

Measuring Technical Efficiency in European Railways: A Panel Data Approach

HENRY-JEAN GATHON AND SERGIO PERELMAN

Department of Economics, University of Liège

Abstract

We estimate a factor requirement frontier for European railways using a panel data approach in which technical efficiency is assumed to be endogeneously determined. This approach has two main outcomes. On one hand, it allows the identification of factors influencing technical efficiency, and on the other hand, it allows the estimation of alternative efficiency indicators free of these influences. In the case under study, a particular attention is devoted to an *autonomy* indicator representing the managerial freedom, with respect to public authorities, experienced by firms, that appears to be positively correlated with technical efficiency.

1. Introduction

Beginning with Klein's [1953] seminal econometric study on the US railway companies, the applied economic literature has given considerable attention to the rail activity. This interest can be explained by the special nature of this transportation activity, which can be considered at first glance as rather simple and easy to formalize. Nevertheless, the experience shows that this is a difficult task, especially because of the multi-output character of the production process and the heavily regulated environment in which these companies operate.

This is particularly the case of the European railway sector which is the object of this article. All the companies we observe, except one, are state-owned. They hold a natural monopoly position on the rail transportation, but in return, their activity is constrained by the public authorities. They have to face various regulations concerning, for instance, their fare policy, their investments, the structure of their network, the management of their workforce, etc.. This means that these firms do not enjoy the same autonomy in order to respond to the changing environmental conditions.

Our main concern in this article is to estimate and compare the levels of technical efficiency reached by each of these railway companies over the last thirty years, taking into account the institutional context in which they operate.

With this context in view we estimate a *factor requirement frontier* for the railway transportation, using a panel data approach. This approach is an extension of the methodology proposed by Schmidt and Sickles [1984] who assumed that specific individual effects can be good indicators of technical efficiency when panel data on productive units are available. These authors so apply an old proposition made by Mundlak in 1961, to the context of recent frontier analysis developments.

Nevertheless, one shortcoming of this approach is that the efficiency indicator, like the other estimated parameters, will be biased if performance is correlated with other variables in the model. To overcome this difficulty it is possible to adopt the Hausman and Taylor [1981] approach that consists in a test and an instrumental variables estimation technique. We use this technique, but paying a particular attention to the possible correlation between the technical efficiency, assumed to be endogenous, and the other explanatory variables in the model, among which we include an *autonomy* indicator representing the managerial freedom experienced by each railway company with respect to the public authorities supervision.

Furthermore, as suggested by Mundlak [1978], this last relation is estimated on the basis of an auxiliary equation allowing to obtain some insight into the process leading to inefficiency and to get an alternative measure of performance, free of the influence of factors that are beyond the firm control.

The results obtained for European railways confirm the case of a high correlation between the individual efficiency indicators and the exogenous factors and illustrates the pertinence of the suggested approach in measuring performance within a panel of firms.

In this article we proceed as follows: in the second section we present some methodological considerations; then, in the third section we deal with the data and the specification of the input constrained frontier; finally, in the last section we report the main results of our estimations.

2. The panel data approach

We begin with the specification of a parametric function representative of the railway activity, to be estimated on the basis of available pooled cross-section and time-series data. For reasons we are going to explain below, we adopt here a factor requirements function (Diewert [1974]) of the form:

$$\begin{aligned} y_{it} &= \alpha_0 + X'_{it} \beta + r_t \tau + Z'_i \gamma + \epsilon_{it}, \\ \epsilon_{it} &= u_i + \nu_{it}, \end{aligned} \quad (1)$$

where $i = 1, \dots, I$ and $t = 1, \dots, T$, indicate firms and periods, respectively; y_{it} is the endogenous input and X_{it} is a vector of exogeneous variables that includes, besides the outputs, some variables characterizing the outputs and the technology; r_t and Z_i indicate time-varying and cross-section variables, respectively; α , β , τ , and γ are the parameters to be estimated; and finally ϵ_{it} is a composed error term that combines the time-invariant latent individual effects u_i and the disturbance terms ν_{it} , assumed to be normally distributed and uncorrelated neither with u_i nor with the explanatory variables in the model.

Following Schmidt and Sickles [1984], we assume that the individual effects u_i are indicators of the firm's efficiency. Furthermore, if we assume $u_i \geq 0$, equation (1) corresponds to a special case of the stochastic frontier model introduced by Aigner, Lovell, and Schmidt [1977], and by Meusen and van den Broek in 1977. The difference lies in the fact that for panel data, equation (1) provides a natural way to discriminate between the efficiency indicator and the noise.

Then, in order to be able to apply the main results of the panel data literature, we transform equation (1) into:

$$y_{it} = X'_{it} \beta + r_t \tau + Z'_i \gamma + \alpha_i + \nu_{it}, \tag{2}$$

where $\alpha_i = \alpha_0 + u_i$.

The two traditional and alternative methods proposed to deal with latent individual effects are the so called *fixed* and *random effects* models. Each of them allows us to calculate an alternative and specific measure of efficiency.

2.1. Fixed and random effects models

In the first case we obtain the *Within group* estimators, indicated by the subscript *W*, by estimating the following equation:

$$\tilde{y}_{it} = \tilde{X}'_{it} \beta_W + \tilde{r}_t \tau_W + \epsilon_{W,it} \tag{3}$$

and

$$\hat{\alpha}_{W,i} = \bar{Y}_i - \bar{X}'_i \hat{\beta}_W,$$

where \bar{k}_i indicates mean values within groups; \sim indicate deviations from individual means, and \wedge estimated values; and $\epsilon_{W,it} = \tilde{\nu}_{it}$ is the normal disturbance term.

