Welfare effects of vertical separation in the Dutch railways

Mark Lijesen[1], *Machiel Mulder and Gertjan Driessen*

Netherlands Bureau for Economic Policy Analysis, CPB, The Netherlands

Paper submitted to 4th Conference on Applied Infrastructure Research, Berlin, 8 October 2005

ABSTRACT

This paper assesses costs and benefits of structural changes in Dutch railways implemented in the late 1990s. Although the performance of Dutch railways has somewhat improved in recent years, the efficiency of the introduction of competition and separation of the industry is still subject to debate. Internationally, net welfare effects of vertical separation in the rail industry are highly debated as well. Vertical separation would only generate a positive net benefit if it has a strong impact on competition. If competition is hardly possible - due to, for instance, low traffic density - the benefits of vertical separation would be negligible while the costs of reduced economies of scope could be considerable. In order to contribute to that debate, we analyse both the magnitude of economies of scope in Dutch railways and the impact of institutional changes on efficiency. The econometric analysis of costs and output of the vertically integrated Dutch railway incumbent suggests that economies of scope were present. These advantages of vertical integration were likely lost by the introduction of vertical unbundling. The introduction of competition in freight transport has increased both efficiency and performance. Passenger transport, however, has had difficulties in realising historical performance levels. The ambiguity of our result implies the need for further research.

Keywords: railways, competition, vertical separation, efficiency; JEL-codes:L51, D61, R48

---

[1] *Address corresponding author:* CPB, P.O. Box 80510, 2508 GM The Hague, The Netherlands, e-mail: mgl@cpb.nl ; homepage: www.cpb.nl/goto/mgl. This paper present preliminary results of a research project on competition in railways which is not yet finished.

1. INTRODUCTION

1.1 Institutional changes in railways

About 15 years, ago, the European Union initiated a process of institutional changes in European railways in order to improve the position of this mode of transport. In the years before, railways had lost market share, both in passenger and in freight, produced less output as well as increasingly needed financial support by governments (Nash et al., 2004). The institutional changes prescribed by the European Union included separation between network management and operation, third party access to the infrastructure, rules regarding the allocation of slots, and rules for pricing of network use. The goals of these changes were to increase the performance of the industry, i.e. improving the market share of railways in traffic, to enhance efficiency of the industry as well as to reduce government involvement.

Following the EU directives, the Dutch government imposed several changes on the Dutch railway industry (see Van de Velde, 2000). The former government agency became more independent, giving it freedom to determine investments and personnel policies, among others. Although full privatisation was the ultimate aim, nowadays this is not pursued anymore. In addition, the formerly fully-integrated railway firm was separated in several units. The owner of the infrastructure was separated from other activities while freight traffic was split from passenger traffic. Initially, the vertical separation was only organisational, i.e. consisted of different departments within one company, but more recently, both type of activities are institutionally separated, i.e. they are conducted by different companies. Another major change comprised the introduction of competition. As experiences with competition *on* the tracks in the passenger market failed, now only competition *for* the market exists, with competitive tendering in the regional transport market and negotiated contract with the operator on the main lines. In the freight market, however, competition *on* the tracks has successfully been introduced.

1.2 Current structure of Dutch railways

On the main part of the network, one firm, NS, is providing passenger services. The licence for the main national lines is given to the firm for the period up to 2015. This operator belongs to the NS-

Group which include several other rail-related firms: NedTrain develops and maintains rolling-stock (also on request of other operators), NS Stations exploits stations, NS Vastgoed exploits real-estate in the direct neighbourhood of the stations and NS International operates international lines.

In two regions, other operators provide passenger services: Syntus[2] in an eastern part of the Netherlands and NoordNed[3] in a northern part of the Netherlands. Both operators offer regionally-integrated public-transport services, including train and bus. All passenger operators have licenses to operate well-defined lines for a certain period of time. The two regional licences are allocated by competitive tendering.

Contrary to passenger transport, freight transport shows competition *in* the market. Here, the institutional changes have resulted in increased competition. Currently, eight freight operators offer services using the Dutch rail infrastructure. The largest of these firms, Railion Nederland BV, which is the successor of the freight division of the former national rail company, now has a market share of more than 80%, implying that this firm has lost almost 20% of the freight market. Railion Nederland BV is a subsidiary of Railion which is active in several European countries. So, at European level a process of concentration has emerged.

The network firm (ProRail) is responsible for management of the infrastructure, traffic management and optimisation of the railways. In addition to the separation from the operator, this firm has undergone several changes, in particular the 100% outsourcing of engineering activities, maintenance and renewal.

1.3 Performance of Dutch railways

The main aims of the institutional changes were to improve the performance of the industry, i.e. to realise a growth in output, and to increase efficiency. Looking at the performance of Dutch railways, a

[2] This firm, established in 1999, is a joint venture of NS Holding, Connexxion Holding and Cariane Multimodel International (CMI).

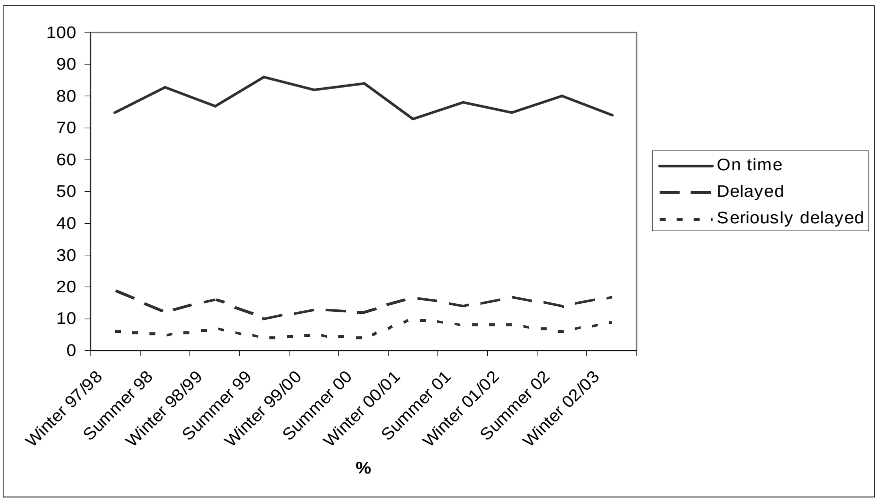
[3] This firm, established as a joint venture of NS and Arriva in 2003, is a 100% subsidiary of Arriva Nederland since 2003. The latter is a 100% subsidiary of Arriva International, a British firm which offers transport services in the United Kingdom and several other European countries.

mixed picture emerges. The network has recently recovered from the worsening performance in the mid 1990s, passenger transport has difficulties in realising historical levels, while freight transport produced a remarkable growth in output.

Passenger transport

The performance of passenger transport, measured in number of travel kilometres, changed since the early 1990s but did not result in a steady improvement: the total annual number of travel kilometres in 2004 about equals the number in 1991. As passenger transport by other modalities, in particular car transport, has increased strongly in this period, the market share of railways in the passenger transport market has declined significantly. One explanation for this deteriorated performance may be found in the worsened level of punctuality (see Van Gent, et al., 2005). At the end of the 1990s, the punctuality of this operator reached historically low levels. The percentage trains being on-time was about 70 while almost 10% of the trains were seriously delayed (see Figure 1.1). This worsening in punctuality was the result of postponed maintenance of the network, and investments in new rolling stock also remaining behind. These developments resulted from uncertainty created by the implementation of the institutional changes.

