: An application of data envelopment analysis. *Journal of Productivity Analysis* 3, 67-84.

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

Coelli, T.J., Rao, P., Battese, G.E., 1998. *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic Press.

Cowie, J., Asenova, D., 1999. Organisational form, scale effects and efficiency in the British Bus industry. *Transportation* 26(3), 231-248

Dalen, D.M., Gomez Lobo, A., 2003. Yardsticks on the road: regulatory contracts and cost efficiency in the Norwegian bus industry. *Transportation* 30, 371-386.

Davis, D., 1971. The Efficiency of public versus private firms: the case of Australia two airlines. *Journal of Law and Economics* 14(1), 149-165.

De Borger, B., Kerstens, K., Costa, A., 2002. Public Transit Performances: what does one learn from frontier studies?. *Transport Reviews* 22(1), 1-38.

European Commission (Ed.), 2005. *Proposal for a Regulation of the European Parliament and of the Council on public passenger transport services by rail and by road*, COM(2005) 319 final, Official Journal of the European Communities, Brussels.

Farrell, M.J., 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society* Series A 120(3), 253-290.

Gagnepain, P., 1998. Structures productives de l'industrie du transport urbain et effets des schémas réglementaires. *Economie et Prevision* 135, 95-107.

Gagnepain, P., Ivaldi, M., 2002. Incentive Regulation Policies: The case of Public Transit Systems in France. *Rand Journal of Economics* 33(4), 605-629.

GART (Ed.), 2002. *Les Chiffres de réference des transports publics*. Paris.

Greene, W., 2000. *Econometric Analysis (4th Ed.)*, London, Prentice Hall

Greene, W., 2005. Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis* 23, pp. 7-32.

Greene, W., 2004. Distinguishing between heterogeneity and efficiency: stochastic frontier analysis of the World Health Organisation's panel data on national health care systems. *Health Economics* 13, 959-980.

Grossman, S.J., Hart, O.D., 1986. The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *Journal of Political Economy* 94(4), 691-719.

Holvad, T., Hougaard, J.L., Kronborg, D., Kvist, H.K., 2004. Measuring inefficiency in the Norwegian bus industry using multi-directional efficiency analysis. *Transportation* 31(3), 349-369.

Jensen, M.C., Meckling, W., 1976. Theory of the firm: managerial behaviour, agency costs and capital structure. *Journal of Financial Economics* 3, 305-60.

Jorgensen, F., Pedersen, P., Solvoll, G., 1995. The cost of bus operators in Norway. *Journal of Transport Economics and Policy* 29, 253-262.

Jorgensen, F., Pedersen, P.A., Volden, R., 1997. "Estimating the inefficiency of the Norwegian bus industry from stochastic cost frontier models", *Transportation* 24, 421-433.

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

Karlaftis, M.G., McCarthy, P.S., 1997. Subsidy and Public Transit Performance: A factor analytic approach. *Transportation* 24, 253-270.

Kerstens, K., 1996. Technical efficiency measurement and explanation of French urban transit companies. *Transportation Research Part* A 30, 431-452.

Kerstens, K., 1999. Decomposing technical efficiency and effectiveness of French Urban Transport. *Annales d'économie et de Statistique* 54, 129-1555.

Khumbhakar, S.C., Lovell, C.A.K., 2000. *Stochastic Frontier Analysis*. Cambridge University Press, N.Y.

Klein, B.; Crawford, R.C., Alchian, A., 1978. Vertical Integration, Appropriable Rents, and the Competitive Contracting Process. *Journal of Law and Economics* 22, 297-326.

Meeusen, W., van den Broeck, J., 1977. Efficiency estimation from a Cobb-Douglas production function with composed error. *International Economic Review* 18, 435-444.

Moutinho, N. Machado, F., Silva, E., 2003. An International Comparison of Productivity Change in Agriculture and the economy as a whole. *Discussion paper 9-2003*. Faculdade de Economia, University of Porto.

