



HAL
open science

Efficiency Evaluation in the Regional Railway Sector: Evidence from a Shift to Line-Level Analysis

Nicolas Fabre, Anthony Martin, Olivier Barlogis, Florent Laroche, Louafi
Bouzouina

► **To cite this version:**

Nicolas Fabre, Anthony Martin, Olivier Barlogis, Florent Laroche, Louafi Bouzouina. Efficiency Evaluation in the Regional Railway Sector: Evidence from a Shift to Line-Level Analysis. 2026. <hal-05591282>

HAL Id: hal-05591282

<https://hal.science/hal-05591282v1>

Preprint submitted on 21 Apr 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License



LABORATOIRE
AMÉNAGEMENT
ÉCONOMIE
TRANSPORTS

TRANSPORT
URBAN PLANNING
ECONOMICS
LABORATORY

WORKING PAPERS DU LAET

NUMÉRO
2026/01

Efficiency Evaluation in the Regional Railway Sector: Evidence from a Shift to Line-Level Analysis

**Nicolas FABRE
Anthony MARTIN
Olivier BARLOGIS
Florent LAROCHE
Louafi BOUZOUINA**

As competition enters the French regional railway sector, assessing how public service operators achieve efficiency has become both a regulatory and strategic priority. Traditional evaluations, often conducted at national or regional levels, tend to overlook the diversity of operating conditions within regions. This paper introduces a novel analytical scale for efficiency evaluation, the line level, which enables genuinely like-for-like comparisons across services that share similar operational characteristics. This is possible due to the existence of accountability to a line-level. Using a unique dataset covering 147 French regional lines between 2015 and 2019, we apply Stochastic Frontier Analysis (SFA) to measure operational efficiency based on cost drivers such as rolling stock, staff, and energy. The results reveal substantial heterogeneity both across and within regions, demonstrating that aggregated indicators conceal meaningful local variations. Beyond its empirical findings, this paper proposes a replicable and policy-oriented framework for evidence-based evaluation at the operational scale, supporting ex-ante, in-itinere, and ex-post assessments of regional rail performance in a liberalised environment.

Keywords: Efficiency Analysis, Regional railway, SFA, Evaluation, Competition

J.E.L. Classification: L91, L92, L32, D24



LABORATOIRE
AMÉNAGEMENT
ÉCONOMIE
TRANSPORTS

TRANSPORT
URBAN PLANNING
ECONOMICS
LABORATORY

NUMÉRO
2026/01

Efficiency Evaluation in the Regional Railway Sector: Evidence from a Shift to Line-Level Analysis

Nicolas FABRE

Chambre régionale des comptes Auvergne-Rhône-Alpes, France
Univ Lyon, ENTPE, Université Lyon 2, LAET, F-69007, LYON, France

Anthony MARTIN

Autorité de régulation des transports

Olivier BARLOGIS

Chambre régionale des comptes Auvergne-Rhône-Alpes, France

Florent LAROCHE

Univ Lyon, Université Lyon 2, LAET, F-69007, LYON, France

Louafi BOUZOUINA

Univ Lyon, ENTPE, LAET, F-69120, VAULX-EN-VELIN, France

Avril 2026

ISSN : 2741-8103

Laboratoire Aménagement Économie Transports
MSH Lyon St-Etienne
14, Avenue Berthelot
F-69363 Lyon Cedex 07 France

Avertissement

Les Working Papers du LAET n'ont pas vocation à être une revue. En conséquent, ils ne sont pas dotés d'un comité éditorial et les propos n'engagent que leur(s) auteur(s) avec ou sans review.

Sans review

Ce WP n'a pas fait l'objet d'une review par ses pairs. Les propos n'engagent que son ou ses auteur(s).

Avec review

Ce WP a fait l'objet d'une review par ses pairs en guise d'amélioration du contenu et non de contrôle éditorial. Les propos n'engagent que son ou ses auteur(s).

Efficiency Evaluation in the Regional Railway Sector: Evidence from a Shift to Line-Level Analysis

Abstract

As competition enters the French regional railway sector, assessing how public service operators achieve efficiency has become both a regulatory and strategic priority. Traditional evaluations, often conducted at national or regional levels, tend to overlook the diversity of operating conditions within regions. This paper introduces a novel analytical scale for efficiency evaluation, the line level, which enables genuinely like-for-like comparisons across services that share similar operational characteristics. This is possible due to the existence of accountability to a line-level. Using a unique dataset covering 147 French regional lines between 2015 and 2019, we apply Stochastic Frontier Analysis (SFA) to measure operational efficiency based on cost drivers such as rolling stock, staff, and energy. The results reveal substantial heterogeneity both across and within regions, demonstrating that aggregated indicators conceal meaningful local variations. Beyond its empirical findings, this paper proposes a replicable and policy-oriented framework for evidence-based evaluation at the operational scale, supporting ex-ante, in-itinere, and ex-post assessments of regional rail performance in a liberalised environment.

Key words

Efficiency Analysis, Regional railway, SFA, Evaluation, Competition

JEL classification

L91, L92, L32, D24

1 Introduction

The Fourth European Railway Package completes the long-running process of rail liberalisation and was adopted in 2016, fifteen years after the first package. From 2019 onward, all Public Service Obligation (PSO) services can be awarded through a competitive tendering process. As of 2023, this becomes the default rule: PSO contracts must be allocated following a competitive procedure.

France is no exception, and competition has already been introduced in several PSO service groups. In France, with the exception of the Paris region, eleven regions act as Public Transport Authorities (PTAs). They define the service i.e., supply levels, investment programmes, and the transport plan. After the tendering process, they also select the operator that will deliver the service and they balance the cost of provision via a subsidy, which is systematically the case under cost-plus arrangements as they currently stand (Desmaris and Monchambert, 2023). In France, as in other European countries, competition for the market (Demsetz, 1968) in railways entails significant changes for the industry. According to the French Court of Auditors (2024), the eleven regional markets will be split into at least forty-five tenders. Until 2023, the average regional market size in France amounted to around 17 million train-kilometres and 1,870 million passenger-kilometres. The breaking up of these regional markets will in effect produce intra-regional markets that are, on average, about 4.6 million train-kilometres and 374 million passenger-kilometres. Consequently, the inherent networked and reticular nature of rail operations is being tested. This is particularly salient given Preston and Robins (2013) who, two decades after the UK railway reforms, concluded that "the privatisation package has been welfare negative and that the most likely cause has been the complex and fragmented nature of the supply-side arrangements".

A second implication of introducing competition in PSO services follows directly from the possibility (and the obligation) for PTAs to choose an operator. PTAs can no longer rely indefinitely on the incumbent;

instead, they must design tenders and organise a more fragmented production system while targeting cost savings and quality-of-service improvements. The incumbent becomes one operator among others and may not share all the acquired expertise out of the scope of the data required for the tendering procedures, compelling PTAs to develop stronger administrative and technical capabilities to govern the service. When there was no choice of contractor, PTAs could readily delegate large segments of production to the operator. In the new setting, however, ensuring an effective regulatory grip becomes a central task if PTAs are to deliver on the promise associated with competition for the market: reducing costs and improving service quality (i.e., achieving greater efficiency).

These two shifts -industrial reorganisation and heightened regulatory requirements- call for new methodologies to assess the efficiency of the railway system. This paper responds to that need by proposing a new framework that mobilises well-established econometric tools to analyse efficiency in a context of emerging competition.

Efficiency analysis is required for at least three complementary evaluative purposes. An ex-ante efficiency appraisal is needed to assess how a system transforms costs into outputs before competition takes place. Such analysis helps establish a factual baseline and informs the design of tenders and the selection of operators. However, horizontal separation reduces the relevance of highly aggregated (national or regional) analyses. With intra-regional unbundling, evaluators need to quantify the costs associated with specific tenders and to identify their cost drivers. As the system becomes less uniformly reticular, the unit of analysis should be updated, ideally to the finest operational granularity: the line level.