The estimated levels of performance $\hat{\mu}_{W,i}$ are then obtained on basis of the estimated fixed effects $\hat{\alpha}_{W,i}$ by assuming that the most efficient firm in the sample corresponds to $\min(\hat{\alpha}_{W,i})$ and that the inefficiency level is given by the distance $\hat{\mu}_{W,i} = \hat{\alpha}_{W,i} - \min(\hat{\alpha}_{W,i})$. Note that when the input requirement function adopts a logarithmic shape—as it will be in our application—the corresponding measure of efficiency will be: $\hat{\mu}_{W,i} = \exp[\hat{\alpha}_{W,i} - \min(\hat{\alpha}_{W,i})]$. Then $\hat{\mu}_{W,i} = 1$ will indicate the most labor efficient firm identified by $\min(\hat{\alpha}_{W,i})$; otherwise, $\hat{\mu}_{W,i} > 1$ will indicate the degree of labor overutilization with respect to the requirement frontier.

In the case of the random effects model, we obtain generalized least squares (GLS) estimators by performing ordinary least squares (OLS) on the transformed variables:

$$(y_{it} - \theta \bar{y}_i) = (X_{it} - \theta \bar{X}_i)' \beta_{GLS} + (r_t - \theta \bar{r}_t) \tau_{GLS} + (1 - \theta) Z'_i \gamma_{GLS} + \epsilon_{GLS,it}, \tag{4}$$

where

$$\theta = 1 - (\hat{\sigma}_\nu^2 / (\hat{\sigma}_\nu^2 + T \hat{\sigma}_\alpha^2))^{1/2}, \epsilon_{GLS,it} = (1 - \theta) \alpha_i + (\nu_{it} - \theta \bar{\nu}_i),$$

and $\hat{\sigma}_\nu^2$ and $\hat{\sigma}_\alpha^2$ can be estimated, for instance, by performing OLS on equation (2) as suggested by Wallace and Hussein [1969].

The random effects are computed as $\hat{\alpha}_{GLS,i} = 1/T \sum_t \hat{\epsilon}_{it}$ on basis of the estimated mean firms residuals obtained also by applying OLS to equation (2). As before, the performance levels are obtained by: $\hat{\mu}_{GLS,i} = \hat{\alpha}_{GLS,i} - \min(\hat{\alpha}_{GLS,i})$.

As expected, these two alternative models yield different results. We thus have to choose the one that most fits the sample and the purpose of the analysis. A way to proceed is to rely on the statistical properties of the sample, in which case we would apply the random effects model for random samples and the fixed effect model otherwise. However, this criteria may be misleading, as shown by Mundlak [1978] when the possibility of correlation between the latent individual effects, α_i , and the other variables in the model is introduced.

Following Mundlak [1978], this relation can be formalized through an auxiliary equation, that in the present case takes the form:

$$\alpha_i = \bar{X}_i' \pi + Z_i' \xi + \omega_i \tag{5}$$

where π and ξ are parameters and ω_i is an error term with the normal properties.

Considering the aim of this article (performance assessment) it is clear that this last relation may have a special interest. Recall that we have chosen to assume (equation 2) that the α_i 's are indicators of efficiency free from the noise which is caught by the ν_i error term. Indeed, among all the variables included in equation (2), we can reasonably consider that those of them characterizing the technology and the environment, including regulation, may affect the technical efficiency of the firm.¹

Furthermore, as Mundlak showed, the random effects model (4) and the fixed effects model (3) are special cases of the general model described by equations (2) and (5). The first one corresponding to the special case of absence of correlation, i.e., $E[\alpha_i | X_{it}, Z_i] = 0$ (or $E[\mu_i | X_{it}, Z_i] = 0$), and the second one corresponding to the opposite case in which the α_i (or μ_i) are correlated with all the variables in the model.

Then we have to determine the most suited model to handle this problem. Hausman and Taylor [1981] (hereafter H&T) provide a method for that purpose. On one hand, they presented a specific test on the correlation between individual-specific effects and explanatory variables, and on the other hand they introduced an instrumental variable estimation technique in order to obtain non-biased and efficient estimators in the case that not all variables will be correlated with the individual effects.

The test performed by H&T is based on the Within and GLS estimators. As mentioned before, these estimators correspond to the extreme cases in which all or none of the variables are correlated with the individual latent effects. We perform:

$$\hat{m} = \hat{q}' \text{cov}(\hat{q})^{-1} \hat{q} \tag{6}$$

where $\hat{q} = \hat{\beta}_{GLS} - \hat{\beta}_w$ and $\text{cov}(\hat{q}) = \text{cov}(\hat{\beta}_w) - \text{cov}(\hat{\beta}_{GLS})$. We test the null hypothesis $H_0 = E[\alpha_i | X_{it}, Z_i] = 0$ against the alternative $H_1 = E[\alpha_i | X_{it}, Z_i] \neq 0$ hypothesis on basis of a χ^2 test. If the null hypothesis is accepted, then IV estimation procedure is required.

2.2. The GLS-Instrumental Variable model

The advantage of the GLS-IV model proposed by H&T is twofold. First, all the instrumental variables can be selected within those already present in the original model. Second, we

can estimate the γ parameters for the variables of Z_i type, estimation which is not possible with the Within effects model.²

In fact, the problem of potential correlation between individual latent effects and the explanatory variables is limited, by definition, to the cross sectional level. This is the reason why in panel data analysis the IV model, designed to deal with this problem, can be applied without adding supplementary variables. The IV technique can be implemented by introducing a distinction between the exogeneous ($X1_{it}, Z1_i$)—non-correlated with α_i —variables³ and the endogenous ($X2_{it}, Z2_i$) ones. We then apply the instrumental variable technique to equation (4) in the GLS model using ($X1_{it}, \bar{X}1_{it}, \bar{X}2_{it}, Z1_i$) as instruments. The only difference is that the variance components are estimated from a previous two stage least squares procedure in which consistent estimators of β and γ can be obtained. Furthermore, in order to avoid identification problems, it is necessary to check that $k_1 > g_2$, where k_1 and g_2 indicate the number of variables in $X1_{it}$ and $Z2_i$, respectively.

Proceeding in this way, we obtain consistent and efficient estimators when the effects are correlated with the explanatory variables (H&T [1981]). Indicating by a suffix IV these estimators, we calculate the individual effects by performing:

$$\hat{\alpha}_{IV,i} = \bar{Y}_i - \bar{X}'_i \hat{\beta}_{IV} - Z'_i \hat{\gamma}_{IV}, \tag{7}$$

and the corresponding measures of efficiency: $\hat{\mu}_{IV,i} = \hat{\alpha}_{IV,i} - \min(\hat{\alpha}_{IV,i})$.