Odeck, J., 2003. Ownership, Scale Effects and Efficiency of Norwegian Bus Operators: Empirical Evidence. *International Journal of Transport Economics* 30(3), 305-325.

Odeck, J., Alkadi, A., 2001. Evaluating efficiency in the Norwegian bus industry using data envelopment analysis. *Transportation* 28, 211-232

Orea, L., Kumbhakar, S., 2004. Efficiency measurement using a latent class stochastic frontier model. *Empirical Economics* 29, 169-183.

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

Porter, M.E., 1979. The structure within industries and companies' performance. *The Review of Economics and Statistics* 61, 214-227.

Quinet, E., Vickerman, R., 2004. *Principles of Transport Economics and Policy*. Edward Elgar, Chaltenham.

Roy, W., Yvrande Billion, A., 2007. Ownership, Contractual Practices and Technical Efficiency: The Case of Urban Public Transport in France. *Journal of Transport Economics and Policy* (forthcoming).

Rumelt, R., 1991. How much does industry matter?. *Strategic Management Journal* 12(2), 167-185.

Teece, D., Pisano, G., Shuen, A., 1997. Dynamic capabilities and strategic management. *Strategic Management Journal* 18(7), 509-533.

Varian, H.R., 1987. *Intermediate Microeconomics: A Modern Approach*. New York, W.W. Norton & Co.

Viton, P.A., 1996. Technical Efficiency in Multi-Mode Bus Transit: A Production Frontier Analysis. *Transportation Research Part B* 31(1), 23-39.

Viton. P.A., 1998. Changes in Multi-mode bus transit efficiency, 1998-1992. *Transportation* 25, 1-21.

Wernerfelt, B., 1984. A resource-based view of the firm. *Strategic Management Journal* 5(2), 171-180.

Williamson, O., 1979. Transaction Cost Economics: The Governance of Contractual Relations. *Journal of Law and Economics* 22, 232-262.

Williamson, O., 1981. The economics of organization: The transaction costs approach. *American Journal of Sociology* 87, 548-577.

Williamson, O., 1985. *The Economic Institutions of Capitalism: Firms, Markets and Vertical Contracting*. Free Press, N.Y.