Figure 1.1 Punctuality of the incumbent train operator (NS)



Source: Rover, www.rovernet.nl

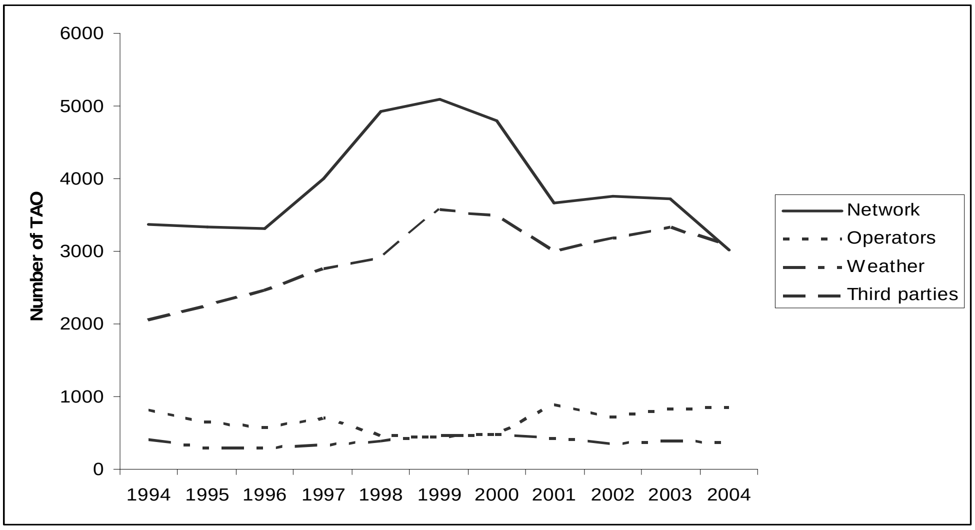
Freight transport

Since the mid 1990s, freight transport by rail in the Netherlands has shown a strong growth, both in historical context as compared to other European countries (Eurostat, 2005). In the years 1990 to 1995, annual total transport of goods by railway hardly changed, but it grew by approximately 50% in the remaining years of that decade. This pace of growth exceeded all other European countries. In the period 1990 - 2003, total freight transport by railway increased by 54% in the Netherlands, while only Portugal produced a growth figure of comparable magnitude (i.e. 44%). The growth of freight transport is largely realised in international transport; domestic freight transport increased by only a few percentages over the whole period. As a result, the relative importance of international transport in the Dutch freight railway industry has increased from 66% in 1990 to 78% in 2004.

Network

During the period the structural changes in the Dutch railway industry were implemented, the number of irregularities in train services increased rapidly (see figure 1.3). In 1999, the number of irregularities caused by network events was at an almost 50% higher level compared to the early 1990s.

Figure 1.3 Number of irregularities in train services by cause, 1994 - 2004



Source: ProRail (TAO stands for (in Dutch) “Treindienst Aantastende Onregelmatigheden”, i.e. irregularities disturbing train services.

This increase resulted from among others insufficient maintenance of the network in the first years after the outsourcing of engineering and maintenance activities. In more recent years, the performance of the network has improved resulting in a level of irregularities below the early 1990s. According to a representative of ProRail, this reduction in network irregularities is due to improved management of outsourced activities.[4] Increased monitoring of usage of tracks, which is needed because of infrastructure charges, seems also to have contributed to the performance of networks (as well as efficiency of train operations), as that encourages in-time maintenance activities. The number of irregularities caused by third parties, including accidents and suicides on the tracks, currently equals the number caused by network events. Note, however, that irregularities caused by network events generally have a larger impact on train services than other types of events. Consequently, this measure does not fully reflect the performance of the network owner.

1.4 Analysis of the performance

Several factors have contributed to development in performance of Dutch railways described above. Some are exogenous, such as macroeconomic growth, others are directly related to the structure of the industry. In this paper, we focus on the impact of structural changes on efficiency in order to assess the net welfare effects on these changes. Our analysis consists of answering the following questions:

a. has the vertical separation between network and operation caused a loss of economies of scope?
b. to which extent have the institutional changes affected incentives for efficiency improvement?
c. which efficiency improvements have occurred and to which extent are they related to the institutional changes?

In the remaining of this paper, we will consecutively give attention to these three questions. Before, we introduce key concepts in the economic analysis of institutional changes in the railway industry and summarise some literature on the economic impacts of vertical separation.

[4] Meeting at ProRail, d.d. 27 June 2005.

2. ECONOMIC ANALYSIS OF INSTITUTIONAL CHANGES IN RAILWAYS

2.1 Efficiency concepts

In the welfare-economic analysis of policy measures, such as institutional changes in an industry, efficiency is the key notion captured in two concepts: allocative efficiency and technical efficiency. Allocative efficiency refers to welfare effects of the allocation of goods and technical efficiency to the costs of supplying goods. If a good is not priced according marginal costs of supplying it, an allocative inefficiency exists. So, allocative efficiency is related to the way goods are priced and allocated. Technical efficiency is related to the incentives firms have to improve productivity. Both efficiency concepts have a static as well as a dynamic dimension. A market is dynamically efficient, if firms have incentives to increase technical efficiency, by reducing costs or improving quality of a good, and to pass on the benefits to consumers. When applying the welfare-economic approach to railways, attention has to be given to the characteristics of that industry.

2.2 Economics of railways

Railways is a network industry like telecommunication and electricity. The fundamental characteristic of network industries is the presence of a network infrastructure, such as pipelines in the gas industry, tracks in railways and copper lines in telecommunications. Network industries have three fundamental, mutually-related characteristics which are related to the presence of the network infrastructure and which make them different from other sectors. These characteristics are a) the presence of network externalities, b) the infrastructure being an essential link in the chain of activities and c) which coincide with substantial economies of scale. Although network industries share these common characteristics, the importance of each characteristic differs among these industries (see Pittman, 2003).

Network externalities

From the perspective of consumers, network externalities occur if “one person’s utility for a good depends on the *number* of other people who consume this good” (Varian, 2003). This holds in particular for the communications industry where each new consumer raises the value of the system to consumers already present. This is a positive network externality. Negative network externalities arise if aggregated demand, resulting from many individual decisions made by consumers, raise the load of the system so much that supply is unable to follow and, hence, system unbalance results. In railways, this kind of negative network externalities can arise during peak hours when demand exceeds available capacity. In addition, new agents benefit from the existence of a well-developed network (in their region) as it reduces the marginal costs of network extension. After all, in a well-developed network, extending the system to more locations within the same area causes relatively low costs due to the small distances which have to be covered.

Essential facility

The network infrastructure forms an essential facility in the industry meaning that the infrastructure is a necessary input for activities of sectors using the infrastructure. Train operators absolutely need tracks to offer their transport services, just as electricity producers need wires to transport power, and suppliers of telecommunication services need an infrastructure such as cables. The essential character of a facility depends, however, on the perspective from which a sector is viewed. A rail operator could use other means of transport, such as busses, if tracks are not available on certain distances. The definition of the essential facility affects the choice which components can to separated from each other. In the London subway, for instance, rolling stock is, just as the tracks, separated from operation, while in Dutch railways, the operators still possess the rolling stock.

Economies of scope

Strongly related to the essential-facility character of networks is the high level of interdependence between users of infrastructure, i.e. operators in the case of railways. Consequently, use of the infrastructure requires much coordination in order to prevent unbalanced gas systems, accidents on the tracks or black outs in the supply of power. According to BTRE (2003), the complexity of interaction between the infrastructure manager and train operators increases disproportionately .[5] These costs stem mainly from the rise in the number of operators, as every additional operator entering the industry

[5] Bureau of Transport and Regional Economics (BTRE), 2003, Rail Infrastructure Pricing: principles and practice, Commonwealth of Australia, Report 109.

results in a multiplication of the number of interfaces. These costs are magnified because of the need to harmonise objectives of the infrastructure owner and the operator. Consequently, close links between infrastructure activities and operational activities could cause economies of scope, i.e. integrating these activities in one firm could be more efficient than conducting these activities in separate firms. In railways, economies of scope appear to be relatively important. A significant part of the costs of rail operators are related to wheel maintenance and, therefore, depends on the wheel-rail interface. Moreover, technological improvements in railway are directed at this interface (Pittman, 2003). Vertical integration enhances an efficient way of dealing with this interface while in case over vertical separation specific arrangements have to be made to deal with the wheel-track interface.