2.3. Measuring net efficiency

Summing up, once we have decided which of the three alternative models will be used to estimate performances, we turn to the analysis of the factors influencing technical efficiency. One possibility, however, is that individual effects may be uncorrelated with the exogeneous variables present in the model. In that case the GLS estimators will be consistent and efficient and the $\hat{\mu}_{GLS,i}$ will be the univocal measures of performance. Otherwise, we turn to the estimation of the auxiliary function (5).

Here, we adapt the estimation procedure proposed by Mundlak [1978] to the case of incomplete correlation, by performing an OLS regression between estimated individual effects ($\hat{\alpha}_{IV,i}$) and all the variables identified as factors influencing technical efficiency, that is $\bar{X}2_i$ and $Z2_i$:

$$\hat{\alpha}_{IV,i} = \bar{X}2'_i \pi + Z2'_i \xi + \omega_i. \tag{8}$$

Then, equation (8) can be used in order to analyze the influence of the factors on efficiency and, as far as these factors are believed to be out of control of the managers, it is possible to re-estimate net efficiency measures purged from these influences by doing: $\bar{\mu}_{N,i} = \hat{\omega}_i - \min(\hat{\omega}_i)$. Where the suffix N indicates the net character of these measures.

Finally, when the hypothesis of correlation between the individual effects and all the explanatory variables in the model cannot be rejected, the net indicators of efficiency must be estimated on the basis of the Within model estimators as indicated by equation (5).⁴

It is important to note that, at this point, our approach differs radically from Schmidt and Sickles [1984]. Dealing with a traditional production function, these authors assume that mismanagement affects the choice of the input intensity. In other words, the causality between technical inefficiency and explanatory variables is assumed to go in the opposite direction as we suppose here. As a consequence, for these authors the efficiency measures obtained under the random effects model ($\hat{\mu}_{GLS,i}$) must be considered as definitive.

Our claim is that the alternative measures of performance presented here are free of bias caused by internal correlation. Therefore we can rank all the firms as if they were free from the constraints imposed by exogeneous conditions. At the same time these evaluations stress the potential capacity of policies designed to lighten these constraints.

2.4. Maximum-Likelihood estimators

We also present the results obtained from the estimation of equation (2) under the assumption that technical efficiency is half-normally distributed. As well known, this type of distribution is generally used for efficiency measurement in the stochastic frontier literature (e.g. Jondrow et al. [1982]). In the present case, we focus on the results obtained by a Maximum Likelihood (ML) procedure proposed by Battese and Coelli [1988]. Specially designed to deal with panel data, this procedure amounts to estimate the conditional expectation, $\mu_{ML,i}$, of the individual efficiencies, as follows:

$$\mu_{ML,i} = E[\exp(\alpha_i) \mid \alpha_i + \nu_i] = 2 \exp(\sigma_\alpha^2/2)[1 - \Phi(\sigma_\alpha)], \quad (9)$$

where $\Phi(\cdot)$ indicates the standard normal distribution function.⁵

These estimations will be performed under the assumption of absence of correlation between the α_i and the other variables present in the model. Nevertheless, they will be included here in order to be compared with those obtained with the alternative procedures described above.

3. Model specification and data

Our main purpose is to estimate an indicator of technical efficiency within a panel composed of 19 European railway companies observed over the period 1961–1988. In table 1, we present the main characteristics of these companies.

As indicated before, we adopt here a parametric approach based on the estimation of a so-called *factor requirement function* (Diewert [1973, 1974]) with labor units as the dependent variable.⁶ Even if it is not a usual way to modelize production activities, it appears to us as a promising field for railways which produce multiple outputs and operate in heavily regulated conditions. In fact, we prefer using this type of function instead of a usual production function because of the problems encountered in obtaining functional forms for multiple output production functions. We could alternatively use a dual cost function. Unfortunately, the cost function approach requires data on the railways' input prices, data which are not very reliable in the present—international and heavily regulated—context.

Table 1. Sample descriptive statistics (mean values over the period 1961-1988)¹

Railways	Country	Train—km				Mean Distances			Load Factors			Autonomy % ²
		Labor (10 ³)	Passenger (10 ⁶)	Freight (10 ⁶)	Lines (km)	Passenger (km)	Freight (km)	Passenger (per train)	Freight (tons per train)	Electrification %		
BLS	Switzerland	1.9	4.9	1.1	237	19.8	44.9	61.4	220.9	100.0	100.0	
BR	United-Kingdom	229.0	343.6	104.2	19439	40.2	118.4	89.4	204.2	17.4	76.3	
CFF	Switzerland	38.5	64.1	28.0	2934	38.2	150.3	131.9	218.1	99.4	66.0	
CFL	Luxembourg	4.0	2.9	1.5	292	21.2	37.7	75.4	412.3	49.1	63.5	
CH	Greece	11.6	12.8	3.2	2534	133.3	217.2	107.7	198.4	0.0	47.3	
CIE	Ireland	8.5	7.6	4.3	2213	58.4	165.8	100.6	123.2	0.0	58.3	
CP	Portugal	24.5	24.5	6.2	3595	27.1	217.6	151.3	146.6	11.0	64.0	
DB	Germany	351.8	384.7	196.3	28949	38.2	199.8	100.7	304.6	30.4	61.0	
DSB	Denmark	19.5	35.7	8.3	2256	30.9	235.3	99.7	200.1	4.5	41.5	
FS	Italy	200.1	210.3	60.3	16366	98.3	324.1	164.9	279.8	50.9	67.0	
NS	Netherlands	27.3	82.6	15.3	2984	44.1	146.3	101.1	211.6	57.6	70.3	
NSB	Norway	16.8	22.9	10.2	4253	57.7	99.4	83.5	240.5	53.9	45.3	
ÖBB	Austria	70.8	56.4	33.6	5834	41.1	201.2	118.9	285.2	44.5	41.8	
RENFE	Spain	78.5	85.0	44.3	13162	80.9	316.1	162.0	209.6	30.0	52.3	
SJ	Sweden	37.9	64.1	40.7	11837	82.2	284.3	84.6	354.3	59.3	80.0	
SNCB	Belgium	54.6	64.4	21.9	4344	40.1	108.3	112.9	331.0	31.4	64.5	
SNCF	France	270.7	261.3	204.1	36305	73.7	303.3	179.4	309.4	26.3	69.8	
TCDD	Turkey	58.4	20.7	18.2	8089	49.4	449.5	247.2	316.7	1.2	60.0	
VR	Finland	24.7	25.2	18.4	5811	86.0	269.3	106.2	353.2	0.6	40.0	

¹The variables are defined in the Appendix (Source: UIC [1961-1988]).