An in-itinere perspective is equally essential once contracts are in operation. Monitoring efficiency trajectories over time allows PTAs to detect early deviations from expected performance, understand the impact of operational shocks, and foster adaptive regulation. An ex-post evaluation is required to assess the benefits, or lack thereof, of competition once contracts mature, and to identify transferable lessons for future tendering cycles. Naturally, because competition is only now emerging in the French regional railway sector, our empirical focus is on the ex-ante dimension. Our research question is therefore how to assess the efficiency of a future fragmented railway system by proposing a new scope of analysis at the line level.

To address this question, we rely on a unique line-level dataset covering French regional railway lines from 2015 to 2019. Working at the line rather than the regional level enables genuine like-for-like comparisons that aggregated analyses cannot deliver, though they are informative (Desmaris and Monchambert, 2023; Link, 2016, 2019; Cantos et al., 2012; Holvad, 2020). Regional indicators inevitably blend very different service profiles with busy suburban corridors and sparsely served rural lines, thus conflating structural heterogeneity with performance outcomes. By contrast, a line-level approach allows for the comparison of truly similar services across different regions (e.g., high-demand commuter lines or low-frequency rural branches) without mixing them with dissimilar ones. This mitigates aggregation bias, enhances the interpretability of efficiency differentials, and increases the diagnostic power of benchmarking: when two lines operate under comparable conditions but exhibit different costs or outputs, the observed gap is more plausibly attributable to managerial practices, contractual arrangements, or technology rather than to compositional effects. This shift in the unit of analysis from regions to lines, is the core motivation of this paper: to ensure genuinely like-for-like comparisons using a dataset of 147 lines for 5 years, which is substantially larger than in previous line-level studies (Laroche and Lamatkhanova, 2022a,b) or in a different context (Le et al., 2022). Our database is derived directly from operational records, which we view as a strength, since the proposed methodology can be easily replicated by public transport authorities (PTAs) and operators. We address the research question using a well-established tool in the railway efficiency literature, Stochastic Frontier Analysis (SFA), for several reasons.

First, SFA provides transparent and interpretable outcomes in the form of efficiency scores, offering a common and intuitive language for dialogue between PTAs and operators. Second, it flexibly accommodates multiple inputs and outputs, allowing us to explicitly control for scale and production factors such as seat-kilometres, staff costs, energy, or rolling stock. Unlike simple cost-to-output ratios, SFA disentangles inefficiency from random shocks or measurement errors, thereby yielding a more robust assessment of performance. We estimate the production frontier using a Cobb-Douglas specification. This functional form provides a parsimonious and easily interpretable benchmark, directly comparable to previous studies on French railways (e.g., Desmaris and Monchambert, 2023). This specification was retained for the analysis due to its robustness and the clarity it provides on efficiency patterns, which is well-suited for this policy-oriented study.

Our primary contribution lies in incorporating intra-regional heterogeneity by shifting the unit of analysis to the line level, whereas most previous studies have focused on inter-regional or cross-national comparisons. The purpose of this paper is not to exhaustively identify all determinants of efficiency. Rather, we provide a diagnostic and exploratory contribution, demonstrating how existing information systems can support a replicable and policy-oriented framework for performance evaluation. Empirically, our results reveal substantial performance variation both across and within regions, demonstrating that aggregated indicators conceal meaningful local variations. We find that even after segmenting lines by production intensity and controlling for regional effects, performance among similar lines varies widely. Furthermore, our analysis of the 2015–2019 period shows divergent dynamic trajectories: 62 lines improved their efficiency while 86 experienced deterioration. These findings have direct policy implications. This need for granular evaluation is critical with the progressive opening of the French regional railway market to competition. Our framework offers practitioners a roadmap to support ex-ante, in-itinere, and ex-post evaluation of service performance in this new environment. This analysis also points to the need for future research to investigate the specific line-level factors (managerial, contractual, or technological) driving the residual variations we identify.

The remainder of this paper is structured as follows. Section 2 reviews the relevant literature. Section 3 details the line-level dataset and the Stochastic Frontier Analysis methodology. Section 4 presents and discusses the empirical results. Finally, Section 5 concludes with policy implications.

2 Background

We review empirical work on railway efficiency with an emphasis on Europe. Three broad strands can be distinguished. A first strand relies on international datasets to study links between sectoral reforms and efficiency. These papers typically compare firms at the company level to enable cross-country benchmarking, with a focus on vertical separation, financial delineation, and competition effects (Holvad, 2020). Covariates are often macro or meso indicators such as population density or GDP per capita, aggregated at national or regional scales. Results commonly point to positive associations between economic development, traffic density, or electrification and measured efficiency (e.g. Lan and Lin, 2003; Benga et al., 2025); to gains associated with separation when paired with complementary reforms (Asmild et al., 2009; Cantos et al., 2012); and to modest or mixed links between liberalisation and efficiency unless effective competitive pressure materialises (Lerida-Navarro et al., 2019; Friebel et al., 2010). More recently, Benga et al. (2025) show with a Data Envelopment Analysis that a nation’s wealth, length of haul and trip, as well as traffic density, have a positive effect on efficiency. A recurring limitation in this strand is the heterogeneity of definitions and data across countries, which complicates comparisons. Crucially, national-level aggregation often blends fundamentally different subsystems such as intercity and high-speed services alongside regional and commuter operations, thereby mixing distinct production technologies, demand profiles, and regulatory environments within a single composite outcome. Some studies find that transactional and organisational factors may outweigh purely structural or competitive variables (Merkert et al., 2010), while scale and labour use remain persistent determinants (Couto and Graham, 2009).

A second strand turns, more recently, to the regional level within a single country, where contractual arrangements, procurement choices, and governance can be observed more closely. Evidence shows that tendered services, gross-cost contracts, and appropriate contract durations are frequently associated with higher efficiency, and that incorporating quality-of-service indicators can reorder performance rankings of Public Transport Authorities (PTAs) (Link, 2016, 2019). Other contributions underline the role of operator autonomy, governance, and socio-demographic context such as population density, delinquency rate in driving efficiency (Desmaris and Monchambert, 2023). This direction connects with earlier observations on the sensitivity of efficiency and cost results to how heterogeneity is modelled and at what level (Wheat and Smith, 2015). Yet, even here, the empirical unit is most often the network, the operator, the country, or the region. As a result, indicators still average over dissimilar services; for example, regional aggregates can combine dense suburban corridors with lightly used rural branches, while national aggregates mix regional and commuter services with intercity and high-speed operations.

Against this background, line-level efficiency studies are rare. Beyond Le et al. (2022) on Japanese lines - which show, among other results, that congestion may depress cost efficiency while raising revenue efficiency - disaggregated applications at the corridor or line scale remain the exception. This scarcity matters. Aggregation to regional or national levels raises the risk of ecological fallacies and blurs the

operational levers available to PTAs and operators. By contrast, line-level data capture corridor-specific demand, geometry, and operating conditions, enabling like-for-like benchmarking across territories. Related research on open access, modal competition, and contracting, although not always framed as efficiency analysis, also indicates that the corridor or line is the relevant locus at which procurement choices, negotiation outcomes, and market structure translate into costs and service design (Laroche and Lamatkhanova, 2022a,b; Blayac et al., 2024).

Overall, the literature has delivered important insights on reform, contracting, and context. However, the dominant empirical lens remains national or, at best, regional, which blends multiple rail subsystems and masks within-region heterogeneity. This leaves space for line-level analyses that treat intra-regional variation as first-order. Our study occupies this space by using a country-wide, line-level dataset to provide cleaner comparisons between similar services across regions and a tighter mapping from institutional and operational choices to measured efficiency.