Economies of scale

Network industries coincide with significant economies of scale due to the high level of fixed costs and (very) low marginal costs. If investments in a network infrastructure have been made, these costs are mainly sunk, i.e. these costs can not be recovered. The huge fixed costs and the scale effects related to it make it uneconomical to double networks in most countries. As a consequence, networks are often natural monopolies. In railways, economies of scale play a major role as here the share of fixed costs is significantly larger than in other network industries. Consequently, creating competition within this industry is fairly complicated, for instance because of the difficulties in determining access prices which take into account both allocative and dynamic efficiency (see Pittman, 2003).

Although the track owner has a monopoly within railways, this does not imply that it does not face any competition. In railways, cross-modal competition is the basic form of competition The intensity of this source of competition partly depends on physical characteristics of a region. In countries having rivers and coast lines, such as the United Kingdom, the Netherlands, Denmark, and Greece, transport of bulky goods by ships is often an efficient manner of transport. As a result, freight transport by railways in these countries face fairly fierce cross-model competition, although for some commodities, e.g. bulk commodities, and in some regions, e.g. in Australia and the USA, rail transport is the only economical mode. In passenger transport, cross-model competition appears to be much weaker.

Economies of density

In addition to the network-industry characteristics, railways has some other peculiarities. Besides economies of scope and economies of scale, both mentioned above, railways are subject to economies of density. This efficiency concept is related to the utilisation of infrastructure as well as rolling stock. If traffic volume increases, i.e. if the use of infrastructure or rolling stock increases, average costs decline up to the degree of utilisation where diseconomies of density emerges. The optimal level of utilisation of infrastructure depends on the characteristics of the rolling stock. If operators use trains with different characteristics - speeds, lengths, axle loads, etc. -, economies of density are limited. The more trains have the same characteristics, the higher utilisation of capacity is possible. The presence of economies of density in railways negatively affects the possibility to introduce competition between operators (Pittman, 2003).

2.3 Welfare effects of vertical separation

In the railway industry, net welfare effects of vertical separation are highly debated. (see e.g. OECD, 2004). Vertical separation would only generate a positive net benefit if it has a strong impact on competition. The benefits of vertical separation depend, therefore, on its impact on above-rail competition. Access of other operators to rail infrastructure generates pressure on the incumbent operator to improve services and reduce costs. If, as a result, the price of rail transport decreases, transport volume increases which could further reduce costs due to economies of density.

If competition is hardly possible - due to, for instance, a low traffic density - the benefits of vertical separation are negligible. Costs of this policy option, however, appear to be significant (see e.g. BTRE, 2003). Separation likely incurs significant costs due to loss of economies of density (less travellers per operator) and increased coordination costs (the complexity of interaction between the infrastructure manager and train operators increases disproportionately). Moreover, costs of utilisation of infrastructure capacity rise if new operators use trains with different characteristics - speeds, lengths, axle loads, etc., making it difficult to use the infrastructure intensively. In addition, in the case of separation, incentives to invest in rail infrastructure depend on the tariffs the manager may charge to operators, creating the risk of suboptimal investments. On the other side, investments of operators - in e.g. longer trains - depend on investments of infrastructure owners - in e.g. longer platforms. This mutual dependency may hamper investments. Finally, separation in railways may lead to asymmetrical incentives for the wheel-rail interface, negatively affecting investments in reliability as this strongly depends on this interface.

3. ECONOMICS OF SCOPE

3.1 Introduction

Economies of scope refer to the effect of simultaneous production of two goods on costs. A very good example of positive economies of scope is that of producing sheep’s milk and wool. If the production of these two goods is combined within a single production process, costs are much lower than in the case of separate production. Note that economies of scope are not related to shared fixed costs, as this is a form of economies of scale.

In the case of railways, economies of scope often refer to the joint production of freight and passenger transport. For the purpose of our analysis however, economies of scope between either passenger of freight transport on the one hand and rail maintenance on the other hand. If economies of scope are present here, they will probably arise because integrated decisions lead to more efficient outcomes than separate decisions. If, for instance, costs of rolling stock wear and tear are taken into account in track investment decisions, this may lead to lower total costs.

The remainder of this section is organized as follows. Section 3.2 provides some theoretical background, followed by a discussion of existing studies in section 3.3 and conclusions in section 3.4.

3.2. Theoretical background

Economies of scope exist if the joint production of two goods is performed at lower costs than the separate production of the same quantities of these goods. To state this mathematically, correcting for scale effects, we may formulate that positive economies of scope exist if the cost elasticity of joint production is smaller than the weighted average of the cost elasticities of the production of each good separately.

Let us look at a simple two-product quadratic cost function to further explore the concept of economies of scope.[6]

$$C = \alpha_1 Q_1 + \alpha_2 Q_2 + \alpha_{11} Q_1{}^2 + \alpha_{22} Q_2{}^2 + \alpha_{12} Q_1 Q_2$$

It is fairly straightforward to see from this equation that the parameter α_{12} tells us something about economies of scope. We can check this by a virtual separation of the cost function. The costs for producing product 1 only (i.e $Q_2 = 0$) are now:

$$C_1 = \alpha_1 Q_1 + \alpha_{11} Q_1{}^2$$

Similarly, the costs for producing product 2 only (i.e $Q_1 = 0$) are:

$$C_2 = \alpha_2 \ln Q_2 + \alpha_{22} Q_2{}^2$$

To check for economies of scope, we can add the costs of producing both goods separately ($C_1 + C_2$) and compare them to the costs of combined production (*C*). From this exercise, it is clear that $\alpha_{12} < 0$ implies positive economies of scope.

[6] Factors prices and the constant are ignored here for simplicity. We use a quadratic form instead of the more popular translog specification for ease of interpretation, especially with zero values. Our conclusions hold for translog costs functions as well.

In the case of railways, one can think of another explanation than the one above however. Let us translate the two-products case two railways and define Q_1 as the output of the operator (e.g. seat kilometres) and Q_2 as the length of tracks. Now in the case of railways, it is pretty safe to say that the length of the network is more or less fixed, especially in the short run.

For illustration purposes, we impose that $\alpha_{12} < 0$, which we identified as positive economies of scope above. Now if Q_1 increases (at any given and fixed value for Q_2), we see that the costs increase by less than was to be expected based on the cost parameters for product 1, α_1 and α_{11}. In network sectors, this phenomenon is known as economies of density. Increasing the output on a given network increases density. In the case of positive economies of density, the increase of density implies a decrease in average costs. The ambiguous interpretation of the cross-term in the cost function renders it impossible to draw a robust conclusion on economies of scope based on cost functions alone.

3.3 Existing studies

Economies of scope between tracks and railway operation are generally not a prime research topic of empirical costs studies of railways. Two studies do however report a cross-term of railway operation output and an indicator for track length, which can be interpreted as an empirical measure for economies of scope between tracks and railway operation.

Preston (1994) estimates a translog cost function for 15 Western European Railways for the period 1971-1990, using total train kilometres (passenger and freight) as an indicator for railway operation output and length of route as an indicator for infrastructure. Preston finds a cross-term that is negative and statistically significant.

Pels *et al.* (2004) estimate a translog cost function for the Dutch railway sector for the period 1951-1993. They use seat miles as an output indicator for railway operations. Although Pels *et al.* (2004) do not take network length into account, they use the value of capital in their analysis, an indicator that is both closely related to and highly correlated with network length. The cross-term between these indicators is positive and statistically not significant.

Note that both studies mentioned above are based on operating costs, implying that possible economies of scope related to capital costs are ignored. Furthermore, as we showed in the previous section, the interpretation of the cross-term is somewhat ambiguous, leading to uncertainty whether the effect found comes from economies of scope or economies of density

3.4 Conclusions

Our results indicate that there is at best weak evidence for positive economies of scope. Only two empirical studies have addressed the issue, and only one of them finds statistically significant results. These results are limited further because the interpretation of the cross-term in the cost-function may be attributed to either economies of scope or economies of density. Further more, since both studies focus on operational costs rather than all costs, we can not be sure that the cost functions tell the entire story. All in all, the empirical proof of economies of scope between tracks and railway operation is troublesome and should at least be backed by findings from case by case investigation of sources of economies of scope.

