

Link, Heike

Article — Published Version

A Two-Stage Efficiency Analysis of Rail Passenger Franchising in Germany

Journal of Transport Economics and Policy

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Link, Heike (2016) : A Two-Stage Efficiency Analysis of Rail Passenger Franchising in Germany, Journal of Transport Economics and Policy, ISSN 0022-5258, University of Bath, Bath, Vol. 50, Iss. 1, pp. 76-92,
<http://www.ingentaconnect.com/content/lse/jtep/2016/00000050/00000001/art00006>

This Version is available at:

<http://hdl.handle.net/10419/142117>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

A Two-stage Efficiency Analysis of Rail Passenger Franchising in Germany

Heike Link

Address for correspondence: Heike Link, German Institute for Economic Research (DIW Berlin), Mohrenstr. 58, 10117 Berlin, Germany (hlink@diw.de).

This research was only possible with the support of numerous people in the German public transport authorities and the transport ministries in the federal states who helped to collect the necessary data. I have also to thank my students Phillip Albert, Linus Haase, and Hannes Fauser, who worked hard in compiling and cross-checking the data provided by public transport authorities, and who conducted extensive and time-consuming searches on alternative sources for supplementing information and data. Finally, I am grateful for helpful discussions with the participants at the Kuhmo Nectar Conference on Transport Economics in Toulouse 2014, and to two anonymous referees for their comments. Responsibility for this paper is of course solely my own.

Abstract

This paper analyses the differences in the efficiency of using subsidies for franchised regional rail services between the federal states in Germany, and provides evidence on the impact of procurement strategies and contractual design on the efficient use of funds. The analysis is based on a 15-year panel data set at the level of the federal states and employs a two-stage efficiency analysis, including a DEA approach and a Tobit panel model. The analysis shows that a higher share of tendering, a higher share of gross contracts, and longer and smaller contracts were efficiency-enhancing factors in the period of analysis.

Final version: September 2015

1.0 Introduction

A major policy problem in rail restructuring is how to ensure an efficient provision of — eventually unprofitable — services in the presence of economies of density. A solution to this problem is competitive tendering for franchises, known under the term competition for the market. This approach is capable of preserving the benefits of economies of density and integrated services while reducing inefficiency (Demsetz, 1968). If appropriately implemented, it has the potential to give incentives to improve service quality, and is consistent with maintaining unprofitable but socially desirable services.

Germany has indeed introduced a franchising system for unprofitable and subsidised regional rail passenger services (PSOs) with the aim to achieve a decent level of services at as low subsidies as possible. Since 1996 the federal states have been responsible for procuring these services from rail companies and for financing them within franchise contracts (so-called regionalisation). They have considerable freedom to organise this process, and consequently the character of franchise contracts in Germany varies regarding the type of procurement (competitive tendering versus direct negotiations), and regarding contract arrangements such as contract duration, network size, and contract type (net versus gross contracts). Financially, the franchising system is based on transfers from the federal budget to the federal states at an annual level of about €7 billion, which can be regarded as a stable and sound public financing for PSOs.

The so-called regionalisation with the agreed sound financing of rail-PSOs in Germany had positive impacts on service provision, patronage, customer satisfaction and service quality (see Link and Merkert, 2011). Germany has the lowest level of tax support per passenger-km compared with Sweden and the UK (see Nash *et al.*, 2013). The overall level of subsidy per unit of output has declined over time, but at the same time, the subsidy per output-unit shows a considerable range between the federal states (see VZBV, 2010). This range has fed into a political debate on a suspected inefficient use of subsidies by the federal states.

The lack of sufficiently detailed data has so far not allowed analysing the efficiency of using public funds to subsidise regional rail services and to identify the determinants of efficiency. Available studies for Germany such as Beck (2011), Lalive and Schmutzler (2008, 2011), and Hunold and Wolf (2012), rather focus on subsamples of contracts awarded under competitive tenders, and study procurement design, potentials, and barriers for market entry. The caveats of these studies is that they refer to rather small samples (in terms of both cross-sectional and longitudinal observations) without any evidence on the degree of representativeness, and that they do not use any of the established tools for efficiency analysis such as frontier techniques.

The analysis presented in this paper contributes to the existing research by analysing the efficiency of using public funding for franchised regional rail services in Germany. It is based on a unique data set which includes for all regional rail services data at the level of federal states on traffic and patronage, monetary data, and data on contract features and type of procurement. Specifically, the analysis is aimed at determining and explaining differences in the efficiency of using subsidies between the federal states, and establishing evidence on the impact of procurement strategies and contractual design on the efficient use of funds.

2.0 Available Research

Although liberalisation of European railways started two decades ago, quantitative evidence on the impact of deregulation, vertical separation, competition, and franchising on the efficiency of railways is limited. This holds in particular for the link between competition and franchising policies to efficiency. Most of the available studies on rail productivity and efficiency focus on the impact of vertical integration versus vertical separation, but without reaching a consensus on the direction of impact. Ivaldi and McCullough (2001), Bitzan (2003), Jensen and Stelling (2007), and Growitsch and Wetzel (2009), are examples of studies that conclude that vertical separation increases costs, while Cantos *et al.* (1999, 2010), and Mituzani and Uranishi (2010), suggest that vertical separation has positive effects of railway efficiency; however, the latter study restricts these positive effects to railways with low train density. All of these studies focus on companies' efficiency rather than on the use of public spending, and the majority are based on international data at country-level, mostly obtained from UIC statistics.