3 Data

Our dataset comprises detailed accounting information for each regional train line across all French regions, with the exception of Île-de-France (Paris) and Corsica, whose organisational arrangements are distinctive. Compiled by SNCF Voyageurs (the French national railway company), the data record train production and associated costs at the line level. The statistical unit is a line-year; the resulting panel contains 735 observations.

Identifiers and line definitions

Each observation carries a train-line identifier (ID) defined by its origin and destination (e.g., LYON - MARSEILLE). Intermediate stops are not enumerated. It is important to note that line definitions may change over time: new stops can be added or removed; some lines can be split, partially reconfigured, or merged into new entities. In several cases, only segments of existing lines are recombined, so the resulting line is not a simple union of predecessors. Because the dataset does not include geospatial coordinates (e.g., lat/long for endpoints), reconstructing precise longitudinal line identities and implementing a defensible historical merge would require additional assumptions. To avoid introducing classification error, we refrain from engineering ex-post merges and treat line-years as observed. In addition, some lines are grouped by operators or PTAs into "étoiles ferroviaires" (hub-and-spoke clusters around urban nodes, e.g., the Reims star), or into thematic bundles such as "urban" or "commuter" lines. To ensure consistency in the estimation of the frontier and to track efficiency changes over the period, we retained only the lines present for all five years (2015–2019). This filtering process results in a balanced panel of 147 lines, corresponding to the 735 line-year observations used in the analysis.

These definitional challenges reflect the institutional reality that line-level information is produced primarily for regional PTAs, each of which may define and bundle lines differently in response to local demand and operational optimisation. Although this heterogeneity complicates harmonisation, it also underscores the operational relevance of the dataset: such accounting artefacts are precisely those that PTAs use to structure tenders and steer services. Working with an operational dataset is therefore, in our view, substantively appropriate for an efficiency study aimed at informing procurement and oversight. However, these data limitations challenge the use of panel data approach, with some observations not being present in the entire panel. We think it is an acceptable risk to lose some observations but keeping the panel structure of the data.

Production metrics (outputs)

We observe standard railway output measures at the line level, notably train-km and seat-km. Unfortunately, passenger-km are unavailable in our dataset, hindering a demand-oriented efficiency analysis. Given this limitation, our empirical approach is supply-oriented: we assess how costs translate into supplied capacity and service production, not into realised passenger usage. Therefore, and since this data is available, we prefer to use seat-km as output, given that seat-km reflect both the number of trains and their seating capacity. This variable primarily measures capacity supplied. Higher seat-km per train can indirectly improve perceived quality, by reducing crowding and increasing the probability of finding a seat, even though seat-km is not a direct quality-of-service indicator. Seat-km is an output used in such supply-oriented analysis (Catalano et al., 2019; Holvad, 2020)

Accounting variables (inputs)

Line-level accounts report various costs. Some categories are initially recorded at the regional level (e.g., ticketing platforms, administrative overheads, marketing) and then allocated to lines via distribution keys, whereas others are directly attributable to a given line, such as energy, rolling-stock maintenance, and on-board staff (inspectors and drivers). To minimise allocation noise, our baseline analysis retains only these three directly attributable cost items. We emphasise that the variables used as inputs are expenditures rather than physical quantities or input prices. In the absence of systematic information on input prices, we treat these expenditures as proxies for input quantities. This relies on the institutional context: the same incumbent operates all regional services, wage scales for operational and maintenance staff are nationally standardised, and major maintenance cost components are subject to national regulatory oversight (Avis 2023-063, 2023). Under these conditions, cross-regional variation in input prices should be limited, so that differences in line-level expenditures primarily reflect differences in the intensity of resource use rather than price heterogeneity. Nonetheless, this remains an assumption and should be borne in mind when interpreting the estimated frontier. To ensure analytical rigor, all financial values were corrected for inflation by standardizing them to 2015 euros using price indices provided by INSEE (the French National Institute of Statistics and Economic Studies). This adjustment facilitates meaningful temporal comparisons over the study period.

Missing variables

As an operational dataset, the file does not contain several desirable covariates such as service-quality indicators, number of employees, detailed rolling-stock inventories, or typologies of lines (rural, urban, mountainous etc.).

Summary Statistics

Table 1 summarizes the descriptive statistics for the key variables used in the frontier model. The statistical unit used in our model is a line for a given year. Consequently, the statistics in Table 1 are calculated on the basis of 147 lines for 5 years.

Table 1: Descriptive statistics of inputs (in millions) and output

Variable	Mean	Median	SD	Min	Max	N
Seat-km (output) in millions	244.5	121.9	350.4	1.1	3,464.0	735
Rolling Stock (input) in M€	4.653	3.451	4.762	0.069	42.358	735
Staff (input) in M€	4.396	3.305	3.897	0.055	25.348	735
Energy (input) in M€	0.709	0.525	0.729	0.007	5.435	735
Line length (z-variable) in km	136.0	114.6	102.2	10.6	530.5	735

The variation observed across all variables, as indicated by the large standard deviations relative to the means and the significant gaps between the minimum and maximum values, highlights the diverse nature of railway operations within our sample. This heterogeneity highlights the methodological advantage of employing a line-level analysis, which is well-suited to capture the nuanced differences across the French railway network.

4 Methodology

Stochastic Frontier Analysis

Stochastic Frontier Analysis (SFA) is employed to estimate the operational efficiency of railway lines by modeling a production frontier that incorporates both random noise and inefficiency components (Holvad, 2020; Battese and Coelli, 1995). This methodological approach enables us to disentangle measurement errors from genuine inefficiency effects and to quantify the deviation of each observation from the best-practice frontier. The primary objective of our analysis is to measure railway operational efficiency at the line level, thereby capturing both intra-regional and inter-regional variations.

Data and Temporal Considerations

Our dataset spans the period from 2015 to 2019. Due to the dynamic nature of the railway network, where lines may merge, split, be closed, or newly created, the composition of the dataset varies over time.

To appropriately account for these temporal effects, we incorporate a time trend into the SFA model. The year variable is scaled to capture potential technological advancements and shifts in the production frontier throughout the analysis period.

Model Specification

We estimate a stochastic frontier model of train service production, using a Cobb–Douglas functional form. Following [Desmaris and Monchambert \(2023\)](#), we model the supply frontier at the line level, relating supply-oriented output (seat-km) to a set of input cost drivers such as staff, energy and rolling stock.

The Cobb–Douglas form is chosen as a baseline specification due to its parsimony and interpretability. It allows for straightforward elasticity interpretation and facilitates comparison with previous studies that used similar data and frameworks. Although it imposes constant output elasticities and a unitary elasticity of substitution among inputs, it provides a convenient benchmark for policy-oriented analysis.

The general Cobb–Douglas production frontier is defined as:

$$Y_{it} = \exp(\beta_0) \times \prod_{k=1}^K (X_{kit})^{\beta_k} \times \exp(v_{it} - u_{it}) \quad (1)$$

where:

- Y_{it} : Output (seat-km) produced by line i in period t ,
- X_{kit} : Inputs k (staff, energy, rolling stock),
- $v_{it} \sim N(0, \sigma_v^2)$: Random noise term capturing measurement errors and exogenous shocks,
- $u_{it} \geq 0$: Non-negative inefficiency term capturing the shortfall from the frontier.