4. INCENTIVES AND EFFICIENCY

4.1 Theory

The process of restructuring, liberalisation and regulation can be seen as a process of improving the incentives to increase efficiency. In the past, the Dutch railway industry, as many others, was fully state owned, vertically integrated and not subject to competition (on or for the tracks). As is described above, the industry has changed strongly over the past decade. In this section, we describe theoretical relationships between different institutional choices and efficiency. In the next section, we analyse how institutional changes have affected incentives.

Regulation

In the regulation of a natural monopoly, as railways, three types of incentive schemes can be distinguished: high powered, intermediate and low powered. In the past, many incentive schemes consisted of cost-of-service regulation which belongs to the last category. This type of schemes does not give incentives to the operators (agents) to improve efficiency. As a result, these schemes have been replaced by price-cap regulation which is a high-powered incentive scheme as the firm is the residual claimant of all efficiency improvements. This regulation scheme also enables the regulator to transfer efficiency gains to consumers by periodically reducing the price cap. Pursuing these two goals, i.e. encouraging efficiency improvement and transferring rents to consumers, generally is subject to a trade-off. The determination of a price caps requires, therefore, careful attention as a too high level causes allocative inefficiencies as well as it gives the firm monopoly rents while a too low level threatens the viability of the industry and can negatively affect incentives for innovation. The optimal contract depends on the degree of information the regulator has (Laffont et al., 1993). If the regulator has perfect information about future costs, a price cap is preferable above cost-of-service regulation as the price can be set on the minimal-cost level while giving incentives to the operator to improve efficiency even more. If, on the other hand, the regulator faces large uncertainties about future costs while the operator has private information, cost-of-service regulation can be used to prevent the operators from making excessive profits by pursuing its own goals (which is called adverse selection). In some cases, alternative options seems to be preferable, such as yardstick competition where periodical price changes are based on exogenous cost movements leaving the firm the efficiency gains of performing better than the industry average (see e.g. Burns, et al, 2004; Farsi et al., 2005).

Ownership

Both theoretical and empirical literature on the relative efficiency of public and private ownership of regulated firms are inconclusive (Laffont et al., 1993). In case of public ownership, the government can easily impose adjustments to the firm in order to let it meet social objectives, while in the case of private ownership the government has to bargain with the private firm. On the other hand, this ability of the government can deter the management of the firm from introducing innovations or making other investments. After all, the management faces the risk the government will relocate the investments to alternative uses. This inefficiency of public ownership is even larger when government decision making is captured by interest groups. Private ownership of regulated firms can cause inefficiencies as a result of conflicting interests between shareholders and regulators. The inefficiencies in both ownership structures are related to the separation of ownership and control, i.e. the separation between principals and agents. In the regulation of an industry as railways this is one of the key issues. In section .., we analyse the incentives given by the contracts between state and railway operators.

Competition

If firms can be subject to competition, regulation and public ownership are not needed to enhance efficiency. In theory, competitive pressure generates optimal allocative efficiency, i.e. in a competitive market the price of a good equals marginal costs. Because of this mechanism, competition also increases productive efficiency on industry level, as only efficient firms are able to survive. The relationship between competition and dynamic efficiency is less straightforward as competition can both enhance and hinder innovation. On the one hand, in a competitive market firms have stronger incentives to innovate (or adapt innovations) as it creates the option to charge prices above marginal costs and earn monopoly prices. On the other hand, competition affects the sustainability of such profits as the presence of profits attracts new entrants. Generally, if firms operating in a competitive market expect to be able to capture the profits of innovations, the market will be dynamically efficient.

4.2 Dutch railways

This section offers a preliminary assessment of the impact of institutional changes on incentives for efficiency improvement in Dutch railways. We consecutively pay attention to incentives for passenger transport, freight transport, and network management.

Passenger transport

Operators in the regional passenger markets are subject to competitive tendering while the operator running the main lines has obtained a license, i.e. a negotiated contract or concession, for the period up to 2015. In the concession for the main lines, performance targets are included combined with financial incentives.

Whether a scheme of competitive tendering or a negotiated contract give incentives to operators to improve efficiency depend on their designs. As is shown above, theoretically, the optimal contract encourages operators to reduce costs or improve quality of the product while allocating the benefits of those actions to consumers. One of the key components in rail concession contracts is the duration. A long duration of a contract or a concession gives the contractor information advantages weakening the bargaining position of the principal in the future or hindering efficient future competition by competitive hindering. Another potential disadvantage of long-lasting contracts is the limited flexibility (or relatively high transaction costs) to adapt contract conditions to changes in costs or market circumstances. When the contractor has not been responsible for investments, short-lasting contracts are preferable above long-lasting contracts. Management tendering is an example of such types of contracts. If a contractor is responsible for investments, however, short-term contracts may lead to underinvestment. This occurs if investments are not easily transferable to new contractors or if the contractor is uncertain about the value the principal will attach to the investments. Consequently, a trade-off exists between the benefits of competition through short-lasting contracts and the benefits of efficient levels of investments through long-lasting contracts.

As is mentioned above, in the Netherlands, the contractor for the main lines has obtained a long-lasting contract (15 years) prescribing performance targets on many issues. In practice, the design of contracts is hampered by information asymmetry resulting in adverse selection. The government faces an information handicap on what realistic performance aims are (Van de Velde, 2000).

Freight transport

Freight operators have been subject to competition, both intermodel and intramodel, for a number of years. This competition gives incentives to the operators to improve efficiency.

Network

In the case of vertical separation, incentives to invest in rail infrastructure in order to increase efficiency depend on the tariffs the manager may charge to operators.

In order to answer the question whether the structural alterations affect the performance of network management, we have to analyse changes in incentives given to the network manager. As a state-owned company not subject to competition, ProRail's incentives still fully depend on regulation and responsibilities allocated to its management. Regarding the latter, incentives to improve efficiency seem not be strong as the company's budget is based on annual political decisions, i.e. ProRail has an input-based budget. In the near future, ProRail will receive performance targets as well as more budget responsibilities. Currently, however, the main incentive for the network owner to improve performance is the threat of more political involvement.

One of the key issues is the pricing of the use of infrastructure. An integrated railway firm would be able to set prices at levels which depend on the characteristics of consumers of railway services. This Ramsey pricing method enables the firm to recover its costs without reducing allocative efficiency. A separated infrastructure owner is hardly able to use this pricing technique as his customers are the operators, not the consumers of train services. In the Netherlands, ProRail charges operators a fixed amount per unit of infrastructure use, which will increase annually in the near future. This method enables ProRail to recoup investments in infrastructure but it also results in a loss of allocative efficiency as prices will exceed both marginal costs and willingness to pay for some consumers.

5. TECHNICAL EFFICIENCY

5.1 Introduction

This section addresses the issue of technical efficiency in railways. Section 3.2 discusses and compares the different methods and approaches in the literature to measure efficiency. As such, first an overview of the various methods will be given, followed by applications on railways. Finally paragraph 3.3 deals with the impact of separation on Dutch railway efficiency. On the basis of firm data of the operator's active in the Dutch railway industry an efficiency analysis is performed. Additionally, the results of this analysis are presented in this paragraph.

5.2 Methodology

This section addresses and compares the various approaches and techniques for measuring productive efficiency.[7] Essentially, four principal methods can be distinguished, namely,

1) total factor productivity (TFP) indices;
2) data envelopment analysis (DEA);
3) deterministic econometric production models;
4) stochastic frontiers analysis (SFA)

For each type, we discuss their use and relative merits. The methods differ according to the type of measures they produce, the data they require, and the assumptions they make regarding the structure of the production technology and the economic behaviour of decision makers. We now provide a brief description of each method.

The first and third approach are most often applied to aggregate time-series data and provide measures of technical change and/or TFP. Both of these methods assume all firms are technically efficient.

[7] This paragraph draws heavily upon a book by Coelli, Prasado Rao and Battesse (1998) and a recent paper by Rivera-Trujillo (2004).