The impact of competitive tendering and franchising systems is even less analysed than the issue of vertical integration versus separation. Driessen *et al.* (2006) show for European and Japanese railways that competition for the market (concessions) increases efficiency more than competition in the market (free entry). Mulder *et al.* (2005) conclude from a cost–benefit analysis for the Netherlands that competition in the freight sector has increased both efficiency and performance of railways, while competition in passenger transport had no effects. Cantos *et al.* (2010) conduct a DEA analysis of European railways and find that open access in freight transport has increased rail efficiency, while the introduction of tendering in passenger transport has had no significant impact. Pollitt and Smith (2002) derive from a social cost–benefit analysis that the restructuring of British rail — which included a franchising system for passenger services — has led to efficiency improvements, lower prices, and better output quality; however, Nash and Smith (2006) identify cost increases after the first round of franchising. Affuso and Newbery (2002) study the influence of contractual arrangements such as contract duration and requests for renegotiation, as well as uncertainty on demand, and its variation and firm characteristics on investments levels into rolling stock. Based on a panel of twenty-five British franchisees between 1997 and 2000, they find that in response to more competition, investments increase and, contrary to theoretical considerations and intuition, shorter contracts generate higher rates of investments. Similar to the aforementioned studies on vertical integration versus separation, the available research on competition and franchising deals with the impact on companies' efficiency rather than on the efficiency in using public funds.

All available studies for Germany are based on subsamples of franchising contracts. Beck (2011) analyses potential barriers for new entrants, based on thirty tenders for rail passenger services during 1997 and 2007. He finds a negative correlation between the number of bidders in a tender and taking over the revenue risk by the bidder (net contracts), and a positive correlation between the number of bidders and taking over the price risk for input factors by the public transport authorities (PTAs). Lalive and Schmutzler (2008) analyse under which contractual arrangements, such as size of franchise, and length and type of network, new entrants are likely to win competitive tenders against the incumbent, based on a subsample of seventy-seven tenders. They find that the incumbent won large

but not necessarily longer contracts. Hunold and Wolf (2012) study the determinants of participation in competitive tenders, and the factors that influence the outcomes of awarding procedures such as the type of winner (new entrant versus incumbent) and the subsidy level, based on a subsample of awarding procedures between 1996 and 2011. They find that the incumbent is more likely to win large and long contracts, as well as those contracts where the operation of used rolling stock was admitted. In line with Beck (2011), they also find that the arrangements regarding the revenue risk influences the number of bidders. Lalive and Schmutzler (2011) analyse differences in procurement prices between lines that were competitively tendered and those which were procured within direct negotiations. They conclude that competitive awarding procedures have led to higher frequencies of services and to lower prices. All the aforementioned studies for Germany refer either to rather small samples or to samples covering rather short time periods, and all of them face the problem that it is unknown whether the subsamples used are representative. None of these studies applies state-of-the-art tools for efficiency analysis.

In summary, there is a lack of research in analysing how external factors, such as socio-demographic characteristics and general economic indicators, franchising strategies such as the procurement mode, and the contractual design of franchise contracts, influence the efficiency of using public subsidies for rail services, at both the national and international level. Against this research background, this paper aims to contribute to the existing research by analysing the efficiency of using public funding of regional rail services within the German rail franchising system and by identifying the major determinants of efficiency.

3.0 Methodology

The research question dealt with in this paper differs from the typical rail efficiency problem insofar as it aims at analysing the efficient use of subsidies in contrast to study rail companies' efficiency. This different research focus implies a different formulation of inputs than in traditional rail efficiency analysis. While the typical input definition is to use labour, material (for example, energy), and capital (for example, rolling stock, other facilities) as variables, this analysis defines monetary subsidies (further disaggregated into operating subsidies and infrastructure charges) and track length as inputs.¹

This paper employs a two-stage bootstrapped efficiency analysis with a DEA application in the first stage and a Tobit panel model for explaining the first-stage efficiency scores. The first stage DEA is a deterministic, non-parametric method to specify the efficient production frontier. Based on observed inputs (w) and outputs (y) for a set of N decision-making units (DMUs), relative efficiency scores for each unit are derived by searching for the most efficient DMUs in the sample — that is, those units with the lowest input levels for a given level of output.² These form the efficiency frontier which

¹Ideally, splitting up subsidies further into spending on labour, material, and capital would enable linking of the analysis to traditional rail-efficiency analysis, and could provide further insights into complementarity and substitution properties of inputs, in particular if the data would allow an econometric cost function approach such as stochastic frontier models. However, this type of disaggregating subsidies appears to be empirically impossible.

²The performance of a DMU is defined as the ratio between the weighted sums of outputs and inputs, whereby the weights are allocated within the DEA model.

envelops all observed data points in the sample. Efficiency of each DMU is measured by the distance Ω between the single observation and the frontier.

The choice of DEA for the analysis in this paper was on the one hand driven by the fact that the necessary data on factor inputs and input prices for stochastic frontier models, the econometric alternative to DEA, was not available.³ Further reasons are the advantages of the DEA technique such as its capability of handling multiple outputs and inputs, and of dealing with data that are on different units. Furthermore, it does not require any assumptions about cost minimisation or profit maximisation, or about the functional form of the frontier, and does not require price information. Caveats of the method include the potential danger of deeming too many observations as inefficient because the concept fails to recognise nonconvexities of the envelope (see, for example, De Borger and Kerstens, 1996). A major problem is the deterministic nature of the method which allocates variations in a firm's performance exclusively to inefficiency, and which lacks statistical inference. This, however, can be solved by bootstrapping procedures to correct for the potential bias in the estimated distance function from the true frontier and to calculate confidence intervals (see, for example, Simar and Wilson, 2000).

Depending on the assumption imposed regarding the ability of firms to influence either inputs or outputs, the distance function can be applied as input-oriented — that is, reflecting a situation where the output set is fixed by exogenous factors — or as output-oriented, which refers to the opposite situation. Given that the federal states have the decision power on the level of subsidies rather than over outputs such as train-km and passenger-km (which are substantially influenced by long-term planning and contracts, as well as by macro-economic and demographic factors), the first-stage DEA approach uses an input-oriented distance function. It should be noted that according to Coelli and Perelman (1999), output- and input-oriented models estimate exactly the same frontier surface, and therefore identify the same set of DMUs as efficient; however, the scores themselves differ. They also show that, for railways, the orientation does not matter as much as it does for other industries.