Figure 1: governance structures

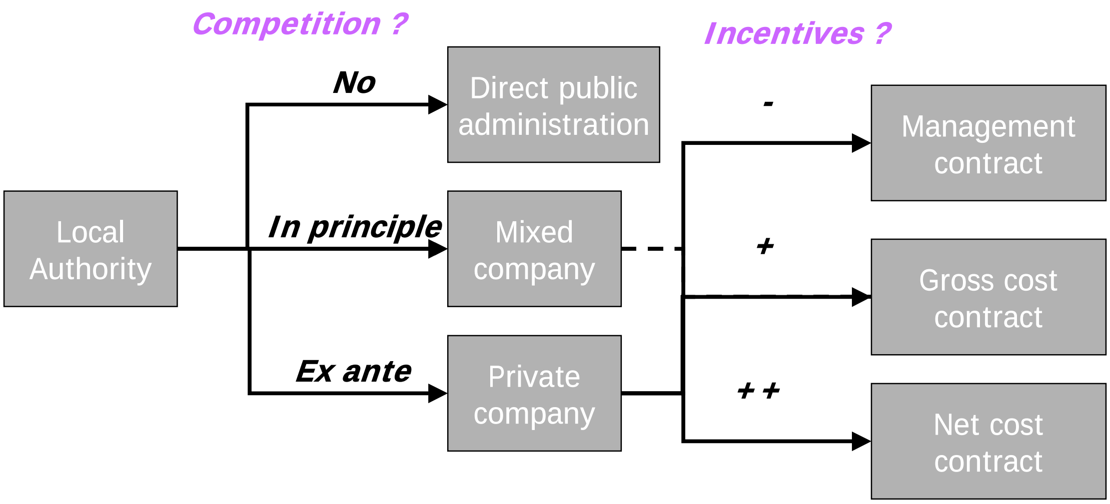

Table 1: Research on bus companies efficiency

Papers	Method	Units	Inputs	Outputs
Karlaftis (2004)	DEA-CCR model and a goal programming model	259 USA transit systems, 1990 to 1994.	(1) Number of employees, (2) fuel used, (3) total number of vehicles operated.	(1) Vehicles-miles, (2) passenger-miles.
Boame (2004)	Two step procedure. First step a DEA-CCR model. Second step, a bootstrap Tobit model	30 Canadian bus companies, 1990 to 1998.	(1) fleet size, (2) liters of fuel, (3) total number of employees hours. Tobit model variables: (1) average speed, (2) peak/base ratio, (3) average fleet age, (4) time trend	(1) revenue vehicle kilometers.
Holvad et al. (2004)	MEA-Multi directional efficiency analysis	17 Norwegian bus companies, 1991	(1) Fuel costs, (2) driver costs, (3) other costs	(1) Seat kilometers
Odeck (2003)	DEA-CCR and DEA-BCC models and Mann-Whitney test.	33 Norwegian bus companies, 1994	(1) Fuel consumption in liters, (2) material, (3) effective driving hours, (4) total number of staff, (5) seat kilometers	(1) Passenger kilometers.
Odeck and Alkadi (2001)	DEA-CCR and DEA-BCC models and Mann-Whitney test.	47 Norwegian bus companies, 1994	(1) effective driving hours, (2) total number of staff; (3) fuel consumption; (4) equipment	Seat kilometers, (2) passenger kilometers, (3) total number of seats.
Pina and Torres (2001)	Two step approach. First step, a DEA-CCR model; second step, a regression analysis and logit model.	Catalonian bus companies	(1) ratio fuel/100 kilometers; (2) ratio cost/km, (3) ratio subsidy/traveller. Regression variables: (1) ratio cost/travellers, (2) ratio fuel/100 kms; (3) ratio km/bus. Logit variables: (1) ratio Km/employees, (2) ratio Km/bus, (3) ratio km/inhabitants, (4) inhabitants, (5) index of accidents, (6) frequency of accidents, (7) ratio fuel/100 km, (8) ratio subsidies/traveler, (9) ratio cost/km, (10) ratio cost/traveler.	(1)ratio bus-km/employees; (2) ratio bus-km (year)/bus; (3) ratio bus-km/inhabitant; (4) accident rate; (5) frequency of accidents; (6) population
Cowie and Asenova (1999)	DEA-CCR and DEA-BCC models and Mann-Whitney test.	British bus companies, 1995/96.	(1) total staff employed, (2) buses with over 35 seats (3) buses with under 35 seats.	Operating revenue,
Viton (1998)	DEA-CCR, DEA-BCC, DEA-Malmquist	USA bus companies. 183 companies in 1988 and 169 companies in 1992	(1) average fleet age; (2) number of directional miles; (3) fleet size; (4) gallons of fuel by mode; (5) persons-hour of transportation; (6) persons-hour of maintenance; (7) persons hour of administrative; (8) persons-hour of capital; (9) costs with tires and material; (10) service cost, (11) utilities costs; (12) insurance cost.	(1) vehicles-miles; (2) passengers trips; (3) vehicle-hours.