Methods two and four, on the other hand, are most often applied to data on a sample of firms (at one point in time) and provide measures of relative efficiency among firms. These latter two methods do not assume that all firms are technically efficient. Note that this ability to control for inefficiency is important, however, one has to acknowledge that this inefficiency measure is merely a relative notion.

Alternatively, the methods can be categorized according to whether the method involves economic estimation of parametric functions. More specifically, it is necessary to specify a particular functional form (e.g., Cobb Douglas, translog, etc) to describe the technology or efficiency frontier. Methods that use parametric functions (methods 3 and 4) are termed “parametric”, while methods that do not use them (methods 1 and 2) are known as “non-parametric”. The former are estimated by using econometric (statistical) methods, while the latter are assessed by applying mathematical programming.

Total factor productivity

The Total Factor Productivity (TFP) method measures changes in total output relative to the change in the usage of all inputs.. In most empirical applications, the Tornqvist index formula is used for purposes of output and input index calculations. TFP captures the effect of factor substitution when relative factor input prices change. Consequently, TFP includes two types of efficiency: 1) allocative efficiency (i.e., usage of inputs according to their relative prices) and 2) technical efficiency (i.e., converting inputs into outputs). TFP estimation is, however, subject to the problem of accurately measuring all inputs. It requires price indices for inputs and outputs. As a result, lack of adequate data makes it often difficult, if not impossible to perform a TFP analysis. In addition, TFP growth can be attributed to several factors (e.g., efficiency gains, technical progress), making it difficult to interpret its results.

Data Envelopment Analysis

Data Envelopment Analysis (DEA), uses linear programming methods to estimate a production frontier by fitting pieces of hyperplanes to envelop an observed set of data formed by the inputs and outputs. Efficiency measures are obtained by estimating the distance of the observations relative to the enveloped surface. DEA is a deterministic approach. In other words, it does not take statistical noise into account. It is not necessary to specify a particular functional form with DEA, therefore this approach is less sensitive to specification errors. Unlike parametric techniques, DEA is able to handle multi-output efficiency measurements. A major disadvantages of DEA is that it does not distinguish between efficiency and statistical noise. For this reason, the results of the analysis are very sensitive to outliers in the data set. An outlier can and does change the apparent efficiency of the other units. Furthermore, with DEA it is not possible to test whether the technical efficiency is statistically significant.

Econometric production models come in a great variety of specifications and techniques. In this approach an error term is used to account for the technical inefficiency. Two different specifications are used, namely, deterministic and stochastic.

Deterministic Economic Production Models

First, the deterministic approach. These type of models mix random variations into the efficiency term. In this way all deviation from the frontier function is taken as a technical inefficiency. The most common econometric techniques applied to estimate these deterministic models are: the Corrected Ordinary Least Squares (COLS), Modified or Displaced Ordinary Least Squares (MOLS) and Maximum Likelihood Estimation (MLE).

Stochastic Economic Production Models

On the other hand, stochastic frontier analysis is employed to estimate technical efficiency. This second econometric approach does not only recognize inefficiency (deviations from the production frontier), it also recognizes stochastic elements outside the control of the producer, which is included by adding an extra error term in the frontier function. In order to estimate these models, it is important to specify a particular distribution for the random variables that are included in the error terms.[8] Further, stochastic models require market prices of inputs and outputs to estimate technical efficiency. However, this is not necessary in the case of production or distance function.

Partial productivity measures

In addition to the four principal methods discussed above a fifth measure is often applied to measure efficiency. These are partial productivity measures. Productivity variation arises from different sources: differences in efficiency, economics of scale and density, differences in network characteristics, and other exogenous factors, that affect performance (e.g., composition of traffic, quality of service, geography) and/or technological changes. To induce the effect of productive efficiency, one must remove the effect on productivity caused by the differences in operating environment and exogenous factors (Oum, Waters II, and Yu, 1999). Some of the methods discussed above take this approach. Partial measures do not take these external factors affecting productivity into account. Therefore, the results are to be regarded with due caution, as they can be inaccurate and/or unreliable. Oum, Waters II and Yu (1999) define partial productivity measures as followed: "Partial productivity measures generally relate a firm's output to a single input factor...the list of potential measures is almost endless. These types of measures are easy to compute, require only limited data, and are intuitively easy to understand. They have widely used by both academics and industry analysts." These types of measures can be financial or physical measures of performance. The authors warn that the productivity of one input depends on the level of other inputs being used. Therefore, high productivity performance of one input may come at the expense of low productivity of other inputs. Another problem is the selection of an adequate output measure. There are many different dimensions of railway output. Hence use of a single output could bias performance comparisons. Despite these shortcomings, the authors still believe in the usefulness of partial productivity measures, as they "can provide useful insights to causes of high and low productivity, and thus provide practical guidance for identifying productivity problems. Partial measures are useful in comparisons of

[8] For variables that represent statistical noise, most studies assume it to be independent and identically distributed with mean zero and constant variance. For the inefficiency term, different options have been applied, for instance, half normal, truncated normal, and the exponential distribution

performance across firms operating in similar operating environments, or over time within a firm when the operating environment and input prices remain relatively stable."

Conclusion

Concluding, a wide range of efficiency measurement approaches have been discussed in this paragraph. Theoretically, econometric stochastic frontier models are most preferable, since, first of all, this approach does not assume that all companies are efficient. Second, SFA has the advantage that it can separate inefficiency from statistical noise. Finally, no specific behavioural objectives are imposed. While, the advantage of SFA over the other methods is clear, it still remains a question whether the technique can actually be applied, as it has strong data requirements. In the following paragraph, we analyse studies that have applied these methods on railways.

5.3 Efficiency analysis of railways: a concise review of literature

In the former paragraph, we described the various approaches for measuring efficiency. All of these approaches have been applied in the literature on estimating efficiency of railways. This paragraph provides a concise overview of these studies, describing methodology as well as results. As a full treatment is beyond the scope of this paper, a selection has been made to provide the reader with a flavour of what is on offer.

The first study we address is by Gathon and Perelman (1992). The authors were among the first to apply SFA to railways. They estimate a factor requirements frontier for 19 European railways using a paned data approach, implicitly assuming the existence of complementarity (fixed proportions) between all the main inputs in production.[9] [10] Technical efficiency is assumed to be endogenously determined. Furthermore, special attention is devoted to an autonomy indicator representing managerial freedom with respect to authorities. They observe a high correlation between individual

[9] Gathon and Perelman consider this assumption to be plausible, since from an empirical analysis they find that labour expenses account for about 90% of the variable cost for all the railways. This result justifies, according to the authors, that the substitution possibilities between labour and energy are highly limited.

[10] A factor requirements function is a production technology in which a single input is expressed as a function of a number of outputs.

technical efficiency and this institutional indicator. Interestingly, higher load factors (measured as the number of passengers/tons by train) reduce efficiency. That is, for a given level of production (measures as the number of train-km), higher demand leads to more input requirements (for instance, more employees). This relationship points at diseconomies of density in the use of trains.

A few years later, the same author and Pestieau (Gathon and Pestieau, 1995) decomposed productive efficiency into a management and a regulatory component. They estimate a translog production frontier to compute these efficiency indicators using DOLS. This method enables the authors to investigate for which part of efficiency slack management can be attributed, and for which part - beyond the control of the rail companies - the governments are responsible. The authors apply this procedure to 19 European railways over the period from 1961-88. Following Nishimizu and Page (1982), they argue that productivity gains are to be divided into two components, technical progress and efficiency changes, having different determinants.[11] Their results show that the Netherlands (NS) has the highest efficiency score in this period. In accordance with the previous paper, the authors find that managerial autonomy is an important determinant of the government owned railway's performance.