The log-linearized form used for estimation is:

$$\ln(Y_{it}) = \beta_0 + \sum_{k=1}^K \beta_k \ln(X_{kit}) + v_{it} - u_{it} \quad (2)$$

In the efficiency effects model proposed by [Battese and Coelli \(1995\)](#), inefficiency is modeled as a function of contextual variables:

$$u_{it} = z_{it}\delta + w_{it}, \quad w_{it} \sim N^+(0, \sigma_u^2) \quad (3)$$

where z_{it} represents a vector of inefficiency determinants (such as regional dummies or environmental characteristics) and δ is the associated parameter vector.

The technical efficiency (TE) for each line and year is then obtained as:

$$\text{TE}_{it} = \exp(-u_{it}) = \frac{Y_{it}}{f(X_{it}; \beta) \exp(v_{it})} \quad (4)$$

Efficiency scores range from 0 to 1, with 1 indicating full efficiency. For instance, a TE of 0.85 indicates that a line operates at 85% of its potential output given its inputs.

Inefficiency effects are modeled following [Battese and Coelli \(1995\)](#), with regional dummies and line length as determinants of efficiency.

Endogeneity Considerations

A potential concern when using operational data is the simultaneity between inputs and output: in many industries, higher production may mechanically induce higher expenditures, making it difficult to distinguish technical inefficiency from the joint determination of costs and output. In the context of French regional railways, this issue is substantially mitigated by the institutional organisation of production. Service levels are defined ex-ante by PTAs through multiyear public service contracts. The operator does not freely choose the amount of output to supply; instead, it allocates staff, rolling stock and energy to deliver a predetermined service plan. As a result, observed input expenditures primarily reflect the resources required to comply with contractual obligations, rather than an endogenous response to output variations. This institutional setting does not eliminate all potential sources of simultaneity

(operational adjustments, disruptions or unforeseen shocks may still affect both inputs and output within a year) but it substantially limits the scope for strong endogeneity biases. Consequently, the relationship estimated in the stochastic frontier should be interpreted as a reduced-form benchmarking tool aimed at assessing how efficiently lines deploy resources to produce the mandated service, rather than as a structural causal production function.

Including the Regions in the Model: Technological or Inefficiency Gap?

In our dataset, each line is associated with a specific region, even in the case of cross-regional services. Hence, heterogeneity across lines also reflects regional policy differences, as highlighted by previous research conducted at the regional level (Desmaris and Monchambert, 2023; Link, 2016, 2019).

From a modelling perspective, the key question is how to represent these regional effects within the stochastic frontier framework. Should regions be considered part of the production technology - implying structural differences in the production process - or should they be treated as determinants of inefficiency? In other words, if regions differ in terms of production technology or productivity, they should enter the production function directly. Conversely, if some regions achieve higher efficiency levels for a given set of inputs, perhaps due to better contract management, governance, or resource allocation, then regional affiliation should be introduced as a z -variable explaining inefficiency.

This distinction is not trivial and has rarely been addressed in railway efficiency studies. Most existing works have compared countries or regions as single entities, rather than individual lines, and thus did not face the same identification problem. The same issue will arise in future studies comparing multiple operators: should operator identity be modeled as part of the production function (reflecting different technologies) or as a determinant of efficiency (reflecting managerial or organizational differences)?

In our specific case, all lines included in the panel are operated by the same incumbent operator, SNCF. Therefore, it is reasonable to assume that the production technology is homogeneous across regions i.e., all lines share the same production frontier. The remaining differences among regions are more likely to stem from allocative or managerial inefficiencies rather than technological gaps. For this reason, we treat regions as z -variables in the inefficiency equation, rather than as shifters of the production function.

Finally, it should be noted that our current dataset does not include detailed environmental or contextual variables. Consequently, the estimated regional effects may capture a mix of unobserved PTA-related characteristics and operator-specific management differences across regions. Future research could disentangle these two dimensions by introducing additional variables describing contract design, infrastructure conditions, or governance quality.

5 Results and discussions

Table 2 reports the estimation results for two sfa models. Two approaches have been applied using two specifications: Cobb-Douglas without z -variables (following Battese and Coelli, 1992) and a second Cobb-Douglas with z -variables (following Battese and Coelli, 1995). Both models include a time trend, that captures technological changes (Henningsen, 2019).

The estimated coefficients for the cost inputs are broadly consistent with those reported in previous studies using regionally aggregated datasets (Desmaris and Monchambert, 2023), thereby reinforcing the validity of our line-level approach. While Desmaris and Monchambert (2023) employed the number of rolling stock units as an input, our analysis uses cost variables, since line-level data on the number of vehicles are not available. The relative alignment of coefficients suggests that, under a single operator, rolling stock costs are highly correlated with fleet size and can thus serve as a reasonable proxy. Because the database covers only one incumbent operator across different regions, variations in input costs are more likely to reflect operational intensity rather than price heterogeneity. This consistency would be less guaranteed in a multi-operator context, where cost structures and accounting standards may differ substantially.

Likelihood ratio tests were performed, clearly rejecting the OLS specification in favour of the stochastic frontier model. Because Model 1 and Model 2 are non-nested, their log-likelihoods cannot be directly compared. We therefore select Model 2 as the preferred specification based on both theoretical and empirical considerations. From a theoretical standpoint, both models allow inefficiency to vary over time through the Battese and Coelli time-decay structure Battese and Coelli (1992, 1995). The key difference is

Table 2: Comparison of SFA models

Variable	Model (1)		Model (2)	
	Estimate	Std. Error	Estimate	Std. Error
Frontier Variables				
Constant	4.4281***	0.3674	1.2987***	0.3238
<i>log(Rolling stocks)</i>	0.1945***	0.0251	0.5099***	0.0353
<i>log(Staff)</i>	0.5646***	0.0360	0.6283***	0.0459
<i>log(Energy)</i>	0.2887***	0.0249	0.0644	0.0394
<i>technological change (time trend)</i>	0.0232***	0.0069	-0.0162	0.0108
Inefficiency effects (Z-variables)				
Constant (Z)	—	—	0.5098	0.2885
Region B (rel. to A)	—	—	-0.6833	0.3882
Region C (rel. to A)	—	—	-0.7572	0.3987
Region D (rel. to A)	—	—	-0.0056	0.2401
Region E (rel. to A)	—	—	-0.3948	0.2115
Region F (rel. to A)	—	—	-0.5579*	0.2270
Region G (rel. to A)	—	—	-0.1266	0.1882
Region H (rel. to A)	—	—	0.0889	0.1911
Region I (rel. to A)	—	—	0.2216	0.1744
Region J (rel. to A)	—	—	-0.8024*	0.4044
Region K (rel. to A)	—	—	0.2972	0.1830
Line length	—	—	-0.0000	0.0004
Variance & Decay Parameters				
σ^2	1.1842***	0.1618	0.2516***	0.0411
γ	0.9872***	0.0020	0.6614***	0.0927
η (time decay)	-0.0369***	0.0068	—	—
Model summary				
Log-likelihood	165.326		-388.495	
Cross-sections	147		147	
Time periods	5		5	
Observations	735		735	
Mean efficiency	0.466		0.656	

Notes: Model (1) is a Battese & Coelli (1992) Error Components Frontier with time decay; Z-variables are not specified. Model (2) is a Battese & Coelli (1995) Inefficiency Effects Frontier using $\log(\text{seat-km})$ as output. Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, $p < 0.1$.

that Model 2 introduces regional z-variables in the inefficiency equation, making it possible to distinguish systematic cross-sectional heterogeneity from the inefficiency term. Model 1, by contrast, attributes all unobserved cross-sectional variation to inefficiency, which can inflate the γ parameter and obscure the distinction between structural differences and genuine inefficiency. By modelling regional effects explicitly, Model 2 provides a clearer and more meaningful decomposition of the sources of performance variation. The empirical results reinforce this choice. Model 1 produces an extremely high gamma ($\gamma = 0.987$), implying that virtually all residual variation is attributed to inefficiency. This is consistent with the absence of controls for regional or environmental differences, and likely results in an artificially low mean efficiency (0.466). By contrast, Model 2 yields a more moderate gamma ($\gamma = 0.661$) and a higher, more plausible mean efficiency (0.656), reflecting the fact that some cross-regional variation is now correctly captured by the Z-variables rather than being misinterpreted as inefficiency.