The distance function can be estimated under alternative assumptions on the type of returns to scale; that is, constant (CRS) versus variable (VRS). The input-oriented DEA model with constant returns to scale is (see Coelli *et al.*, 2005):

$$\begin{aligned}
 & \text{Min wrt } \Omega_i, \lambda: \Omega_i \\
 & \text{st } \sum_k \lambda_k y_k - y_i \geq 0, \quad k = 1, \dots, N, \\
 & \quad \Omega_i w_i - \sum_k \lambda_k w_k \geq 0, \quad k = 1, \dots, N, \\
 & \quad \lambda_k \geq 0,
 \end{aligned} \tag{1}$$

where λ_k represent the optimal weights for inputs and outputs which are searched for each DMU by the optimisation programme. The value of Ω_i is the input-oriented technical

³The collection of factor inputs and input prices, or a subsidy given to a franchisee, are also complicated due to the fact that it cannot simply be derived from the companies' profit-and-loss statements.

efficiency score for the i th DMU.⁴ A VRS model is obtained by imposing the additional convexity constraint:

$$\sum_k \lambda_k = 1. \quad (2)$$

Since the German rail sector is characterised by imperfect competition, budget restrictions, and regulatory deficits regarding the entry of new companies, the CRS assumption appears to be not appropriate. A further argument against the CRS assumption is that this paper does not analyse a technical relationship between outputs (passenger-km and train-km) and inputs such as labour, material, and capital, but between outputs and monetary public subsidies as discussed at the beginning of this section. Therefore, this paper presents efficiency scores both under CRS and VRS, with the latter being the preferred assumption. As is common in DEA studies, a bootstrapping approach based on Simar and Wilson (1998) is applied in order to obtain confidence intervals, and to overcome the potential problem of biased results in the second-stage regressions. The bias-corrected results for the efficiency scores were obtained from 1,000 iterations.

For the analysis presented in this paper, two output variables and three input variables were defined. As it is common practice in studies of rail efficiency, this paper uses train-km and passenger-km as output variables. The definition of inputs contains two monetary variables, namely operating subsidies (to be influenced directly by the DMUs) and infrastructure charges to be paid (exogenously given by the infrastructure provider), and one physical variable, the track length. The paper presents results for four different DEA models. All of them use train-km and passenger-km as output variables, but differ regarding the specification of the input variables: V0 considers only one input which is the total subsidy as the sum of operating subsidy and infrastructure charges; V1 includes two monetary inputs (operating subsidy and infrastructure charges to be paid); V2 includes total subsidy and track length; and V3 includes the two monetary and the physical input variable.

Following previous studies such as Merkert *et al.* (2010) and Cantos *et al.* (2010), this paper regresses the efficiency scores obtained from the first-stage DEA approach against a set of independent explanatory variables in order to explain differences in the efficiency between the federal states. The explanatory variables contain two types of data. First, a set of policy variables which reflect the decisions of the federal states on how to organise rail franchising. To these variables belong the share of tendered train-km, the share of train-km under net contracts, the average size of contracts (expressed in train-km), and the average contract duration. Second, environmental variables which cannot be influenced by the DMUs such as the population density, car density, and the rate of unemployment.

The efficiency scores are bound in the (0,1) interval and represent a censored variable which requires an appropriate econometric treatment by means of corner solution — that is, Tobit models (named after Tobin, 1958). The general formulation of a one-limit Tobit model with 0 as the limit is usually given as an index function (or latent variable model):

$$\begin{aligned} y_i^* &= x_i' \boldsymbol{\beta} + \varepsilon_i, \varepsilon_i \sim N(0, \sigma^2), \\ y_i &= \max(0, y_i^*), \end{aligned} \quad (3)$$

⁴Note that most DEA studies use the Farell (1957) measure, which is obtained by the reciprocal of the Shephard distance function.

where x'_i represent the explanatory variables, β is the vector of parameters to be estimated, y_i is the observed dependent variable, and y_i^* is the unobserved, latent variable. The log-likelihood for this model is a mixture of discrete and continuous distributions as follows:

$$\log L = \sum_{y_i > 0} -\frac{1}{2} \left[\log(2\pi) + \log \sigma^2 + \frac{(y_i - x'_i \beta)^2}{\sigma^2} \right] + \sum_{y_i = 0} \log \left[1 - \Phi \left(\frac{x'_i \beta}{\sigma} \right) \right], \quad (4)$$

where the first term refers to the classical regression for the non-limit observations and the second term gives the relevant probabilities for the limit observations (see Greene, 2002). The Tobit model can be estimated by MLE procedures. Marginal effects are obtained by:

$$\frac{\partial E(y_i|x_i)}{\partial x_i} = \Phi \left(\frac{x'_i \beta}{\sigma} \right) \beta, \quad (5)$$

and indicate that a change in x'_i affects the conditional mean of y_i^* in the positive part of the distribution and, in addition, also the probability that the observation falls in that part of the distribution.

For the type of data used in this paper, the basic Tobit model has to be extended to a panel context. As common, the two standard approaches for incorporating individual heterogeneity are available; that is, the fixed and random effects model. For incorporating individual heterogeneity, the basic Tobit model is modified to:

$$\begin{aligned} y_{it}^* &= \alpha_i + x'_{it} \beta + \epsilon_{it}, \epsilon_{it} \sim N(0, \sigma^2), \\ y_{it} &= \max(0, y_{it}^*), \end{aligned} \quad (6)$$

where α_i represents unobserved individual heterogeneity. While the fixed effects model allows correlation between α_i and x'_{it} , the random effects model assumes that heterogeneity is uncorrelated with the regressors. The fixed effects model is estimated by including a set of dummy variables d_i , ($i = 1, \dots, N$), which indicate membership in group i :

$$\begin{aligned} y_{it}^* &= \sum_{i=1}^N d_{it} \alpha_i + x'_{it} \beta + \epsilon_{it}, \epsilon_{it} \sim N(0, \sigma^2), \\ y_{it} &= \max(0, y_{it}^*). \end{aligned} \quad (7)$$