Oum and Yu (1994) perform a comparative efficiency study of the OECD countries' railways. Like the last two studies, this study predates the reforms in Europe. Their data deals with the period 1978-89. The aim of this study is to identify the implications of public subsidy and the degree of managerial autonomy in technical performance. The authors estimate technical efficiency by using a DEA model, assuming constant returns to scale. Two alternative output measures are used: 1) revenue-output measures (passenger-/ton-km) and 2) available output measures (passenger/freight train-km). In order to estimate the effects of policy and other variables beyond the control of management, a Tobit regression model is applied. The main (new) finding of this paper was that railway systems with high

[11] Efficiency is linked to the quality of management and to the institutional setting of railway operations, whereas technical progress is linked to R&D. Gathon and Pestieau estimate technical progress from the coefficients associated with a trend variable in their model.

dependence on public subsidies are significantly less efficient than similar railways with less dependence on public funds.[12] Again, the NS is among the most efficient performers.

Cowie and Riddington (1996) examine the methods of assessing rail efficiency. Their analysis of using these different methods to estimate efficiency has led them to be very critical. Especially DEA is not recommended: "We could not recommend any reliance on results generated by this measure". In addition, the authors note that as there is effectively no international trading and no common accounting practise, comparative international efficiency is best based on physical measures rather than value measures.[13] Cowie and Riddington, commenting on the studies previously mentioned, argue that there are clear reasons why the Dutch railways are more efficient than the Austrian railways (usually very low efficiency). According to them, this result follows from the high utilization of the infrastructure in the Netherlands. Finally, the authors point at the problem that the results of the studies investigated in their article do not correlate with each other (except from the Dutch vs. Austrian regularity).

Using data on 17 European railway companies during 1988-93, Coelli and Perelman (2000) estimate multioutput distance functions using COLS. This approach is advocated because it avoids making unrealistic assumptions of firm behaviour, while at the same time it is able to handle the multiouput nature of railways. The distance function results are compared with those obtained from single-output production functions. They find that the results differ substantially between the two methods. As a result, the authors doubt the reliability of the single-output methods.

A paper by Cantos, Pastor and Serrano (2000), investigates the importance of output specification. Their results show that alternative output specifications lead to different results. Nonetheless, these

[12] Following Nash and Rivera-Trujillo (2004), a word of caution is in order. It could be that the direction of causality is the other way around. That is, inefficient railways require high subsidies to survive, whilst high costs and low productivity might be the result of public service obligations to provide services such as peak commuter services which are costly but socially desirable.

[13] As most European railways are not free to operate on purely commercial terms, the output measure should therefore not only reflect the physical nature of output, but also the public service obligations and the product they are actually selling (e.g., quality of service).

differences can be brought substantially closer when output variables are corrected to account for the impact of the load factor. In this study, the authors use DEA.

As mentioned in the introduction of this paper, during the long period (1950s-1990s) of regulation, the railway companies improved their productivity level.[14] In contrast to this positive development, the financial state of affairs, defined as revenue over operating costs, got worse. These developments can only be explained by revenue growth falling behind the advances of productivity. Cantos and Maudos (2001) try to explain these facts by estimating both cost and revenue frontier functions (SFA). In so doing, the authors are able to calculate the losses associated with both cost and revenue inefficiencies. The technique used to estimate efficiency is maximum likelihood. Their empirical analysis shows the existence of significant potential losses of revenue. They reason that this is caused by the strong policy of regulation and intervention by the governments in this period.[15] Consequently, Cantos and Maudos propose a light form of price regulation and a service adapted to market conditions if the companies' financial burden are to be reduced.[16] They argue that it is time for a re-orientation from cost efficiency and productivity towards a policy focus on revenue. Policies such as concessions/franchises are regarded as positive, since they are compatible with the recommendations above (see our discussion on the design of contracts in section 4.)

There are relatively very few studies which extend efficiency analysis to the impact of rail restructuring in the 1990s. A study by Friebel, Ivaldi and Vibes (2003) investigates to what extend third-party access, independent regulation and the separation of infrastructure from operations affects railway performance. The authors find that reforms that have efficiency improving effects are implemented sequentially, while reforms introduced in a package have at best neutral effects. Moreover, their results show that full separation is not a necessary condition for increasing efficiency.

[14] This period stretches from the period of nationalisation in the 1950's to the deregulation measures undertaken in the 1990s.

[15] This argument can be explained as follows. Tightly regulated companies are not free to operate their companies on a purely commercial basis. As such, despite significant productivity improvements, companies may not increase their fares. The improvement in productivity may result in additional volume of traffic which reduces the average costs as result of economies of density, but this benefit may be insufficient to recover the investments generating the productivity improvements. As a result, the firm may face financial difficulties in spite of improved efficiency. Relaxing regulation, in particular regarding prices, would enable the firms to raise prices in order to increase financial returns. As traffic demand if highly inelastic, i.e. consumers have a high willingness to pay for train services, a rise in the price will hardly result in a reduction of traffic volume.

[16] Except for service that are socially desirable but not economically viable.

This result seems to conflict with the firm believe of many policy-makers. Interestingly, Friebel, Ivaldi and Vibes find that all smaller countries, except for the Netherlands, have been able to keep or raise their efficiency levels. In addition, they argue that better research is needed. Especially, quality measures of output are not available.

A recent study by Lan and Lin (2004) examines railway efficiency, effectiveness and productivity, while controlling for environmental effects, data noise and slacks. In order to make these adjustments and to overcome the shortcomings of traditional DEA models, the paper proposes a four-stage DEA approach. Due to the non-storable nature of railway services, technical efficiency (a transform from outputs from inputs) and technical effectiveness (a transform of consumptions from inputs) represent two distinct measurements. The authors find that the major decline of the rail industry (market share of passenger rail transport for the EU has declined from 32% in 1970 to 12% by 1999) should not be attributed to rail's poor performance in technical efficiency or service effectiveness, rather it is the consequence of higher level-of-service of other modes. In fact, their results indicate that the rail industry has a positive progress in recent years (1995-2001).

A SFA approach is applied by Rivera-Trujillo (2004). This methodology is justified by the arguments that it does not require a high availability and quality of data (which is complicated by rail reforms) and specific behaviour (e.g., cost minimisation). Furthermore, in order to take into account the multioutput characteristic of the rail industry, an input distance function is used. The study concentrates on freight transport, which is the most important segment in North and South American railways. The results show that a great part of productivity improvement was due to technological change rather than technical efficiency change. Rivera-Trujillo notes that further research is needed to the selection and specification of the variables in order to obtain internationally agreed performance measures in the rail industry, as well as, on the whole period in which the recent rail reforms took place to determine their degree of success.

Taken together, various methods have been applied in different ways. As a result, it is difficult to compare the results of the studies. Measuring technical efficiency is a difficult task with many shortcomings. Therefore most studies should be taken with due caution and results are to be interpret prudently. However, in general a picture emerges that the Netherlands (NS) have one of the most efficient railway companies in the world. Yet, this edge is deteriorating in the last decade. Other rail companies are "catching-up".

One of the most relevant findings for this paper comes from Friebel, Ivaldi and Vibes (2003). They obtain the result that full separation of infrastructure from operations is not a necessary condition for improving railroad efficiency. Note that, as with all the literature, this result must be taken with a grain of salt for various reasons.

This subsection outlined the most important literature that examined railroad efficiency. In subsection 5.4, we perform an efficiency analysis of the Netherlands.

5.4 Dutch railways

In this section, we concentrate on partial indicators of efficiency. Conducting an efficiency analysis as described above is not yet possible, as we were not able to collect the necessary data.[17] Both freight and passenger transport are investigated. In addition, a comparison with other countries is made.

Below, we present some partial productivity indicators. First, we have a look at passenger transport. Then, we investigate the developments of the freight sector. Finally, some concluding remarks are made. The results below are to be regarded with caution. The caveats of partial productivity measures discussed above in subsection 5.2 apply.

[17] After the reforms in the Dutch railway industry, companies became secretive about their performance. This is a common problem as very few studies are available which extend to the impact of rail restructuring beyond the 1990s.