In Model 2, rolling stock and staff appear as positive and significant drivers of output, while energy shows a positive but non-significant effect ($p=0.10$). Evaluated at the mean, the sum of estimated elasticities is approximately 1.20, which implies increasing returns to scale at the line level. This suggests that doubling all inputs would lead to roughly 20% more output than proportional input-costs increases. Line length, included as an inefficiency determinant, has no statistically significant effect. It is not the length of a line that influences efficiency, but its volume of operation (i.e., density), which is captured by the increasing returns to scale. However, a lack of additional variables makes it difficult to draw definitive conclusions about scale and density while not taking heterogeneity fully into account (Wheat and Smith, 2015). The inefficiency equation shows that several regional dummies have negative and significant coefficients. A negative coefficient indicates a reduction in the inefficiency term, implying that these regions are structurally more efficient and operate systematically closer to the production frontier than the reference region. This pattern strongly suggests that unobserved factors, such as superior contract management or infrastructure quality, have a tangible impact on performance, even under the same incumbent operator. Finally, the estimated technological change (time trend) in the frontier is negative but not significant. It does not suggest a statistically significant decline in productivity over the 2015–2019 period on average.

Differences Among and Within Regions and Among Same Type of Line

Efficiency scores have been calculated using Shephard’s output-oriented technical efficiency methodology, with values ranging between 0 and 1. A score of 0.65 means that, given its current level of inputs, the line produces 65% of the output it could achieve if it were operating on the efficient frontier.

Figure 1 presents boxplots that illustrate the distribution of efficiency scores across different years and regions.

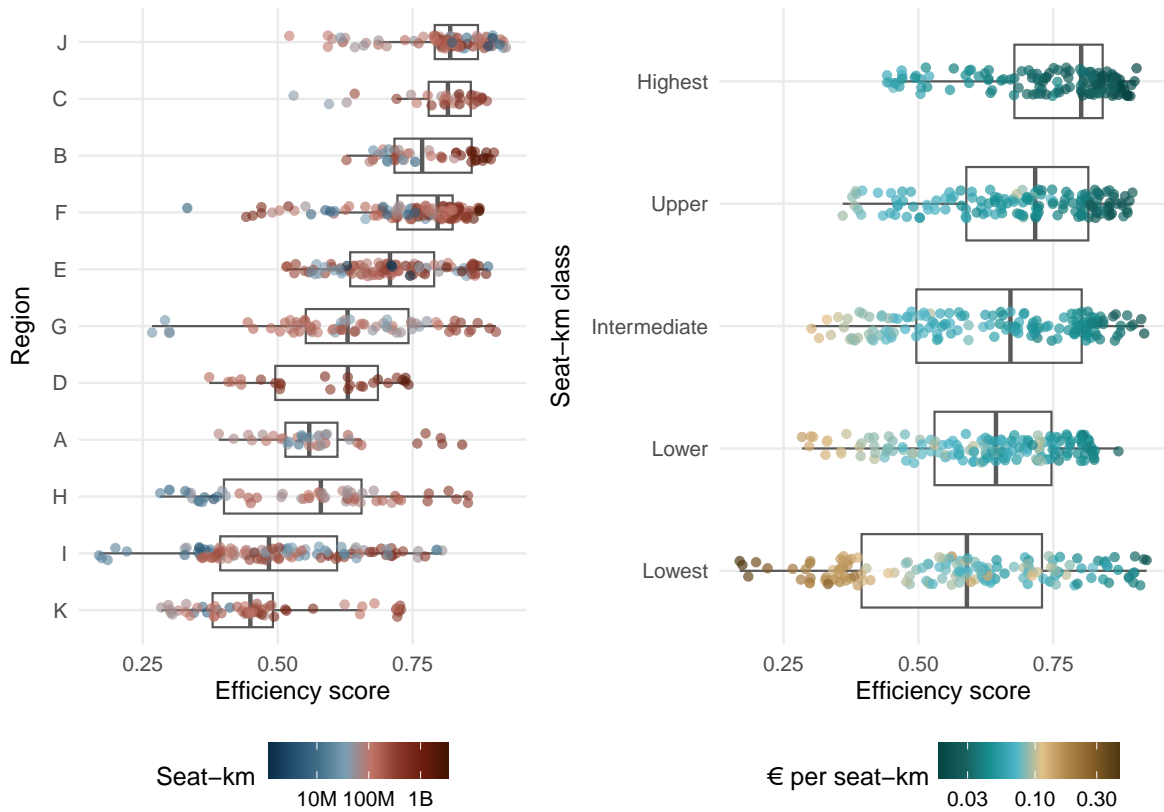


Figure 1: Efficiency Scores (M2) by Region and Intensity of Use Class (seat-km). € per seat-km is computed as the sum of energy, staff, and rolling-stock costs.

We plotted the efficiency scores against two segmentation variables: the administrative region and the seat-km class. The latter divides lines into five homogeneous groups according to their annual volume of seat-km. These two visualizations are presented in Figure 1.

The first plot highlights, as expected, substantial intra-regional heterogeneity. Prior research at the regional level (e.g., [Desmaris and Monchambert, 2023](#); [Link, 2016, 2019](#); [Affuso et al., 2009](#)) has underlined the influence of factors such as contract management, governance structures, population density, and service configuration in explaining inter-regional efficiency gaps. However, the fact that efficiency levels vary widely within the same contractual and governance framework suggests that other determinants, more operational and environmental in nature, play a non-negligible role. Except for some regions, intra-regional dispersion remains limited: when a region performs well (or poorly), most of its lines follow the same trend. This reinforces previous findings on the importance of efficient contract design and optimized network organization as drivers of overall efficiency.

The remaining variation likely stems from unobserved technical and environmental factors, such as rolling stock characteristics ([Fabre et al., 2025](#)) or meteorology. This observation raises a central question: why do similar lines operating within the same regional framework exhibit different efficiency scores? Although our dataset does not allow for a complete answer, this result demonstrates that real-world railway data still offer untapped potential for efficiency analysis at the line level.

Such analysis provides a basis for objective discussions between regional authorities and operators. When two comparable lines within the same region show divergent efficiency scores, this discrepancy deserves to be examined. Operators possess detailed knowledge of operational constraints, which can be modelled through methods such as SFA. In this sense, benchmarking efficiency at the line level can support both competitive tendering and continuous performance monitoring. Moreover, when a line changes operator, regional-level analyses are often insufficient to capture the true impact on efficiency, while line-specific comparisons before and after the change can offer valuable insights. As regions are now required to design

competitive tenders (effectively decomposing the regional network into smaller service bundles) such fine-grained evaluations will be key to assessing the effects of competition and identifying best practices. The objective is not to determine whether competition is "good" or "bad", since it is a mandated reform, but to understand why some operators or tenders perform better than others and under which operational conditions. Identifying which determinants are discretionary (i.e., under the control of the operator or the authority), can greatly enhance learning and foster performance improvements benefiting both the public transport authorities, the operators, and ultimately, passengers.

The second plot groups lines by their seat-km class, i.e., their intensity of use. Although there is a slight positive relationship between seat-km volume and efficiency, it is far from deterministic. Some low-volume lines perform well, while certain high-volume ones remain inefficient. This dispersion within classes mirrors what we observe across regions: similar lines can perform very differently even after controlling for regional and governance effects. Such findings are particularly relevant given the information asymmetry between regional authorities and the incumbent: the latter has access to nationwide operational data, while each region typically knows only its own network data. Addressing this asymmetry through systematic benchmarking could improve competitive outcomes and avoid "winner's curse" situations observed in the British case (Cowie, 2009). Understanding, through objective and transparent measures, why some lines outperform others offers a promising path to enhance the effectiveness of competition and to strengthen the evidence-based management of regional rail services.