²Values obtained on the basis of a survey conducted in 1988-89 (Gathon [1991]).

The use of the labor function as a valid representation of the productive process of the railroads requires an assumption of complementarity (fixed proportions) between all the main inputs (labor, capital and fuel) needed to supply the rail transportation. This assumption can be considered as plausible not only from a technical but also from an empirical viewpoint.⁷ Furthermore, for our sample, labor expenses account for about 90% of the variable cost⁸ throughout the whole period and for all the railways, that allow us to assume that the substitution possibilities between labor and energy are highly limited.

3.1. The factor requirement function

The function we estimate is log-linear in outputs. It can be represented by equation (2), except that we include only one time-invariant variable z_i :

$$y_{it} = \alpha_i + X'_{it} \beta + r_t \tau + z_i \gamma + \nu_{it}, \quad (i = 1, \dots, 19; t = 1, \dots, 28), \quad (9)$$

where

y_{it}	: labor
X_{it}	: 8 time- and firm-varying variables
$x_{1,it}$: train-km (passengers)
$x_{2,it}$: train-km (freight)
$x_{3,it}$: km of lines
$x_{4,it}$: mean distance (passengers)
$x_{5,it}$: mean distance (freight)
$x_{6,it}$: load factor (passengers)
$x_{7,it}$: load factor (freight)
$x_{8,it}$: electrification
r_t	: the trend (firm-invariant)
z_i	: autonomy (time-invariant).

All these variables, but the trend, are expressed in natural logarithms.⁹

The first two variables in X_{it} are the main outputs of the railways, i.e., the transportation of passengers and freight, evaluated in train-kilometers (train runs performed during one year). The third variable is the length of the network which is here assumed to be an output since an important part of the railroad staff is assigned to the maintenance of tracks and stations independently of the rail traffic intensity.

Variables $x_{4,it}$ to $x_{7,it}$ are output characteristics that allow us to take into account the nature of the demand side in the analysis. We comply here to a tradition¹⁰ that consists in the consideration of some output characteristics as factors determining the production technology. Variables $x_{4,it}$ and $x_{5,it}$ are two indirect indicators of the density of the network, as postulated by Caves et al. [1981]. The load factors $x_{6,it}$ and $x_{7,it}$ (passengers and tons by trains) are used as proxies for the intensity of passengers and freight services demand.

The electrification variable $x_{8,it}$ (% of the electrified network) is assumed to be a good indicator of the technology chosen by the railways. The trend variable, r_t , is included in

order to catch the general improvement in productivity not controlled by the other variables in the model. In accordance with Nishimizu and Page [1982], we assume here that productivity gains have two different sources: the results of progress in technology and in technical efficiency. As we will see hereafter, this is a way of relaxing the assumption of constancy of the general level of performance overtime.

Finally, the variable *autonomy*, which is an index of the regulatory and institutional environment that the companies have to face, has been constructed from a survey we conducted among these companies in 1988–1989 (Gathon [1991]). This survey aims at evaluating the autonomy enjoyed by the railways management with respect to public authorities. The numerical value of this index ranges between 40 and 100. The more autonomous the management, the higher is the value of this index.

A difficulty with this variable is the fact that it was observed at the end of the analyzed period. In spite of this, we include this autonomy index in equation (9), next to the other variables, assuming that the institutional environment of the national railways moves little and very slowly.¹¹

3.2. *The panel*

On table 1, we reproduce the mean individual values of these variables for all the companies. As it can be seen, the sample includes the Turkish company, partially outside Europe and two companies for Switzerland, one of them (BLS) being private. This aside, the panel consists of European national railways that are for a large part interconnected. The length of the observed period (1961–1988) gives us an excellent basis for comparison and allows us to obtain more confident statistical results.

Some interesting facts can be outlined from table 1.

First, the large scale variations across railroads. The largest firms, BR, DB, and SNCF are more than one hundred times bigger, in terms of lines, labor or outputs, than the smallest companies, BLS and CFL.

Second, we also observe some important differences across firms, in the output composition. Some companies such as SNCF, VR and TCDD, display a fair balance between passengers and freight transportation. Other railroads, such as NS and DSB, are specialized in passenger services, leaving freight traffic to other transportation modes.

Third, mean distances and load factors vary substantially across countries. These variables seem to be partially correlated with the size of the network, but, at the same time, with its density and structure. Concerning the electrification process, we observe that both Swiss companies reach a 100 percent level, while on the opposite, some companies continue to operate a network with fuel traction only.

Note that these exogeneous variables, together with the index of regulation, can be at the same time factors of the production process and factors influencing the performance. However, we do not introduce a priori information about which variables should influence the technical efficiency, except for *autonomy* which is, by construction, expected to be correlated with performance.

4. Estimation and main results

As stated in section 2, we proceed first with the estimation of the Within and GLS models for the labor constrained function. The results of these estimations are presented in columns (a) and (b) of table 2.