The corresponding log likelihood function is given by:

$$\begin{aligned} \ln L &= \sum_{i=1}^N \sum_{t=1}^T (1 - c_{it}) \ln \Phi(-\eta_i - x'_{it} \gamma) + c_{it} \ln [\theta \phi y_{it} - \eta_i - x'_{it} \gamma], \\ c_{it} &= 1, \quad \text{if } y_{it} > 0, \\ c_{it} &= 0, \quad \text{otherwise,} \end{aligned} \quad (8)$$

with $\eta_i = \alpha_i/\sigma$. Following Olsen's (1978) transformation, the parameters γ and θ are:

$$[\gamma, \theta] = \left[\frac{\beta}{\gamma}, \frac{1}{\sigma} \right]. \quad (9)$$

The random effects model is:

$$\begin{aligned} y_{it}^* &= \alpha + x'_{it}\boldsymbol{\beta} + u_{it} + \epsilon_{it}, \epsilon_{it} \sim N(0, \sigma^2), \\ y_{it} &= \max(0, y_{it}^*). \end{aligned} \quad (10)$$

Since in the random effects model the regressors are assumed to be uncorrelated with heterogeneity, a different approach to estimation is necessary. The conditional log likelihood in the presence of random effects is given by:

$$\ln L_{\text{cond}} = \sum_{i=1}^N \ln \prod_{t=1}^T [\Phi(-\tau w_i - x'_{it}\boldsymbol{\gamma})]^{1-c_{it}} [\theta\varphi(\theta y_{it} - \tau w_i - x'_{it}\boldsymbol{\gamma})]^{c_{it}}, \quad (11)$$

where $\tau = \sigma_{\alpha}/\sigma$, $w_i \sim [0, 1]$, and $\boldsymbol{\gamma}$, θ , and τ are unknown parameters to be estimated. Equation (11) includes the unobserved random effect and can thus not serve as the basis for estimation. The unconditional log likelihood is:

$$\ln L = \sum_{i=1}^N \int_{-\infty}^{\infty} \prod_{t=1}^T [\Phi(-\tau w_i - x'_{it}\boldsymbol{\gamma})]^{1-c_{it}} [\theta\varphi(\theta y_{it} - \tau w_i - x'_{it}\boldsymbol{\gamma})]^{c_{it}} \varphi(w_i) dw_i. \quad (12)$$

While the random effects model has the advantage that it is more manageable than the fixed effects model, its disadvantage is the probably unrealistic assumption of no correlation between heterogeneity and regressors. Therefore, Woldridge (2010) and Greene (2008) suggest applying the Mundlak (1978) correction by including the group means of regressors as additional explanatory variables. This leads to the correlated random effects model:

$$\begin{aligned} y_{it}^* &= \alpha_i + x'_{it}\boldsymbol{\beta} + \epsilon_{it}, \epsilon_{it} \sim N(0, \sigma^2), \\ \alpha_i &= \text{mean}(x'_{it})\boldsymbol{\pi} + w_i, w_i \sim N(0, 1), \\ y_{it} &= \max(0, y_{it}^*). \end{aligned} \quad (13)$$

Since it cannot be decided *a priori* which of these panel models is appropriate, this paper reports and compares the results of the three versions of a Tobit panel model summarised above — that is, a fixed effects, a random effects, and a correlated random effects model. Finally it has to be noted that, following Greene (1993), the efficiency scores Ω_i were censored at zero by transforming:

$$y_i = 1 - \Omega_i, \quad (14)$$

in order to avoid concentrating of variables at unity. Both the DEA models and the Tobit regressions were performed in NLOGIT.

4.0 The Data

Potentially, an analysis of efficiency in subsidising regional rail passenger services could focus on three levels of disaggregation, which imply the definition of the respective decision-making units (DMUs):

- (1) The level of federal states which are responsible for the general decisions on franchising strategies and designs, for the adoption of passenger fares, and which are also involved in decisions on the closure of unprofitable rail lines.

- (2) The level of PTAs which are responsible for the day-to-day management of franchise contracts.
- (3) The contract level, which would allow to study more specifically contract design, patronage, and, in addition, reliability, punctuality, and customer satisfaction (for which data covering all PTAs and all federal states are not available).

In Germany, centrally held and publicly accessible data is not available for any of these disaggregation levels. Therefore, a comprehensive data collection work was conducted whereby the level of federal states was chosen. The data set includes variables on the volume of regional rail services (train-km and passenger-km), on the level of subsidies and fare revenues, and on the characteristics of the contracts (net versus gross contracts, tendered versus non-tendered contracts, contract size, and contract duration). Data collection work included questionnaires and interviews with PTAs and the transport ministries of the federal states. This information was supplemented by data from various other sources such as business reports of rail companies (including DB, Deutsche Bahn), data from the Association of PTAs (BAG SPNV) and from the Association of Public Transport Operators (VDV), and official statistics of the federal states, as well as their own calculations. All monetary data was deflated with the price index for regional rail passenger services at constant prices of 2010.

Comprehensive own calculations were necessary for two types of data. First, infrastructure charges (that is, charges paid for the use of tracks and stations) which are part of the subsidy paid within the franchise contract had to be quantified and allocated to federal states. Information on revenues of DB Netz and DB Station (the track provider and the operator of stations in Germany) from user charges is only available as overall figures, and had in a first step to be broken down to traffic types, based on earlier work on different versions of the track access charging schemes and the charging schemes for rail stations during 1996–2010 (see Link, 2004, 2013). In a second step, infrastructure charges from regional rail services were allocated to federal states based on information from profit-and-loss statements of several rail companies, partially available regional information, the train-km of different traffic types, and by considering the regional surcharges as applied in DB's access charging scheme from 1999 to 2011.