Dynamic Analysis: Which Lines Improved?

Another key perspective for dialogue between public transport authorities and operators concerns the dynamic evolution of efficiency over time. This evolution reflects both external environmental conditions (such as strikes, infrastructure works, or service disruptions) and the operator's own managerial and operational efforts. Among the lines in our sample, 59 improved their efficiency score between 2015 and 2019, while 88 experienced a deterioration. The distribution of these changes is shown in Figure 2.

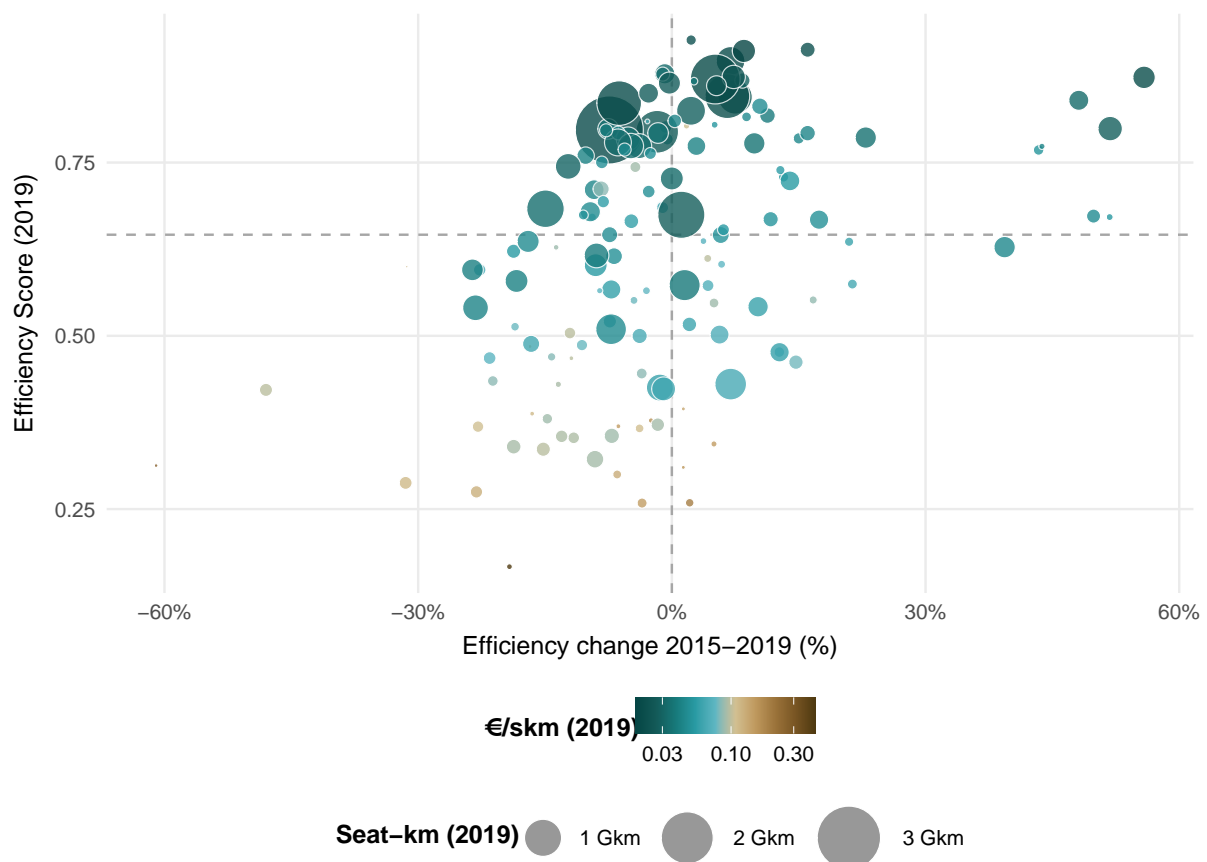


Figure 2: Efficiency scores and line-level efficiency evolution.

In this figure, the horizontal axis represents the change in efficiency between 2015 and 2019, while the vertical axis shows the 2019 efficiency level. The lines are sized by their respective seat-km for 2019 and dots are colored by the cost per skm for 2019 as well. This two-dimensional representation helps identify benchmark lines, located in the top-right quadrant. These lines not only achieve higher efficiency scores than others but also show positive progress over time. Interestingly, these benchmark lines come from different regions and exhibit diverse operating profiles: some are high-intensity lines with large train volumes, while others are smaller or less intensive lines that have nonetheless improved. Conversely, lines located in the bottom-left quadrant show both low efficiency and negative evolution. Understanding why some lines deteriorated while others improved is crucial. Although our dataset does not include all the operational or contextual variables needed for a full explanation, such analyses could be conducted by regional authorities, who have access to more detailed data. The bottom-left quadrant is predominantly composed of low-intensity lines with high costs per seat-km. This is not really surprising, since railway industry benefits from scale efficiency.

From an operational standpoint, the analysis of cost dynamics reveals a pattern. Lines whose efficiency deteriorated tend to have experienced growth in operating costs per seat-km, particularly in staff and Rolling Stock expenses. The only component showing a decrease is energy expenditure. Conversely, lines that have improved show a decrease in cost per seat-km for every expenditures. These patterns are summarized in Table 3.

Table 3: Inputs and Output Changes (2015 vs. 2019)

Group	Change in Cost / skm (%)				Output (%)
	Staff	Rolling Stock	Energy	Total	Seat-km
Improved Efficiency	-12.4	-14.6	-11.4	-13.4	4.4
Degraded Efficiency	12.7	21.5	-1.0	15.8	1.5

These results are descriptive and should not be interpreted as evidence of a causal relationship between changes in efficiency and cost structure. Rather, they serve to illustrate how efficiency dynamics may be influenced by a range of both controllable and uncontrollable factors (captured by a regional effect in the SFA).

The purpose of this paper is to propose a framework for efficiency assessment in a competitive context, especially relevant in France, where regional rail competition is being progressively introduced. The econometric tools presented here, such as the SFA, offer regulators a way to objectify discussions with operators by disentangling what stems from the public authority, the operator, or the external environment. Future research should include additional covariates (e.g., rolling stock characteristics, maintenance intensity, type of line) to deepen these insights (Fabre et al., 2025). Ultimately, the value of this approach lies in establishing a analytical roadmap for assessing regional railway performance that helps identify best practices, inform contract design, and enhance transparency in the forthcoming competitive landscape.

A Roadmap for Practitioners and Future Research

Only the French regulatory body, the Autorité de régulation des transports (ART), currently has access to comprehensive national railway data as part of its institutional mandate. The ART has significantly enhanced transparency in the sector and promotes sunshine regulation as a promising tool for oversight and accountability. Recent reports by both the ART (Autorité de régulation des transports, 2023) and the Regional Court of Auditors (Chambre Régionale des Comptes, 2023) have incorporated line-level analyses, representing a major step forward in promoting transparency and fostering evidence-based dialogue between regulators and operators. The European Union could also benefit from such analysis. Extending this type of line-level methodology to a European scale would enable more robust cross-country comparisons and a clearer identification of country-specific effects, thereby strengthening evidence-based regulation and improving the design of incentives. Building on this perspective, we identify two key levels of analysis that could be pursued by practitioners, along with the necessary data structuring efforts required to sustain them.