Table 2. Estimated parameters for alternative models.¹

<i>Dependent variable: labor; number of observations: 532</i>					
Variables	Parameters	(a) Within	(b) GLS	(c) GLS-IV	(d) ML
<i>Train-km:</i>					
passengers	β_1	0.223 (0.052)	0.380 (0.041)	0.254 (0.045)	0.269 (0.048)
freight	β_2	0.128 (0.041)	0.149 (0.039)	0.159 (0.062)	0.136 (0.039)
Km of lines	β_3	0.493 (0.074)	0.553 (0.054)	2 0.610 (0.060)	0.515 (0.057)
<i>Mean distance:</i>					
passengers	β_4	0.089 (0.089)	-0.142 (0.062)	0.185 (0.055)	-0.070 (0.075)
freight	β_5	-0.210 (0.062)	-0.200 (0.052)	2 -0.183 (0.059)	-0.192 (0.055)
<i>Load factor:</i>					
passengers	β_6	0.015 (0.040)	0.131 (0.040)	0.019 (0.072)	0.041 (0.041)
freight	β_7	0.067 (0.040)	0.097 (0.042)	0.071 (0.068)	0.063 (0.041)
Electrification	β_8	-0.015 (0.005)	-0.020 (0.004)	-0.013 (0.008)	-0.016 (0.005)
Trend	τ	-0.010 (0.001)	-0.011 (0.001)	-0.010 (0.002)	-0.011 (0.001)
Autonomy	γ	—	-0.175 (0.135)	0.359 (0.081)	-0.054 (0.089)
Intercept	α_0	—	1.66 (0.69)	-0.471 (0.59)	1.94 (0.44)
χ^2 (d.f. = 1)		—	105.7	6.63	—

¹Standard errors are presented in brackets.

²Instrumental variable.

As can be observed, the sum of the estimators corresponding to the three outputs (passengers and freight train-km and km of lines) indicates some economies of scale¹² in the Within model ($\hat{\beta}_{W,1} + \hat{\beta}_{W,2} + \hat{\beta}_{W,3} = 0.844$) and diseconomies of scale in the GLS model ($\hat{\beta}_{GLS,1} + \hat{\beta}_{GLS,2} + \hat{\beta}_{GLS,3} = 1.082$). Note that the Within estimators can be interpreted as short-run estimators, since they are performed on variables expressed in deviations from individual means.

Furthermore, concerning the variables controlling for the output characteristics, we observe that overall they present the expected signs. On one hand, average passengers and freight distances appear, in all cases but one, as labor saving factors. On the other hand, the two load factors variables are associated with positive coefficients. Given that the railway output is measured in terms of train-km, a higher loading of the trains is related to a higher demand for the supplied transportation capacity and probably implies more labor consumption. It is important to note, the two coefficients are only statistically significant under the GLS model.

The coefficients associated with the electrification variable (a proxy for the technology chosen by the railways) and with the trend, unambiguously indicate on the one hand that the electrification of the network is labor saving and on the other hand that the sector experienced productivity growth at a rate of about 1 percent each year. Recall that we will interpret the effect of the trend variable as reflecting a technological progress that also includes an overtime general technical efficiency improvement within the sector.¹³

Finally, under the GLS model, the autonomy variable, which is the only time-invariant variable of the model, behaves as a shift factor of the estimated function, influencing labor use favorably, i.e., negatively, but not in a significant way.

However, as indicated in section 2, the Within and GLS estimators correspond to two extreme models, characterized by general or zero correlation between latent individual effects and explanatory variables, respectively. In order to choose between these models we test the null hypothesis of zero correlation as proposed by H&T [1981].

We proceed to the estimation of the statistic \hat{m} (equation (6)) using the Within and GLS estimators presented before. On the basis of a χ^2 test ($df = 1$) and the value of \hat{m} ($\hat{m} = 105.7$), we reject the null hypothesis of zero correlation. Therefore, in order to obtain unbiased and efficient estimators we use the GLS-IV procedure presented in section 1. Note that this procedure has the particularity that the instrumental variables will be chosen between the explanatory variables present in the model. Consequently, we select three of them—km of lines, mean freight distance, and the trend—as instruments following a search path aiming at minimizing the value of the statistic \hat{m} obtained from GLS-IV and Within estimators comparisons.

The results of this estimation are reported in column (c) of table 2. Given that we retain only three instrumental variables the results are, not surprisingly, very close with those corresponding to the Within model. Only two differences may be outlined. On one hand, the coefficient associated with the size of the network is higher (0.609) and thus the labor requirement function is now characterized by nearly constant economies of scale ($\hat{\beta}_{IV,1} + \hat{\beta}_{IV,2} + \hat{\beta}_{IV,3} = 1.023$). On the other hand, the *autonomy* variable, missing under the Within model, presents a positive coefficient.

Nevertheless, these results are not cautioned by the H&T test. When we compare the GLS-IV estimators versus the Within estimators, the value of \hat{m} is now equal to 6.63 and has a weak probability of 1 percent to be accepted on the basis of a χ^2 test ($df = 1$). That means that some correlation between the individual effects and the explanatory variables subsist in the GLS-IV results. Given this fact, we are inclined to consider the estimations obtained under the Within model as very probable too.¹⁴

In the last column of table 2 we report the estimated coefficients corresponding to the maximum likelihood (ML) model developed by Battese and Coelli [1988]. Recall that this model was estimated, like the GLS model, under the hypothesis of absence of correlation between the efficiency indicators and the explanatory variables. The coefficients obtained are, unexpectedly, very close with those corresponding to the fixed effects model.