Passenger transport

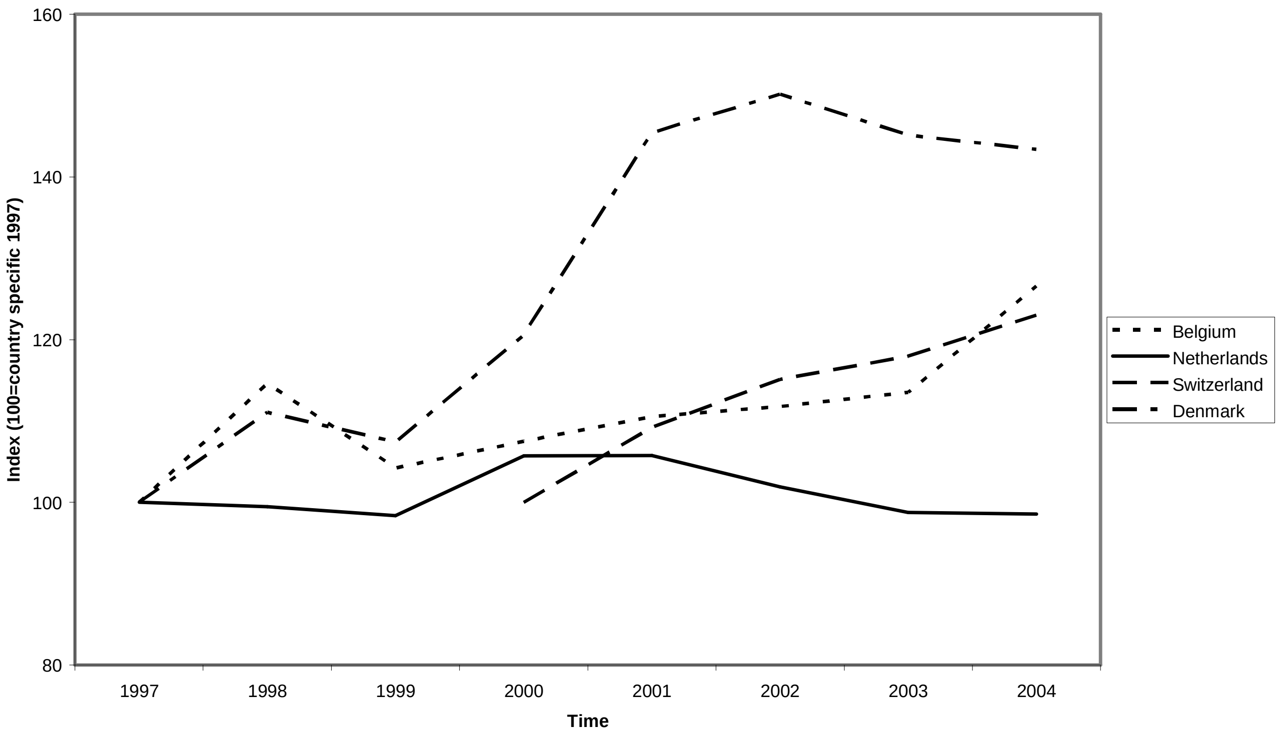
Probably one of the most frequently used productivity measures is labour productivity. Figure 5.1 displays this indicator. From 1987 towards 1992, labour productivity increased by more than 70 percent. After this leap, productivity has not increase much. In order to get a good perspective on the development of labour productivity, the volume of passenger traffic is also displayed in figure 5.1. The figure clearly shows a big growth in the early nineties. Hereafter, passenger traffic remained fairly stable. Thus, it seems that labour productivity has increased as a result of the same workforce producing a higher output. As this improvement took place before any major institutional changes were undertaken, we conjecture that the reforms did not affect labour productivity in a significant way.

Figure 5.1: Labour productivity and passenger rail transport in the Netherlands, 1980-2004

Source: World Bank Railway Database and annual reports of the Netherlands Railways (NS).

Having in mind all the caveats in comparing countries with each other, figure 5.2 shows the development of labour productivity in 5 countries in the period 1997-2004. The 1997 level of productivity of each country is indexed as 100. Surprisingly, all countries, except the Netherlands, have improved their labour productivity with more than 20 percent. An explanation could be that the Netherlands were already on a high level of passenger kilometres per employee, whereas the other countries had the opportunity to "catch-up". Rail restructuring has an effect on productivity if the railway companies are forced to increase productivity by competition or contracts. In the Netherlands, there has not been effective pressure on the incumbent (see section 4 on incentives).

Figure 5.2: Labour productivity in four countries, index (1997=100), 1997-2004



Source: annual reports of companies and national bureaus of statistics.

Reduction of staff is one of the instruments most frequently used to improve productivity. However, these staff cuts can lead to a drop in quality (e.g., punctuality), which is an important criterion of the success of reforms. In addition, there are other factors, besides institutional changes, which have a decisive, and possibly even greater, influence on productivity. Furthermore, it might still be too early to expect improvements in market performance as a results of structural changes. Overall, productivity and efficiency did not change after the reforms. We conjecture that this is could be the result of the Dutch policy to award the NS a concession until 2015 (see section 4). Nonetheless, this could change when we come closer to the end of this concession.

Freight Transport

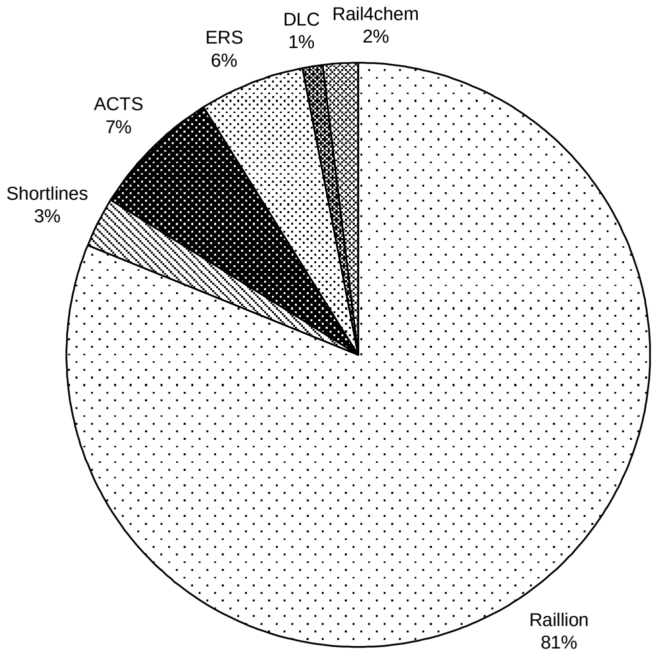
In contrast to passenger transport, freight transport has experienced severe competition from within the market as well as from other modalities (road and ship). In 1995, the Netherlands Railways (NS) was reorganised. As a result, passenger traffic and freight traffic were split up into separate divisions: NS Reizigers and NS Cargo, respectively. In 2000, NS Cargo became part of Railion Benelux, which is in turn a holding of Die Bahn. Meanwhile, competition appeared on the track. Freight transport grew remarkably, since 1993 by 80 percent. The globalisation of the economy and the increasing integration of the European economies have led to considerable growth in the sector.

Figure 5.3: Freight rail transport in the Netherlands, 1980-2004

Source: World Bank Railway Database and Statistics Netherlands.

It is interesting to see whether the former incumbent has lost market share due to the institutional changes. Besides Railion, seven other suppliers of freight transport by rail are active in the Netherlands. In figure 5.4, the market shares of the different suppliers are displayed. A few of the these suppliers are too small to show up in this figure. One can see that the reforms in freight transport have been relatively successful, as the former monopolist has lost a considerable percentage of its market share.