The first level of analysis should be carried out by individual regions. Regional authorities have only limited access to national datasets, yet the econometric tools presented in this paper, particularly the

SFA, are specifically designed for industries such as railways. Regions should therefore examine both the determinants of efficiency among the lines within their jurisdiction, and the differences in efficiency between similar lines operating under comparable conditions. For instance, a regional authority could legitimately ask its operator why two similar lines display markedly different efficiency scores or opposite trajectories over time. Such questions, when supported by operational data shared by the operator, can help objectify the discussion and focus it on key managerial issues. While statistical analysis cannot capture all the complexity of railway operations, it can serve as a starting point for informed dialogue. For example: is a line’s decline in efficiency due to insufficient operator effort, or to underinvestment in rolling stock or infrastructure? These questions are far from trivial, especially in the context of competitive tendering. Ex ante efficiency assessments could help regions identify why certain lots or service bundles are more efficient than others, and could guide the evaluation of bids during the tendering process.

The second level of analysis should be undertaken jointly by regions, through data sharing, or by the independent regulator (ART) itself. This broader comparative approach would aim to explain efficiency differences among similar lines located in different regions. In our sample, for instance, Region J appeared to be the most efficient, and prior research has identified potential reasons for such inter-regional gaps (Desmaris and Monchambert, 2023; Link, 2016, 2019). However, these studies did not delve into the line-level, and it would be interesting to understand why a similar line in Region J is more efficient than in Region K, the less efficient one. Such inter-regional analyses to a line-level would enrich discussions between operators and public authorities and could be applied ex-ante, in-itinere, and ex-post to evaluate, for example, whether competitive tendering improves efficiency once other covariates are controlled for. These analytical perspectives require strong data governance and consistent data structures. Databases must ensure continuity over time, so that line-level accounting and production indicators remain comparable before and after competition, whether or not the same operator is in charge. Additional variables should be incorporated to enrich these datasets such as technical data on rolling stock types and characteristics, service quality indicators and broader environmental or contextual parameters.

In this study, we introduced a segmentation by seat-km class, but we acknowledge that it is insufficient to fully capture the diversity of line characteristics. Future research should move toward finer-grained segmentation, integrating non-traditional variables such as delinquency rates (Desmaris and Monchambert, 2023), intermodal competition, and socio-demographic profiles of the served population (Link, 2016; Laroche and Lamatkhanova, 2022a,b; Blayac et al., 2024). Additional factors such as line topology, meteorological conditions, or even user characteristics at the line level could provide valuable insights, especially for linking operational efficiency with user behavior in a more demand-oriented or marketing perspective. Ultimately, improving our understanding of these multiple determinants serves not only the economic efficiency of the railway system but also contributes to the broader societal goal of encouraging a modal shift from car to rail.

6 Conclusion and Policy Implications

The progressive opening of regional rail markets in France marks a structural turning point for the governance of public service obligations. Competition for the market is reshaping the traditional, network-based model into a constellation of tendered service bundles. In such a fragmented environment, evaluating efficiency at aggregated regional or national scales becomes increasingly inadequate. Understanding performance requires moving to the operational unit where production decisions are made: the line.

This paper proposes a new empirical level of analysis for railway efficiency evaluation and demonstrates its feasibility using a unique dataset of French regional lines between 2015 and 2019. By applying a stochastic frontier approach at the line level, we revealed substantial heterogeneity not only across regions but also within them, even under a single operator and a uniform regulatory framework. These intra-regional differences highlight the limitations of regional averages and call for a renewed evaluative lens. While our model does not explicitly estimate the determinants of these variations, the observed dispersion of efficiency scores suggests that factors operating below the regional level, such as differences in rolling-stock composition, maintenance intensity, or service configuration, could play a role. Identifying and quantifying these influences remain important avenues for future research.

Methodologically, the study shows that Stochastic Frontier Analysis (SFA) can be applied directly to operational accounting data to produce interpretable, reproducible, and policy-useful efficiency scores. The results indicate increasing returns to scale at the line level suggests that economies of density exist. Beyond its econometric contribution, the paper advances an evaluation framework tailored to the realities

of competitive tendering. Efficiency assessment should be continuous and cyclical: ex-ante to establish a factual baseline and support tender design; in-itinere to monitor performance during contract execution; and ex-post to evaluate outcomes and extract lessons for future tenders. Such a system would allow public transport authorities (PTAs) to move from reactive to anticipatory regulation, using evidence to design contracts, benchmark operators, and justify public funding.

Several limitations temper our conclusions. The absence of passenger-km data, service quality indicators, and detailed environmental covariates means that part of the estimated inefficiency likely reflects contextual heterogeneity rather than pure managerial performance. Nevertheless, this constraint mirrors the informational reality of most PTAs, for whom line-level data remain the most operationally accessible layer of information. In that sense, our approach offers a pragmatic and replicable method to enhance transparency and accountability in regional rail governance.

Future research should enrich this framework by integrating technical data on rolling-stock age and type, infrastructure quality, and demand indicators. The transition toward multi-operator environments will also call for the estimation of SFA models that disentangle persistent heterogeneity from inefficiency. At a broader scale, developing a coordinated national line-level database would enable consistent benchmarking across regions and operators. Ultimately, this study aims to contribute to a culture of measurement and learning in the regional railway sector, even at a European scale. As competition expands, the challenge is not merely to identify who performs better, but to understand why. By grounding evaluation in robust econometric methods and operationally meaningful units, public authorities can turn competition into a mechanism for collective improvement rather than fragmentation. Evidence-based benchmarking at the line level thus becomes a cornerstone of the emerging regulatory paradigm that reconciles efficiency, transparency, and the long-term sustainability of regional rail services.

References

- Affuso, L., Angeriz, A., & Pollitt, M. (2009). The Impact of Privatisation on the Efficiency of Train Operation in Britain. *Working Papers*, 28. URL: <https://ideas.repec.org//p/cgs/wpaper/28.html>
- Alexandersson, G. (2009). Rail Privatization and Competitive Tendering in Europe. *Built Environment (1978-)*, 35(1), 43–58.
- Asmild, M., Holvad, T., Hougaard, J. L., & Kronborg, D. (2009). Railway reforms: Do they influence operating efficiency? *Transportation*, 36(5), 617–638. <https://doi.org/10.1007/s11116-009-9216-x>
- Autorité de régulation des transports. (2023). *Bilan ferroviaire 2023*. URL: https://www.autorite-transport.fr/wp-content/uploads/2025/02/bilan_ferroviaire_2023_integrale.pdf
- Autorité de régulation des transports. (2023). *Avis portant sur la fixation des redevances relatives aux prestations régulées fournies par SNCF Voyageurs dans les installations de service des centres d'entretien pour les horaires de service 2024 et 2025*. URL: https://www.autorite-transport.fr/wp-content/uploads/2024/01/avis-2023-063_projet-vnc.pdf
- Battese, G. E., & Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productive Analysis*, 3(1), 153–169. <https://doi.org/10.1007/BF00158774>
- Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics*, 20(2), 325–332. <https://doi.org/10.1007/BF01205442>
- Benga, A., Delgado Rodríguez, M. J., De Lucas Santos, S., & Zeneli, G. (2025). Efficiency analysis of European railway companies and the effect of demand reduction. *Journal of Rail Transport Planning & Management*, 34, 100516. <https://doi.org/10.1016/j.jrtpm.2025.100516>
- Blayac, T., Bougette, P., & Laroche, F. (2024). An analysis of intermodal competition and multiproduct Incumbent's strategies in the French market: What drive high-speed trains' prices and frequencies? URL: <https://doi-org.inshs.bib.cnrs.fr/10.1016/j.tranpol.2024.07.016> *Transport Policy*, 156, 25–42.
- Cantos, P. (2001). Regulation and efficiency: The case of European railways. *Transportation Research Part A*.