On table 3, we report the levels of technical efficiency corresponding to the four alternative models (Within, GLS, GLS-IV, and ML) and the 19 railway companies in the sample. These measures were calculated as indicated in section 1. In all the cases, except for the ML model, we obtain these indicators by normalizing the specific individual effects to the most performant company, as it was first suggested by Greene [1980]. Moreover,

Table 3. Labor technical inefficiency of railway companies (period 1961-1988).¹

Railways	Country	Net measures of efficiency											
		(a)		(b)		(c)		(d)		(e)		(f)	
		Rank	$\hat{\mu}_{w,i}$	Rank	$\hat{\mu}_{GLS,i}$	Rank	$\hat{\mu}_{IV,i}$	Rank	$\hat{\mu}_{ML,i}$	Rank	$\hat{\mu}_{w,i}$	Rank	$\hat{\mu}_{IV,i}$
BLS	Switzerland	1	1.00	4	1.28	4	1.65	3	1.04	8	1.24	9	1.15
BR	United-Kingdom	15	3.33	11	1.62	10	2.54	15	2.41	15	1.39	15	1.36
CFE	Switzerland	12	2.62	12	1.77	11	2.85	12	2.12	7	1.23	14	1.33
CFL	Luxembourg	8	1.86	17	2.28	17	3.52	10	1.92	17	1.48	17	1.54
CH	Greece	5	1.32	7	1.47	5	1.67	5	1.17	11	1.29	12	1.28
CIE	Ireland	3	1.21	5	1.29	3	1.57	4	1.09	15	1.39	16	1.37
CP	Portugal	11	2.51	14	1.81	12	2.99	11	2.09	10	1.26	6	1.06
DB	Germany	18	4.14	12	1.77	15	3.16	17	2.89	14	1.36	11	1.23
DSB	Denmark	9	2.18	8	1.50	14	3.09	8	1.78	6	1.21	4	1.02
FS	Italy	19	4.26	18	2.53	16	3.19	19	3.12	19	1.77	19	1.71
NS	Netherlands	7	1.84	6	1.30	7	1.96	7	1.50	2	1.06	2	1.01
NSB	Norway	2	1.19	2	1.03	2	1.48	1	1.03	4	1.12	7	1.09
ÖBB	Austria	16	3.54	16	2.18	19	4.13	16	2.77	18	1.64	18	1.58
RENFE	Spain	10	2.42	10	1.56	9	2.15	9	1.83	12	1.32	13	1.29
SJ	Sweden	4	1.28	1	1.00	1	1.00	1	1.03	13	1.33	5	1.03
SNCB	Belgium	13	2.81	15	1.88	12	2.99	14	2.29	5	1.13	9	1.15
SNCF	France	14	3.15	9	1.56	8	2.09	13	2.22	1	1.00	2	1.01
TCDD	Turkey	17	3.65	19	2.59	18	3.65	18	2.94	9	1.25	7	1.09
VR	Finland	6	1.48	3	1.21	6	1.70	6	1.21	3	1.09	1	1.00
Mean			2.41		1.67		2.49		1.92		1.29		1.23
Std. dev.			1.05		0.47		0.87		0.71		0.19		0.21
Table of correlation ²		Within	1.0		0.812		0.807		0.982		0.289		0.328
		GLS		1.0		0.918		0.879		0.416		0.539	
		GLS-IV			1.0		1.0		0.851		0.372		0.396
		ML				1.0			1.0		0.319		0.394
		Net Within								1.0			0.835
		Net GLS-IV									1.0		1.0

¹The values are normalized; the value of the most efficient firm equals to 1.

²Spearman rank correlation coefficients, except for columns (d).

given the reduced number of firms, the estimated levels of inefficiency are consistent for relative comparative analysis but not reliable as indicators of absolute levels of inefficiency.¹⁵ Finally, note that in the context of the labor requirement function we estimate here, technical efficiency corresponds to level 1.0, otherwise, the degree of labor over-utilization is indicated by values greater than 1.0.

A first lecture of table 3 shows an extremely wide dispersion in the Within estimators of performance ($\hat{\mu}_{W,i}$) that in addition appear also highly correlated with the scale. The most and least efficient firms are the smallest (BLS) and two out of the biggest (DB and FS) of the panel, respectively. These results confirm the fact that individual fixed effects are not reliable as indicators of efficiency. They tend to catch differences between firms, other than efficiency, existing at the cross-section level. That is also the case of measures obtained under the ML procedure ($\hat{\mu}_{ML,i}$), presented in column (d) of table 3.

At the opposite, under the random effects model, the efficiency scores ($\hat{\mu}_{GLS,i}$) seem to be free of scale influence, and the variance seems to be more reliable. Moreover, let's remember that these results were obtained under the hypothesis that $E(\mu_i | X_{it}, Z_{it}) = 0$, an hypothesis that was severely rejected by the H&T test.

The table of rank correlation coefficients reproduced at the bottom of table 3 shows the degree of convergence between these measures of technical efficiency. Unexpectedly, we obtain high correlation coefficients between the Within, GLS and ML measures. Nevertheless, one can see some noteworthy changes in efficiency measures, from column (a) to column (b), for large railways such as BR and SNCF on the one hand, and for small companies such as BLS and CIE, on the other hand.

Then, we turn to the efficiency estimators obtained using the GLS-IV procedure presented in section 2.2. These indicators, indicated as $\hat{\mu}_{IV,i}$, are reported on column (c) of table 3. It can be noted that the measures of technical efficiency derived from the instrumental variables procedure present high mean and dispersion levels as under the Within model but appear as highly correlated (0.918) with the GLS ranks.

However, as we explained in section 2.3, the indicators of efficiency corresponding to the Within and GLS-IV models cannot be considered as definitive if they are not purged from the influence of the explanatory variables. In order to do that we estimate the auxiliary functions (5) and (8) corresponding to the Within and GLS-IV models, respectively. The results are presented on table 4.

What can we learn from these results? As argued before, the estimated $\hat{\pi}_k$ and $\hat{\xi}$ coefficients can be interpreted as the effect of explanatory variables on the unpurged technical inefficiency indicators. Then, if we adopt this point of view and consider the two main outputs, we can note that more passengers transportation leads to increased inefficiency and vice-versa for freight transportation but with a hardly significant coefficient. Note that, apart from some cases that we will indicate, the coefficients presented in columns (a) and (b) of table 4 are closed each together and that the coefficients corresponding to the two selected instruments (km of lines and mean freight distance) have no significant influence on efficiency under the Within model.