Viton (1996)	DEA-CCR, DEA-BCC.	217 USA bus companies, 1990	(1) average speed, (2) average fleet age, (3) number of directional miles, (4) fleet size, (5) gallons of fuel by mode; (6) persons-hour of transportation; (7) persons-hour of maintenance; (8) persons hour of administrative; (9) persons-hour of capital, (10) costs with tires and material, (11) service cost, (12) utilities costs; (13) insurance cost.	(1) vehicles-miles; (2) passengers trips;
Jorgensen et al. (1997)	Cobb-Douglas stochastic frontier approach, Technical Efficient Effects Model with Z variables	170 Norwegian Bus companies, 1991	(1) total cost per vehicle-kms	(1) vehicle-kms produced, (2) average bus size measured by the sum of seats and standing places, (3) passengers per vehicle-km, (4) a dummy for the case where the bus operator is also engaged in sea transport, (5) a dummy for companies operating the coastal area.
Kerstens (1996)	Two step procedure. First step DEA model and FDH model. Second step, Tobit model	114 French in 1990 transport companies	(1) number of vehicles in use; (2) employees; (39 total fuel consumption; Tobit model variables: (1) a dummy for ownership, (2) a dummy for group; (3) mean line length, (4) average distance between stops, (5) speed, (6) population density, (7) average age of the bus vehicles, (8) agreement type, (9) duration of the agreement contract.	(1) number of vehicle kilometers; (2) number of seat kilometers
Jorgensen et al. (1995)	Cobb-Douglas stochastic frontier approach	170 Norwegian Bus companies, 1991	(1) total cost per vehicle-kms	(1) vehicle-kms produced, (2) average bus size measured by the sum of seats and standing places, (3) passengers per vehicle-km, (4) a dummy for the case where the bus operator is also engaged in sea transport, (5) a dummy for companies operating the coastal area.

Table 2: Descriptive Statistics of the data 1995-2002

Variable	Description	Min	Max	Mean	Std dev
ln (VehicleKm)	The number of kilometers cover by vehicles	5.251	7.056	6.169	0.445
ln (Labour1)	The number of equivalent employees-drivers	1.945	6.537	4.118	1.063
ln (Labour2)	The number of equivalent employees non-drivers	-0.693	4.036	2.997	1.562
ln (Energy)	The quantity of diesel m^3 consumed	1.799	6.779	4.777	1.659
ln (Vehicles)	The number of vehicles	1.609	4.127	3.819	2.103
ln (Network length)	The total lines length	2.639	5.319	4.790	0.805
ln (Population)	The population in the area covered by the network	4.141	12.848	11.261	0.729
ln (Journeys)	The number of journey[3]	5.153	10.553	8.170	—
Public	Public company	0	1	0.065	—
Spublic	Semi-public company (SEM)	0	1	0.2528	—
Private	Private company	0	1	0.681	—
Managed	Private or semi-public company operating under management (cost-plus) contract	0	1	0.265	—
Net	Private or semi-public company with net cost contract	0	1	0.417	—
Gross	Private or semi-public company with gross cost contract	0	1	0.253	—

[3] A trip usually involves more than one single journey. Typically, if somebody travel in two different buses to reach her destination (a unique trip), the number of journey counted is two.

Table 3: Latent Translog panel production frontier (dependent variable: ln Vehicle-Km)

Non-random parameters	Latent class 1 Coefficients (t-ratio)	Latent class 2 Coefficients (t-ratio)	Latent class 2 Coefficients (t-ratio)
Constant	0.214 (1.521)	0.351 (3.218)*	1.038 (3.591)*
Trend	0.255 (3.126)*	0.314 (4.217)*	0.419 (3.567)*
Trend²	-0.052 (-3.218)*	-0.051 (-4.216)*	-0.048 (3.038)*
ln labour1	0.073 (3.128)**	0.125 (3.673)*	0.129 (3.214)*
ln Labour2	0.321 (2.788)*	0.318 (3.782)*	0.325 (3.035)*
ln Energy	0.521 (3.627)*	0.451 (3.752)*	0.402 (3.318)*
ln Vehicles	0.487 (3.523)*	0.528 (4.128)*	0.507 (4.072)*
ln Network length	0.215 (3.127)*	0.207 (4.129)*	0.225 (3.521)*
ln Population	0.218 (4.234)*	0.251 (3.218)*	0.262 (4.519)*
ln Journey	0.052 (4.215)*	0.065 (4.032)*	0.075 (4.273)*
1/2 ln Labour1* ln Labour1	0.127 (3.783)*	0.153 (2.832)*	0.187 (3.282)*
1/2 ln Labour2* ln Labour2	0.521 (1.945)	0.637 (3.218)*	0.574 (0.378)
1/2 ln Energy* ln Energy	0.832 (4.278)*	0.763 (3.219)*	0.915 (3.021)*
1/2 ln Vehicles * ln Vehicles	0.832 (3.219)*	0.917 (3.178)*	1.021 (2.917)*
1/2 ln Network Lengh * ln Network Length	0.415 (3.218)*	0.521 (4.128)	0.718 (3.016)
1/2 ln Population * ln Population	0.518 (3.812)*	0.485 (2.184)**	0.632 (3.218)
1/2 ln Journey * ln Journey	0.127 (2.583)	0.145 (2.832)	0.183 (3.015)
ln Labour1 * ln Labour2	-1.021 (-0.128)	-0.893 (-1.037)	-0.905 (-0.896)
ln Labour1 * ln Energy	-0.075 (-1.056)	-0.083 (-2.153)	-0.091 (-3.017)*
ln Labour1 * ln Vehicles	0.208 (1.344)	1.551 (2.816)	3.526 (1.231)
ln Labour1 * ln Network Length	0.454 (7.036)	0.369 (5.675)	0.421 (3.781)*
ln Labour1 * ln Population	0.432 (3.887)	0.459 (2.838)	0.0321 (2.219)**
ln Labour1 * ln Journey	0.058 (2.448)	0.110 (3.054)	0.127 (4.381)*
ln Labour2 * ln Energy	0.469 (3.488)*	0.214 (2.950)	0.314 (4.214)*
ln Labour2 * ln Vehicles	0.448 (3.260)	0.368 (3.021)	0.416 (4.783)*
ln Labour2 * ln Network Length	0.488 (3.781)	0.511 (3.966)	0.521 (4.232)*