- Cantos, P., Pastor, J. M., & Serrano, L. (2012). Evaluating European railway deregulation using different approaches. *Transport Policy*, 24, 67–72. <https://doi.org/10.1016/j.tranpol.2012.07.008>
- Chambre Régionale des Comptes. (2023). *Évaluation de la politique du matériel roulant ferroviaire de la Région Auvergne-Rhône-Alpes*. URL: https://www.ccomptes.fr/sites/default/files/2024-07/ARA%20202420_4.pdf
- Catalano, G., Daraio, C., Diana, M., Gregori, M., & Mattucci G., (2019) *Efficiency, effectiveness and impacts assessment in the rail transport sector: A state-of-the-art critical analysis of current research*. *International Transactions in Operational Research*, 26, 5–40. URL: https://onlinelibrary.wiley.com/doi/epdf/10.1111/itor.12551?getft_integrator=sciencedirect_contenthosting&src=getftr&utm_source=sciencedirect_contenthosting
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429–444. [https://doi.org/10.1016/0377-2217\(78\)90138-8](https://doi.org/10.1016/0377-2217(78)90138-8)
- Coelli, T., & Perelman, S. (1999). A comparison of parametric and non-parametric distance functions: With application to European railways. *European Journal of Operational Research*, 117(2), 326–339. [https://doi.org/10.1016/S0377-2217\(98\)00271-9](https://doi.org/10.1016/S0377-2217(98)00271-9)
- Cour des comptes. (2024). *Les transports express régionaux à l'heure de l'ouverture à la concurrence*. URL: <https://www.ccomptes.fr/fr/publications/les-transports-express-regionaux-lheure-de-louverture-la-concurrence-0>
- Couto, A., & Graham, D. J. (2009). The determinants of efficiency and productivity in European railways. *Applied Economics*, 41(22), 2827–2851. <https://doi.org/10.1080/00036840801949782>
- Cowie, J. (2009). The British Passenger Rail Privatisation: Conclusions on Subsidy and Efficiency from the First Round of Franchises. *Journal of Transport Economics and Policy*, 43(1), 85–104.
- Cowie, J., & Riddington, G. (1996). Measuring the efficiency of European railways. *Applied Economics*, 28(8), 1027–1035. <https://doi.org/10.1080/000368496328155>
- Demsetz, H. (1968). Why Regulate Utilities? *Journal of Law and Economics*, 11(1), 55–65. <http://www.jstor.org/stable/724970>
- Desmaris, C., & Monchambert, G. (2023). Regional passenger rail efficiency: Measurement and explanation in the case of France. *Journal of Transport Economics and Policy*, 57(1), 22–58.
- Directive (EU) 2016/2370 of the European Parliament and of the Council of 14 December 2016. Amending Directive 2012/34/EU as regards the opening of the market for domestic passenger transport services by rail and the governance of the railway infrastructure. *Official Journal of the EU*, L 352. URL: <http://data.europa.eu/eli/dir/2016/2370/oj/eng>
- Fabre, N., Laroche, F., & Bouzouina, L. (2025). Rolling Stock Management and Railway Efficiency: Evidence from a Line-Level Analysis of Regional Rail Services in France. *Working paper*. <https://hal.science/hal-05387769>
- Farrell, M. J. (1957). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society, Series A (General)*, 120(3), 253–281. <https://doi.org/10.2307/2343100>
- Friebel, G., Ivaldi, M., & Vibes, C. (2010). Railway (De)Regulation: A European efficiency comparison. *Economica*, 77(305), 77–91. <https://doi.org/10.1111/j.1468-0335.2008.00739.x>
- Fried, H. O., Schmidt, S. S., & Lovell, C. A. K. (1993). *The Measurement of Productive Efficiency: Techniques and Applications*. Oxford University Press.
- Gathon, H.-J., & Perelman, S. (1992). Measuring technical efficiency in European railways: A panel data approach. *Journal of Productivity Analysis*, 3(1–2), 135–151. <https://doi.org/10.1007/BF00158773>
- Guihéry, L. (2014). Competition in regional passenger rail transport in Germany (Leipzig) and lessons to be drawn for France. *Research in Transportation Economics*, 48, 298–304. <https://doi.org/10.1016/j.retrec.2014.09.056>
- Henningsen, A. (2019). *Introduction to Econometric Production Analysis with R*. Department of Food and Resource Economics, University of Copenhagen, Denmark.

- Holvad, T. (2020). Efficiency analyses for the railway sector: An overview of key issues. *Research in Transportation Economics*, 82, 100877. <https://doi.org/10.1016/j.retrec.2020.100877>
- Lalive, R., & Schmutzler, A. (2008). Exploring the effects of competition for railway markets. *International Journal of Industrial Organization*, 26(2), 443–458. <https://doi.org/10.1016/j.ijindorg.2007.02.005>
- Lalive, R., Schmutzler, A., & Zulehner, C. (2015). Auctions vs negotiations in public procurement: Which works better? <https://doi.org/10.5167/UZH-115738>
- Lan, L. W., & Lin, E. T. J. (2003). Technical efficiency and service effectiveness for railways industry: DEA approaches. *Journal of the Eastern Asia Society for Transportation Studies*, Vol.5, 2003. https://www.researchgate.net/publication/228361269_Technical_Efficiency_and_Service_Effectiveness_for_Railways_Industry_DEA_Approaches
- Laroche, F., & Lamatkhanova, A. (2022). Exploring effects of competitive tenders on users in the regional railway market: Evidence from Europe. *International Journal of Transport Economics*, XLIX(1). <https://doi.org/10.19272/202206701002>
- Laroche, F., & Lamatkhanova, A. (2022). Effects of open access competition on prices and frequencies on the interurban railway market: Evidence from Europe. *Research in Transportation Business & Management*, 43, 100705. <https://doi.org/10.1016/j.rtbm.2021.100705>
- Le, Y., Oka, M., & Kato, H. (2022). Efficiencies of the urban railway lines incorporating financial performance and in-vehicle congestion in the Tokyo Metropolitan Area. *Transport Policy*, 116, 343–354. <https://doi.org/10.1016/j.tranpol.2021.12.017>
- Lerida-Navarro, C., Nombela, G., & Tranchez-Martin, J. M. (2019). European railways: Liberalization and productive efficiency. *Transport Policy*, 83, 57–67. <https://doi.org/10.1016/j.tranpol.2019.09.002>
- Link, H. (2016). A two-stage efficiency analysis of rail passenger franchising in Germany. *Journal of Transport Economics and Policy*, 50(1), 76–92.
- Link, H. (2019). The impact of including service quality into efficiency analysis: The case of franchising regional rail passenger services in Germany. *Transportation Research Part A*, 119, 284–300. <https://doi.org/10.1016/j.tra.2018.11.019>
- Merkert, R., Smith, A. S. J., & Nash, C. A. (2010). Benchmarking of train operating firms: A transaction cost efficiency analysis. *Transportation Planning and Technology*, 33(1), 35–53. <https://doi.org/10.1080/03081060903429330>
- Preston, J., & Robins, D. (2013). Evaluating the long term impacts of transport policy: The case of passenger rail privatisation. *Research in Transportation Economics*, 39(1), 14–20. <https://doi.org/10.1016/j.retrec.2012.05.019>
- Simar, L., & Wilson, P. W. (1998). Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44(1), 49–61. <https://doi.org/10.1287/mnsc.44.1.49>
- Wheat, P., & Smith, A. S. J. (2015). Do the usual results of railway returns to scale and density hold in the case of heterogeneity in outputs? *Journal of Transport Economics and Policy*, 49. <https://www.jstor.org/stable/jtranseconpoli.49.1.0035>