The density factor (mean trip) appears as an exogeneous factor improving efficiency. In other words, long passengers trips or, equivalently, a low density of the network, facilitates the achievement of the efficiency goal. At the opposite, the load factors (defined as the number of passengers or tons by train) reduces efficiency. As expected, for a given

Table 4. The auxiliary function: effect of exogeneous variables on technical inefficiency¹

Variables	Parameters	(a)	(b)
		Within	GLS-IV
Train-km:			
passengers	π_1	0.362 (0.138)	0.332 (0.144)
freight	π_2	-0.087 (0.197)	-0.244 (0.141)
Km of lines	π_3	-0.005 (0.181)	2
Mean distance:			
passengers	π_4	-0.477 (0.159)	-0.708 (0.142)
freight	π_5	-0.206 (0.180)	2
Load factor:			
passengers	π_6	1.029 (0.237)	0.760 (0.181)
freight	π_7	0.467 (0.223)	0.548 (0.225)
Electrification	π_8	-0.043 (0.025)	-0.037 (0.026)
Autonomy	ξ	-0.265 (0.228)	-0.524 (0.230)
Intercept		-5.53 (1.85)	-1.94 (1.94)
R ²		0.904	0.810
n		19	19

¹Standard errors are presented in brackets.

²Instrumental variable.

production—evaluated by the number of train-km and the length of the network—higher demand implies more input requirements as it was confirmed by the estimation of the labor requirement function (see table 2). Moreover, it appears that this phenomenon is also partially caught by the efficiency indicator and this fact justifies the computation of an indicator free of this influence.

The electrification rate, that we interpreted as an indicator of technology, here again appears as an element in favor of efficiency. Finally, the autonomy variable has, as expected, a negative impact on technical inefficiency (less significant under the Within model). To some extent, this confirms the existence of a positive correlation between the autonomy enjoyed by a firm and its performance.

Finally, in columns (e) and (f) of table 3, we present the net measures of inefficiency estimated on the basis of the residuals ($\hat{\omega}_i$) of the auxiliary functions. Some notable cases aside—CP, DSB, SJ, and TCDD—most of the companies present similar ranks and levels of inefficiency under the two models. Otherwise, as expected, these indicators show a lower dispersion than those obtained before and dramatic changes in evaluation and reclassification of railway performances that confirms the need and rectitude of net measures of efficiency to take into account the environment in which these firms operate.

5. Conclusions

In this article we have presented a method in order to evaluate the technical efficiency of regulated firms such as the European railroads companies. We have stressed and illustrated

the fact that, dealing with panel data, the correlation between efficiency and explanatory variables must be taken into account and alternative net measures of efficiency are needed. We believe that research in the field of firms' performance and frontier analysis should be more oriented in this way.

Furthermore, the factor requirements function used in our analysis appears as a simple and convenient way of modelling the productive activity of firms such as the railways, that are highly regulated and do not display strong substitutability possibilities between inputs.

Finally, we observed a negative correlation between institutional managerial autonomy and technical inefficiency. Of course this result requires additional confirmation but it shows at least that when this factor, and other variables affecting the performance, are taken into account the spectrum of inefficiency across firms narrows and becomes more reliable.

Acknowledgments

We are grateful to P. Barla, G.E. Battese, C.A.K. Lovell, G. N'Gbo, P. Pestieau, B. Sak, B. Thiry, and J.P. Urbain for helpful discussions. D.H. Good, P. Schmidt, R. Sickles and the participants at the Third Conference on Panel Data, ENSAE, Paris, June 11–12, 1990 provided constructive comments on an earlier draft.

Appendix. Variables definition and sources

Variable	Description
Labor	Annual mean railway staff (total number of workers assigned to the rail operation)
Mean distance:	
passengers	Average length of passenger journey
freight	Average length of haul of one ton of freight
Train-km:	
passengers	Train kilometers by passenger trains
freight	Train kilometers by freight trains
Load factor:	
passengers	Number of passengers per train
freight	Number of freight tons per train
Km of lines	Length of rail lines operated at the end of year
Electrification	Percentage of electrified lines
Trend	Trend: 1, ..., 28 (= 1 in 1961; ...; = 28 in 1988)
Autonomy	Autonomy enjoyed by the railway management with respect to public authorities, index of regulatory and institutional environment

All the data (except *autonomy* are from the International Railways Statistics yearbook (UIC [1961–1988]) published from 1961 to 1988 by the International Railway Union (UIC). Our sample of observations is of cross-section-time-series nature: 19 European railroads over the 1961–1988 period. These railroad companies were selected on two grounds: availability of data and comparability. When gathering data from each individual company, the UIC tries to insure the highest homogeneity and comparability in the definition and the measurement of both inputs and outputs. The variable *autonomy* has been constructed from a survey, as indicated in the text (section 2).

Notes

1. The introduction of environment factors affecting efficiency and the estimation of net efficiency measures were previously analyzed by Perelman and Pestieau [1988] and by Deprins and Simar [1988], in the context of deterministic production frontier analysis.
2. This model was first introduced in frontier analysis by Schmidt and Sickles [1984]. Note that the GLS-IV estimators must be checked against the Within estimators by the H&T test in order to verify the choice of the instrumental variables (Hausman and Taylor [1981] pp. 1388–89).
3. In order to simplify the presentation, the time-varying regressor r , that by definition is an exogenous variable, is included here in vector X_1 .
4. Note that in this case, the parameters on the auxiliary function (5) can be also obtained by $\hat{\pi}_W = \hat{\beta}_B - \hat{\beta}_W$ and $\hat{\xi}_W \equiv \hat{\gamma}_B$, where the suffix B indicates the *Between* estimators obtained by performing OLS on equation (2) with all the variables expressed in individual means (Mundlak [1978]).
5. Note that we assume here a half-normal distribution for α_i ; this is a particular case of the more general model suggested by Battese and Coelli [1988] that considers also normal distributions truncated at points different from the mean. The program we use has been developed by T. Coelli who kindly transmitted it to us (see Coelli [1989]).
6. For an application of this type of function, also called *inverse production function* in the case of single input-multiple output technology, see for instance Bjurek et al. [1990].
7. Perelman and Pestieau [1987] estimated a production frontier for the European railways using capital and labor as the only inputs. They found little evidence of substitutability between capital and labor.
8. Labor and energy costs.
9. Details about the definition and the source of the variables are given in the Appendix.
10. See, for instance, Wang Chiang and Friedlander [1984].
11. One could, of course, argue about this point. From our survey among the railways, it appears that this institutional environment did vary frequently over the analyzed period for companies such as BR, CFF, and DSB. Nevertheless, for a vast majority of the railways, this environment was about the same in the early sixties as in the late eighties.
12. We estimate a labor requirement function under the hypothesis of strict complementarity of inputs. Therefore, the term *economies of scale* has the usual meaning.
13. Alternative treatments with time-varying individual efficiency effects has been proposed by Kumbhakar [1988], Cornwell et al. [1990], and Battese and Coelli [1991].
14. We perform also alternative GLS-IV models as suggested by Amemiya and MaCurdy [1986] and Breusch et al. [1989]. The results obtained confirm the case of severe correlation between technical efficiency and all the explanatory variables in the model.
15. Note that Schmidt and Sickles [1984] stated the conditions for the consistency of estimators (other than the efficiency indicator). They are met by our sample that is composed by nearly 30 years.