ln Labour2 * ln Population	0.1362 (4.032)	0.1225 (3.079)	0.314 (4.521)*
ln Labour2 * ln Journey	0.081 (3.675)*	0.437 (5.260)*	0.225 (4.367)*
ln Energy * ln Vehicles	0.057 (3.791)*	0.032 (0.321)	0.128 (0.762)
ln Energy * ln Network Length	0.072 (5.321)*	-0.142 (4.403)*	0.021 (4.218)*
ln Energy * ln Population	0.0321 (3.821)*	-0.142 (-3.289)*	0.073 (4.452)*
ln Energy * ln Journey	0.0217 (3.783)*	0.942 (6.574)*	0.073 (4.452)*
ln Vehicles * ln Network Length	0.072 (5.321)*	-0.142 (4.403)*	0.0452 (3.295)*
ln Vehicles * ln Population	0.0321 (3.821)*	-0.142 (-3.289)*	0.0375 (3.219)*
ln Vehicles * ln Journey	-0.0217 (-3.783)*	-0.942 (-6.574)*	-0.088 (-3.563)*
ln Network Length * ln Population	0.630 (9.517)*	0.282 (5.970)*	0.325 (4.378)*
ln Network Length * ln Journey	0.135 (3.741)*	0.273 (4.233)`*	0.218 (4.893)*
ln Population * ln Journey	0.098 (2.255)*	0.126 (3.673)*	0.218 (3.174)*
Public	0.031 (3.156)*	0.038 (3.125)*	0.052 (1.015)
Spublic	0.027 (3.211)*	0.053 (1.782)	0.085 (4.352)*
Managed)	-0.041 (-3.217)*	-0.055 (-1.918)	-0.053 (-3.152)*
Net	-0.014 (-3.232)*	-0.028 (-1.145)	-0.024 (-3.129)*
Gross	-0.018 (-3.532)*	-0.052 (-1.128)	-0.045 (-3.137)*
Probability of class membership	0.527 (5.317)*	0.317 (4.215)*	0.156 (3.173)*
$\sigma = \left[\sigma_v^2 + \sigma_u^2\right]^{1/2}$	0.338 (3.852)*	0.218 (4.518)`*	0.125 (4.387)*
$\lambda = \sigma_u / \sigma_v$	0.127 (3.218)*	0.102 (3.215)*	0.091 (4.145)*
Log likelihood	1252.132	—	—
Nobs	981	—	—

(t-statistics) in parentheses are below the parameters. Those followed by * are significant at 1% level. Those followed by ** are significant at 5% level.