A second issue refers to the treatment of fare revenues because all federal states use a mixture of net and gross contracts within their franchising strategies. This implied the need for an adjustment of subsidies paid for regional rail services because the amount of public money finally to be spent by the PTAs for these services is influenced by the chosen contract type — that is, the different treatment of fare revenues in net and gross contracts. While under gross contracts the PTA pays an overall subsidy to the rail operator and receives all fare revenues, in a net contract regime the rail operator keeps all fare revenues and receives a subsidy to compensate for those costs not covered by rail fares. Analysing the efficiency of subsidising regional rail passenger services therefore requires considering the fare revenues collected by PTAs as a subsidy-minimising income of PTAs. Based on data from the VDV, fare revenues differentiated by federal states were allocated to passenger-km and train-km operated under both types of contracts, and the amount of subsidies paid was adjusted by subtracting fare revenues collected from gross contracts.

As a result of this comprehensive data collection work, a unique panel data set with the federal states as cross-section units and a coverage of the period from 1996 (the start of franchising) to 2010 was elaborated.

5.0 Estimation Results

Table 3 summarises the estimated and the bias-corrected scores for technical efficiency obtained with the four DEA models under different assumptions regarding returns to scale. The average bias-corrected technical efficiency scores, as well as the corresponding standard deviations, are for the CRS specification smaller than the uncorrected scores and for the VRS model roughly equal. The CRS results confirm earlier studies (Growitsch and Wetzel, 2009; Merkert *et al.*, 2010), which indicate that DEA models without bootstrapping tend to overestimate technical efficiency, even though the difference in this application is not as distinctive as in other studies. The level of DEA scores is in line with previous literature on rail efficiency (see Growitsch and Wetzel, 2009; Cantos *et al.*, 2010; Merkert *et al.*, 2010).

The main result from the first-stage DEA is that, based on the preferred specification of variable returns to scale, the average federal states could save subsidies in a range of 25 per cent to 40 per cent. This, however, has to be interpreted as the maximum possible saving if all federal states would face identical conditions (population density, level of infrastructure charges to be paid, and so on) as the most efficient states. The considerable variation of these indicators in Table 1 shows this is obviously not the case. To underline this argument, Table 3 shows disaggregated technical efficiency scores for different types of states. First, a distinction between the federal states of Hamburg, Bremen, and Berlin (the so-called city states), and the remaining federal states. The city states represent large urban

Table 1
Descriptive Statistics

	<i>Variable</i>	<i>Unit</i>	<i>Mean</i>	<i>Standard deviation</i>	<i>Minimum</i>	<i>Maximum</i>
Population density	P_SKM	inhabitants/skm	667.8	1,012.8	71.0	3,897.0
Car density	CAR	Cars per 1,000 inhabitants	508.95	62.58	317.11	624.6
Rate of unemployment	UNEMPL	%	11.75	4.45	4.10	20.5
Track length	TRACKS	km	2,440.3	1,809.7	134.8	6,589.2
Train-km	TR_KM	million km	37.43	27.81	2.24	112.04
Passenger-km	P_KM	million km	2,546.61	2,377.21	180.39	8,725.81
Share of tendered train-km	TEND	%	6.87	9.63	0.00	45.66
Share of train-km under net contracts	NET	%	93.40	13.27	32.70	100.00
Average contract size	SIZE	million train-km p.a.	5.31	6.90	0.44	88.80
Average contract duration	DUR	years	9.55	3.71	1.41	17.50
Subsidies*	S	million €	225.86	158.00	10.37	721.47
Share of infrastructure charges** in total subsidies	INFRA	%	53.98	9.84	24.98	88.77

Notes: $N = 240$. *, At 2010 prices. Adapted by fare revenues from gross contracts. **, Infrastructure charges paid for the use of tracks and stations represent items which are paid by rail operators to rail rack providers (typically DB Netz). They are passed through to PTAs and are part of the subsidies paid by PTAs to operators within franchise contracts.

Table 2
DEA Results

	<i>VRS</i>				<i>CRS</i>			
	<i>ES</i>		<i>ES—bias corrected</i>		<i>ES</i>		<i>ES—bias corrected</i>	
	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>	<i>Mean</i>	<i>SD</i>
V0	0.6006	0.2015	0.5989	0.1925	0.5395	0.1652	0.5177	0.1548
V1	0.6856	0.1784	0.6859	0.1865	0.6405	0.1610	0.6186	0.1523
V2	0.6982	0.2009	0.6988	0.1914	0.6427	0.1740	0.6185	0.1625
V3	0.7496	0.1828	0.7513	0.1763	0.6976	0.1658	0.6747	0.1573

Notes: V0: input is total subsidy. V1: inputs are operating subsidy and infrastructure charges. V2: inputs are total subsidies and track length. V3: Inputs are operating subsidy, infrastructure charges, and track length.

areas with high population density, allowing a bundling of rail traffic and interchanges with other public transport which makes regional rail traffic attractive. Second, Table 3 shows disaggregated efficiency scores for East and West German states. As to be expected, the disaggregated efficiency scores for the city states are higher than for the remaining states. Another observation is that the federal states in the Western part of Germany are more efficient in using rail subsidies than the East German states. There are several potential reasons for this difference; among them is the transformation process with the related economic problems in East Germany and the lower population density. Furthermore, the East German PTAs have to pay higher infrastructure charges due to the regional surcharges levied by DB Netz from 1998 to 2011 for low-density lines, in particular in East Germany.