References

- Aigner, D.J., C.A.K. Lovell, and P. Schmidt. (1977). "Formulation and estimation of stochastic frontier production function models." *Journal of Econometrics* 6, pp. 21–37.
- Amemiya, T. and T.E. MaCurdy. (1986). "Instrumental-variable estimation of an error-components model." *Econometrica* 54, pp. 869–880.
- Battese, G.E. and T.J. Coelli (1988). "Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data." *Journal of Econometrics* 38, pp. 387–399.
- Battese, G.E. and T.J. Coelli. (1991). "Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India." Department of Econometrics, University of New England, Australia.
- Bjurek, H., L. Hjalmarsson, and F.R. Forsund. (1990). "Parametric and nonparametric estimation of efficiency in service production: a comparison." Frontier analysis, parametric and non parametric approaches, edited by A.Y. Lewin and C.A. Knox Lovell in *Journal of Econometrics Annals 1990-4*, 46-1/2, pp. 213–228.

- Breusch, T.S., G.E. Mizon, and P. Schmidt. (1989). "Efficient estimation using panel data." *Econometrica* 57, pp. 695-700.
- Caves, D.W., L.R. Christensen, and J.A. Swanson. (1981). "Productivity growth, scale economies and capacity utilization in U.S. railroads, 1955-1974." *The American Economic Review* 71, pp. 994-1002.
- Coelli, T.J. (1989). "Estimation of frontier production functions: A guide to the computer program "Frontier"" Working Papers in Econometrics and Applied Statistics, University of New England, Australia, 34..
- Cornwell, C.P., P. Schmidt, and R.C. Sickles. (1990). "Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels." Frontier analysis, parametric and non parametric approaches, edited by A.Y. Lewin and C.A. Knox Lovell in *Journal of Econometrics Annals* 1990-4, 46-1/2, pp. 185-200.
- Deprins, D. and L. Simar. (1988). "Mesure d'efficacité des réseaux de chemins de fer." In CORE (ed.) *Gestion de l'économie et de l'entreprise: l'approche quantitative*, Bruxelles, DeBoek-Wesmael, pp. 321-344.
- Diewert, W.E. (1973). "Functional forms for profit and transformation functions." *Journal of Economic Theory* 6, pp. 284-316.
- Diewert, W.E. (1974). "Functional forms for revenue and factor requirements functions." *International Economic Review* 15-1, pp. 119-130.
- Gathon, H.J. (1991). "La performance des chemins de fer européens: Gestion et autonomie." Ph.D. thesis, Department of Economics, University of Liège.
- Greene, W.H. (1980). "Maximum likelihood estimation of econometric frontier functions." *Journal of Econometrics* 13, pp. 27-56.
- Hausman, J.A. and W.E. Taylor. (1981). "Panel data and unobservable individual effects." *Econometrica* 49, pp. 1377-1398.
- Jondrow, J., C.A.K. Lovell, I.S. Materov, and P. Schmidt. (1982). "On the estimation of technical inefficiency in the stochastic frontier production function model." *Journal of Econometrics* 19, pp. 233-238.
- Klein, L.R. (1953). *A textbook of econometrics*. New York: Row Peterson.
- Kumbhakar, S.C. (1988). "Production frontiers, panel data and time-varying technical inefficiency." Frontier analysis, parametric and non parametric approaches, edited by A.Y. Lewin and C.A. Knox Lovell in *Journal of Econometrics Annals* 1990-4, 46-1/2, pp. 201-212.
- Meeusen, W. and J. van den Broeck (1977). "Efficiency estimation from Cobb-Douglas production functions with composed error." *International Economic Review* 18, pp. 435-44.
- Mundlak, Y. (1961). "Empirical production function free of management bias." *Journal of Farm Economics* 43, pp. 44-56.
- Mundlak, Y. (1978). "On the pooling of time-series and cross-section data." *Econometrica* 49, pp. 69-86.
- Nash, C. (1985). "European railways comparisons—What can we learn?" In K.J. Bulton and D.E. Pitfield (eds.) *International Railway Economics*. London: Gower House.
- Nishimizu, M. and J.M. Page. (1982). "Total factor productivity growth, technological progress, and technical efficiency change: dimensions of productivity change in Yugoslavia, 1967-1978." *Economic Journal* 92, pp. 920-936.
- Perelman, S. and P. Pestieau. (1988). "Technical performance in public enterprises, a comparative study of railways and postal services." *European Economic Review* 32, pp. 432-441.
- Schmidt, P. and R.C. Sickles. (1984). "Production frontiers and panel data." *Journal of Business and Economics Statistics* 2, pp. 367-374.
- UIC (1961-1988). *International Railway Statistics*. Statistics of individual railways, Paris.
- Wallace, T.D. and A. Hussein. (1969). "The use of error components models in combining cross-section with time-series data." *Econometrica* 37, pp. 55-72.
- Wang Chiang, S.J. and A.F. Friedlander. (1984). "Output aggregation, network effects, and the measurement of trucking technology." *Review of Economics and Statistics* 20, pp. 267-276.