Figure 5.4: Market shares in freight transport in the Netherlands in 2004

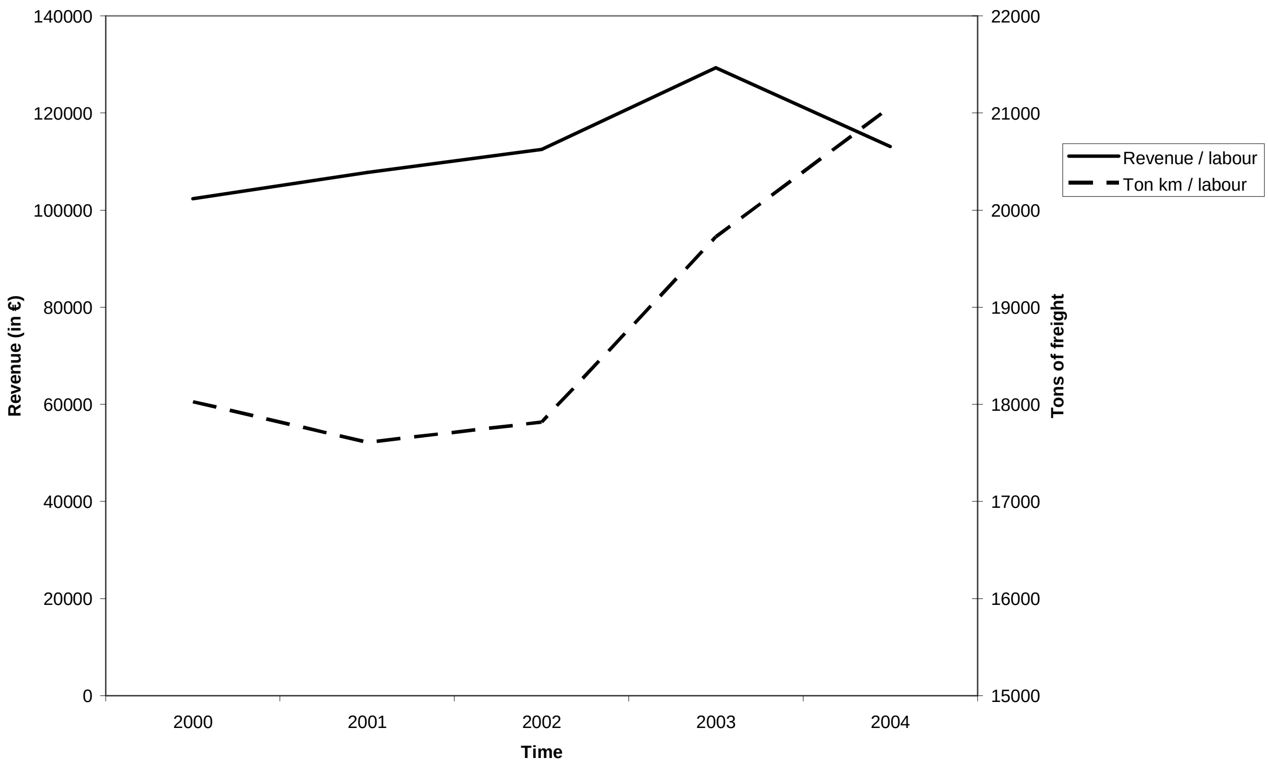


Source: Rail Cargo Information Netherlands

Next to these market developments, we investigate the productivity of the incumbent. If the structural changes have resulted in competition, we expect to see an improvement in productivity, as the company will struggle to keep its market share. In figure 5.5, the development of labour productivity of Railion Benelux is displayed. While labour productivity measured in physical units improved strongly in this period, labour productivity measured as revenue per worker increased only marginally.

This fits good with the previous story, Railion needs to improve its labour productivity, as competition evolves which puts tariffs and profit margins under heavy pressure.

Figure 5.5: Labour productivity Railion Benelux, 2000-2004



Source: annual reports Die Bahn

5.5 Conclusion

This section has examined the different methods to measure efficiency of railway operators. Due to lack of adequate data, a comprehensive efficiency analysis has not yet been performed. For this reason, we adopted a partial efficiency analysis. Despite its shortcomings, we were able to find some interesting results. For Dutch passenger transport, we find that the institutional changes have not improved the efficiency of the main operator, the Netherlands Railways (NS). As regard to freight transport, structural changes certainly have their effects on this market. Productivity increased and the market share of the incumbent declined by more than 20 percent.

6. Discussion

The results of our research indicate that the net benefit of the structural changes are probably positive in the case of freight transport. For passenger transport, the result is indecisive, as we can not give a robust quantification of possible loss of economies of scope. The ambiguity of this result implies the need for further research. Our analysis so far suggests that further research should be aimed at the following issues:

1. Economies of scope. The results so far are ambiguous, whereas this may be a key element in assessing the costs of vertical separation.

2. Coordination costs also form a part of the costs of vertical separation and need to be assessed properly in order to weigh costs and benefits.

3. Organisation of downstream competition. This encompasses questions regarding whether the passenger market is contestable at all, if and how market design could help reap the benefits of competition and if and how information asymmetry problems in competitive tendering may be solved.

REFERENCES

Bureau of Transport and Regional Economics (BTRE), 2003, Rail Infrastructure Pricing: principles and practice, Commonwealth of Australia, Report 109.

Burns, P, et al.,, 2004, Yardstick Competition - a win-win setting?, Conference on Applied Infrastructure Research, Berlin.

Cantos, P., Pastor, J.M., L. Serrano, 1999, Productivity, efficiency and technical change in the European railways: a non-parametric approach, Transportation 26(4), 61-68.

Cantos, P., J. Maudos, 2001, Regulation and efficiency: the case of European railways, Transportation Research Part A, 35.

Coelli, T., Prasada-rao, D.S., G.E. Battese, 1998, An introduction to efficiency and productivity analysis, Kluwer Academic Publishers.

Coelli, T., S. Perelman, 2000, Technical efficiency of European railways: a distance function approach, Applied Economics, vol. 32, 1967-1976.

Cowie, J., G. Riddington, 1996, Measuring the efficiency of European railways, Applied Economics, vol. 28, 1027-1035.

Eurostat, 2005, Statistics in Focus: Transport, 3/2005.

Friebel, G., Ivaldi, M., C. Vibes, 2003, Railway (de)regulation: a European efficiency comparison, IDEI report no 3 on passenger rail transport, University of Toulouse.

Gathon, H.J., S. Perelman, 1992, Measuring technical efficiency in European railways: a panel data approach , The Journal of Productivity Analysis, vol. 3, 131-151.

Gathon, H.J., P. Pestieau, 1995, Decomposing efficiency into its managerial and its regulatory components: The case of European railways, European Journal of Operational Research 12, 500-507.

Gent, H.A. van, P. Rietveld, 2005, Deregulering railvervoer duldt geen zigzagbeleid, in: ESB-dossier “Tien jaar mededingings- en reguleringsbeleid”, ESB, 30 juni, nr. 4464.

Lan, L.W., E.T.J. Lin, 2004, Measuring the Railway Efficiency, Effectiveness, Productivity and Sales Force with Adjustment of Environmental Effects, Data Noise and Slacks.

Laffont, J.-J., J. Tirole, 1993, A Theory of Incentives in Procurement and Regulation, The MIT Press, Cambridge/London.

Nash, C., et al., 2004, Rail regulatory reform in Europe - principles and practice, paper presented at the STELLA Focus Group 5 synthesis meeting, Athens, June.

Nishimizu, M., J.M. Page, 1982, Total factor productivity growth, technological progress, and technical efficiency change: dimensions of productivity change in Yugoslavia, 1967-1978, Economic Journal 92, 920-936.

Oum, T.H., C. Yu, 1994, Economic efficiency of railways and implications for public policy, Journal of Transport Economics and Policy 38, 120-138.

Oum, T.H., Waters II W.G., C. Yu, 1999, A Survey of Productivity and Efficiency Measurement in Rail Transport, Journal of Transport Economics and Policy 33, 9-42.

Pittman, Russell, 2003, Railroads are different, paper presented at Conference “Structural separation and access pricing: a new synthesis”, Amsterdam, 21 November.

Rivera, C., 2004, Measuring technical efficiency in North and South American Railways using a stochastic frontier model: an international comparison, University of Leeds, Forthcoming.

Velde, Didier van de, 2000, Dutch and Japanese Railway Reforms and Exchanges, Japan Railway & Transport Review 24, July.