Tables 4 and 5 report the results of four Tobit panel models which use the efficiency scores from DEA model V3 under the VRS assumption as a dependent variable: 1) a fixed effects model with individual effects only; 2) a fixed effects model with individual and time effects; 3) a random effects model; and 4) a correlated random effects model. The scale factors for obtaining marginal effects are given at the bottom of Tables 4 and

Table 3
Disaggregated Efficiency Scores

	<i>DEA-V0</i>		<i>DEA-V1</i>		<i>DEA-V2</i>		<i>DEA-V3</i>	
	<i>ES</i>	<i>SD</i>	<i>ES</i>	<i>SD</i>	<i>ES</i>	<i>SD</i>	<i>ES</i>	<i>SD</i>
Hamburg, Bremen, Berlin	0.8244	0.2058	0.8875	0.1865	0.7984	0.2012	0.9785	0.1789
Other federal states	0.5489	0.2060	0.6694	0.1871	0.6750	0.2008	0.6949	0.1792
Out of these:								
East German states	0.4364	0.2059	0.5497	0.1869	0.6622	0.2009	0.5918	0.1791
West German states	0.6177	0.2061	0.6917	0.1872	0.6807	0.2007	0.7539	0.1796

Notes: Bias-corrected efficiency scores, based on VRS assumption. V0: DEA model with total subsidies as input. V1: DEA model with subsidies and track length as inputs. V2: DEA model with operating subsidies and infrastructure charges as inputs. V3: DEA model with operating subsidies, infrastructure charges, and track length as inputs. All figures are based on VRS assumption and are bias-corrected scores obtained from 1,000 bootstrapping iterations.

Table 4
Results of the Second-stage Tobit Regressions — Fixed Effects Models

Variable	Fixed effects (individual effects only)			Fixed effects (individual and time effects)		
	Coefficient	Std error	P(z) > Z	Coefficient	Std error	P(z) > Z
P_SKM	-0.00057	0.00051	0.2645	-0.00077	0.00049	0.1185
DUR	-0.00841***	0.00253	0.0009	0.00059	0.00339	0.8606
NET	0.51488***	0.07623	0.0000	0.30493***	0.08234	0.0002
TEND	-0.49429***	0.09016	0.0000	-0.14052	0.09784	0.1509
SIZE	0.00427***	0.00096	0.0000	0.00244**	0.00099	0.0140
Sigma	0.07680***	0.00388	0.0000	0.06806***	0.00345	0.0000
Scale factor	0.999			1.000		
Log L	206.5			230.9		

*** Significant at 1% level. ** Significant at 5% level.

Table 5
Results of the Second-stage Tobit Regressions — Random Effects Models

Variable	Random effects			Correlated random effects		
	Coefficient	Std error	P(z) > Z	Coefficient	Std error	P(z) > Z
P_SKM	-0.00017***	0.00002	0.0000	-0.00055	0.00100	0.5818
DUR	-0.00720***	0.00108	0.0000	-0.00841***	0.00272	0.0020
NET	0.51295***	0.07766	0.0000	0.50850***	0.05562	0.0000
TEND	-0.50932***	0.03914	0.0000	-0.48824***	0.11267	0.0000
SIZE	0.00439***	0.00101	0.0000	0.00438***	0.00137	0.0014
D_East	0.10040***	0.03822	0.0086	0.13735**	0.06185	0.0264
Const	-0.03181	0.10471	0.7613	0.03257	0.40199	0.9354
Sigma(v)	0.07826***	0.00263	0.0000	0.07823***	0.00366	0.0000
Sigma(u)	0.09436***	0.01948	0.0000	0.08246**	0.03670	0.0247
Scale factor	0.9986			0.9986		
Log L	177.8			181.3		

*** Significant at 1% level. ** Significant at 5% level.

5 and are near unity, indicating that the coefficients can be interpreted as almost equal to the marginal effects (evaluated at the mean of the exogenous variables).

Since the dependent variable was obtained from the transformed efficiency scores, negative signs of the parameters indicate that the respective variables have a positive impact on efficiency. In all four models, most of the coefficients are significant at 1 per cent level and have the same sign.⁵ However, the sizes of coefficients in particular for the policy variables TEND and NET differ between the fixed effects model with individual and time effects, and the remaining three models. Choosing one model as the preferred one is difficult because each of the four models has methodological advantages and disadvantages. As Greene (2008) points out, there is no simple test with known

⁵Models that included the variables car density and rate of unemployment suffered from collinearity problems, and therefore these variables were removed from the models.

properties to choose between fixed and random effects specification. A Hausmann test of the random effects alternative against the fixed effects null hypothesis is inappropriate. A Likelihood Ratio (LR) test of the null hypothesis of the fixed effects model, with individual effects only against the alternative of the fixed model with individual and time effects, does not reject the null. Within the random effects and correlated random effects alternatives, an LR test rejects the null hypothesis of random effects. This would favour the correlated random effects, even though two variables of interest, the population-density, and the contract duration are not significant.

Overall, the estimation results are intuitively plausible and confirm expectations from theory. Those federal states that use more competitive tendering for contract procurement use their available funds more efficiently than others, a finding which is in line with studies such as Driessen *et al.* (2006), and Lalive and Schmutzler (2011), but which contradicts studies which found no effect of franchising and competitive tendering (see Mulder *et al.*, 2005; Cantos *et al.*, 2010). When interpreting and comparing these results, it should be noted that this paper analyses the effect of competitive tendering on the efficiency in using subsidies, while the aforementioned studies focus on the effect of tendering on the efficiency of rail companies.

A further result is that federal states which award a higher share of train-km under gross contracts use their funds more efficiently than others. This finding differs from Hunold and Wolf (2012), which, however, did not adjust the subsidy paid by fare revenues collected. At a first glance, this finding seems to be counterintuitive, given that net contracts give incentives to the operator to increase patronage. However, the net contract framework involves potential information asymmetries between operators and PTAs, if the operators possess information on demand and revenue of the respective lines. Potential new entrants without this information might even be discouraged from market entry, a finding which is also reported by Beck (2011).⁶ A further explanation can also be found in the special German situation in particular during the first half of the observation period, where the majority of train-km was operated by the incumbent DB.⁷ During this period the by far largest share of train-km was awarded within direct negotiations without competitive pressure,⁸ and net contracts were the dominant contract type.⁹ These overlapping effects might have led to a situation where operators under net contracts (and here mostly DB) were able to receive rather high net subsidies. PTAs had finally to spend more public money for net contracts than for gross contracts for which the fare revenues collected by the PTA reduce the subsidy level — even though the net subsidy per train-km is per definition lower than the gross subsidy when not considering fare revenues.

⁶PTAs in Germany actually favour net contracts due to the incentives which this type of contract provides to increase patronage, and DB has committed itself to provide the necessary information from 2009 onwards. However, PTAs have so far considered the information provided by DB as not sufficient to use net contracts as standard procedure.

⁷In 1998, DB operated 95.5 per cent of all train-km in regional rail passenger transport. In 2003, DB's share still amounted at 90.1 per cent and in 2010 it amounted at 78.4 per cent (see Mofair, 2009, 2013).

⁸During 1996–2002, when tendering was not compulsory but optional, between six and 19 million train-km were tendered p.a. This translates into a share between 1 per cent and 3 per cent of all regional rail passenger services in Germany (Link, 2004).

⁹The share of net contracts declined only slowly from 98 per cent in 1996 to 95 per cent in 2003, and to 82 per cent in 2010.

The results indicate, furthermore, that those federal states that awarded contracts of a longer duration and a lower size (expressed in train-km p.a.) spent their subsidies more efficiently. The efficiency-enhancing effect of longer contract duration contradicts the finding from Affuso and Newbery (2002),¹⁰ but is in line with the ongoing critiques of rail operators which are not able to amortise rolling stock within shorter contract periods, and which have consequently to negotiate a higher subsidy. The positive effect of smaller contracts seems to be counter-intuitive. However, it might reflect the existence of a considerable number of very large contracts granted to the incumbent DB during the period of analysis, in particular for reasons of practicability at the beginning of the regionalisation (lack of new entrants with sufficient capability to operate large contracts), but also in 2003 where the federal states awarded large service contracts with DB in a second round without competitive tendering. In combination with the direct awarding and the aforementioned problem of information asymmetries in net contracts, these large contracts might have negatively affected the efficient use of public money. However, it should be noted that the size of the coefficients for both the duration and the size variables is rather small. In addition, they have to be interpreted within the range of contract duration and contract size in German contracts during 1996–2010 (mean contract duration: 9.6 years; mean contract size: 5.3 mill. train-km p.a.), and do not imply that a continuous increase of contract duration and decrease of contract size would continuously increase efficiency.

As expected, a higher population density has a positive effect on the efficient use of subsidies, although this variable is only significant in the random effects model. Finally, the positive sign for the variable *D_East* (a dummy variable indicating that the federal state is an East German state) confirms the results from the disaggregated efficiency scores that East German states use their subsidies less efficiently than West German states. Explaining this effect requires taking into account the transformation process in East Germany. In the 1990s, the East German states still possessed a very dense rail network with a decreasing patronage as a consequence of individual motorisation, changing settlement structures, negative demographic effects, and high levels of unemployment.

A dummy variable for urban areas was not significant and is not reported here.

6.0 Conclusion

This paper has for the first time provided a systematic two-stage efficiency analysis for subsidising regional rail services in Germany. It considers the total of these services and the relevant data for the period from 1996 to 2010. As in any econometric application, the results naturally reflect the specific empirical situation under study — in this paper the German rail franchising during the period from 1996 to 2010. Bearing this note of caution in mind, the following conclusions can be drawn. First, the findings suggest that a higher share of tendered rail services would significantly increase the efficiency in using the available public funds to provide these services. Second, the paper provides evidence that a higher share of gross contracts has a positive impact on efficiency. One potential

¹⁰Interestingly, the mean of contract duration in Affuso and Newbery (2002) is 9.1 years, comparable to the mean contract duration of 9.6 years in this paper.

reason for this is the fact that the majority of net contracts was awarded without any competitive tendering, and the effect could therefore be overlapped with the effect of lacking competitive pressure. A further reason is the information asymmetry regarding patronage and fare revenues which has so far disfavoured PTAs in net contracts, and which might have prevented market entry of new entrants.

Third, the analysis suggests that contract duration should be longer than currently (on average 9.6 years) to guarantee a more efficient use of public funds, a finding which confirms debates and critiques in Germany on too short contract periods for amortising rolling stock. Furthermore, smaller contracts than in the period of analysis (on average 5.3 million train-km p.a.) would potentially increase efficiency, although it should be borne in mind that there is certainly a minimum contract size below this efficiency that is negatively affected.

The analysis has also shown that rail subsidies are less efficiently used in East Germany than in the Western part of the country. Potential reasons for this are the negative demographic trend in East Germany and the transformation process. It has to be discussed at the political level to what extent East German states can achieve a more efficient use of subsidies in future, and to what extent the public interest of providing a (defined and agreed) level of regional rail services justifies a higher subsidy per train-km and per passenger-km in East Germany than in the Western part of the country.

Finally, there are some issues for future research. First, an increase of efficiency in the sense of reducing subsidies and maintaining rail services should be reflected in an increase of companies' efficiency in using their factor inputs. With new data collection, in future this could be tested and serve as a complementary evidence of the effect of rail franchising. Second, the analysis presented here does not consider the quality of regional rail services. Quality indicators such as traveller satisfaction, punctuality, and cleanness of trains would be an important variable in the first stage DEA, but are currently only available for a small set of federal states and a short time period. The federal states have been in the process of establishing contract controlling and reporting systems which include quality indicators, and future research should take quality into account in modelling.

References

- Affuso, L. and D. Newbery (2002): 'The impact of structural and contractual arrangements on a vertically separated railway', *The Economic and Social Review*, 33, 83–92.
- Beck, A. (2011): 'Barriers to entry in rail passenger services: empirical evidence for tendering procedures in Germany', *European Journal of Transport and Infrastructure Research*, 11, 20–41.
- Bitzan, J. (2003): 'Railroad costs and competition', *Journal of Transport Economics and Policy*, 37, 201–25.
- Cantos, P., J. M. Pastor, and L. Serrano (1999): 'Productivity, efficiency and technical change in the European Railways: a non-parametric approach', *Transportation*, 26, 337–57.
- Cantos, P., J. M. Pastor, and L. Serrano (2010): 'Vertical and horizontal separation in the European railway sector and its effects on productivity', *Journal of Transport Economics and Policy*, 44, 139–60.
- Coelli, T. and S. Perelman (1999): 'A comparison of parametric and non-parametric distance functions: with application to European railways', *European Journal of Operational Research*, 117, 326–39.
- Coelli, T. J., D. S. P. Rao, C. J. O'Donnell, and G. E. Battese (2005): *An Introduction to Efficiency and Productivity Analysis* (second edition) Springer US, New York.
- De Borger, B. and K. Kerstens (1996): 'Cost efficiency of Belgian local governments: a comparative analysis of FDH, DEA, and econometric approaches', *Regional Science and Urban Economics*, 26, 145–70.

- Demsetz, H. (1968): 'Why regulate utilities?', *Journal of Law and Economics*, 11, 55–66.
- Driessen, G., M. Liejesen, and M. Mulder (2006): 'The impact of competition on productive efficiency in European railways', CPB Discussion Paper no. 71, The Hague.
- Farrell, M. J. (1957): 'The measurement of productive efficiency', *Journal of Royal Statistical Society*, Series A, 120, 253–81.
- Greene, W. H. (1993): 'The econometric approach to efficiency analysis', in H. Fried, C. A. K. Lovel, and S. S. Schmidt (eds), *The Measurement of Productive Efficiency and Application*, Oxford University Press, New York.
- Greene, W. H. (2002): *Econometric Analysis*, Prentice-Hall, New Jersey.
- Greene, W. H. (2008): 'Limited dependent variables — truncation, censoring and sample selection'. Available at <http://people.stern.nyu.edu/wgreene/Lugano2013/Greene-Chapter-19.pdf> [Accessed 20 April 2014].
- Growitsch, C. and H. Wetzel (2009): 'Testing for economies of scope in European railways: an efficiency analysis', *Journal of Transport Economics and Policy*, 43, 1–24.
- Hunold, M. and C. Wolf (2012): 'Competitive procurement design: Evidence from regional passenger railway services in Germany'. Paper presented at the Kuhmo Nectar Conference on Transport Economics, Berlin.
- Ivaldi, M. and G. J. McCullough (2001): 'Density and integration effects on class 1 US freight railroads', *Journal of Regulatory Economics*, 19, 161–82.
- Jensen, A. and P. Stelling (2007): 'Economic impacts of Swedish railway deregulation: a longitudinal study', *Transportation Research E*, 43, 516–34.
- Lalive, R. and A. Schmutzler (2008): 'Entry in liberalised railway markets: the German experience', *Review of Network Economics*, 7, 37–52.
- Lalive, R. and A. Schmutzler (2011): 'Auctions versus negotiations in public procurement: which works better?', CEPR Discussion Paper no. 8538.
- Link, H. (2004): 'Rail infrastructure charging and on-track competition in Germany', *International Journal of Transport Management*, 2, 17–24.
- Link, H. (2013): 'Unbundling, public infrastructure financing and access charge regulation in the German rail sector', *Journal of Rail Transport Planning & Management* 12/2012, 2, 63–71.
- Link, H. and R. Merkert (2011): 'Success factors and pitfalls of the German rail franchising approach', *International Journal of Transport Economics*, XXXVIII, 171–98.
- Merkert, R., A. Smith, and C. A. Nash (2010): 'Benchmarking of train operating firms — a transaction cost efficiency analysis', *Transportation Planning and Technology*, 33, 35–53.
- Mituzani, F. and S. Urinishi (2010): 'Does vertical separation reduce cost? An empirical analysis of the rail industry in OECD countries', Discussion Paper 2010-48, Kobe University.
- Mofair (2009): *Wettbewerber-Report Eisenbahn 2008/2009*, Berlin.
- Mofair (2013): *Wettbewerber-Report Eisenbahn 2013/2014*, Berlin.
- Mulder, M., M. Lijesen, G. Driessen, and D. Van de Velde (2005): 'Vertical separation and competition in the Dutch rail industry. A cost–benefit analysis'. Paper submitted to the 3rd Conference on Railroad Industry Structure, Competition and Investments, 21–22 October, Stockholm.
- Mundlak, Y. (1978): 'On the pooling of time series and cross sectional data', *Econometrica*, 56, 69–86.
- Nash, C. A. and A. Smith (2006): 'Passenger rail franchising — British experience'. In ECMT workshop on competitive tendering for passenger rail services, 12 January, Paris.
- Nash, C. A., J. E. Nilsson, and H. Link (2013): 'Comparing three models for introduction of competition into railways', *Journal of Transport Economics and Policy*, 47, 191–206.
- Olsen, R. (1978): 'A note on the uniqueness of the maximum likelihood estimator in the Tobit model', *Econometrica*, 46, 1211–15.
- Pollit, M. and A. Smith (2002): 'The restructuring and privatisation of British Rail: was it really that bad?', *Fiscal Studies: The Journal of the Institute for Fiscal Studies*, 23, 463–502.
- Simar, L. and P. W. Wilson (1998): 'Sensitivity analysis of efficiency scores: how to bootstrap in non-parametric frontier models', *Management Science*, 44, 49–61.
- Simar, L. and P. W. Wilson (2000): 'A general methodology for bootstrapping in non-parametric frontier models', *Journal of Applied Statistics*, 27, 779–802.
- Tobin, J. (1958): 'Estimation of relationships for limited dependent variables', *Econometrica*, 26, 24–36.

- VZBV (2010): *Der Schienenverkehr braucht klare Ziele, mehr Transparenz und Wettbewerb, eine effektive Regulierung und eine wirksame Verbrauchervertretung*. Hintergrundpapier des Verbraucherzentrale Bundesverbandes VZBV. Available at www.vzbv.de/mediapics/schienenverkehr_hintergrundpapier_2010.pdf [Accessed 14 July 2014].
- Woldridge, J. M. (2010): *An Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, MA.