REPORT OF THE ADVISORY BOARD 2018
BENCHMARKING AND REGULATION IN THE TRANSPORT SECTOR

ART Advisory Board

December 2018

Autorità di regolazione dei trasporti

The Transport Regulation Authority (Autorità di Regolazione dei Trasporti – ART) was set up under article 37 of Decree-Law No 201 of 6 December 2011 and established within the framework of the public utilities regulatory bodies pursuant to Law No 481 of 14 November 1995.

The Authority's main mission is the economic regulation in the field of transport and access to related infrastructures and services. Further, ART is competent for defining the quality levels of transport services and the minimum content of the rights that may be claimed by users against infrastructure managers. The Authority reports annually to the Parliament highlighting the state-of-the-art of the liberalization measures which have been already adopted and those which remain to be defined.

The Authority is an independent administrative authority and operates in full autonomy, in accordance with the EU legislation and in compliance with the subsidiarity principle and with the powers of regions and local authorities.

It operates as a collegial body composed of the President and two Commissioners designated by Decree of the President of the Republic, upon decision of the Council of Ministers and proposal of the competent Minister, and voted in by at least two-thirds of the members of the competent parliamentary committees. The President and the Commissioners have a seven-year non- renewable mandate.

The Authority was established on 17 September 2013 and is based in Turin at the "Lingotto" building in via Nizza 230.

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This Report aims at discussing the role and potential use of benchmarking in a pro-competitive regulatory strategy in the transport sector.

There is a general preference amongst economists for competition 'in the market' as the best way to provide incentives for static and dynamic efficiency. However, this may not always work well in the transport sector. Infrastructure is often seen as a natural monopoly, and there may be economies associated to the integration of services, as shown for instance by the experience of competition in the market for bus services in the UK.

Competition 'for the market' may provide similar incentives, provided that the contract ensures that the franchisee bears both revenue and cost risk. But there may also be circumstances in which this is not efficient, for instance when there are very high costs to switching supplier or where isolating exogenous revenue risk from risks under the control of operators is challenging, leading to problems of winner's curse. In this case, regulation may be the best way of ensuring efficiency and preventing abuse of monopoly power.

Which of the three scenarios (competition in the market, competition for the market or regulation in the absence of competition) should be pursued in a specific case requires an in-depth analysis of the economic features of the relevant markets.

Benchmarking, which is the process of comparing the performance of different decision-making units to determine what is an efficient level of costs for the process concerned, can play a role in each of these settings. In the absence of competition, since the regulator needs to know what would be the costs of an efficient operator for the service in question, benchmarking is central. However, benchmarking may also be valuable in testing whether competition in or for the market is working effectively, and in designing competitive tendering arrangements.

A policy of yardstick competition based on benchmarking faces various challenges. Firstly is the need to find comparable situations (or to use appropriate statistical methods to allow for differences); which in turn implies good quality and comparable data of sufficient quantity in order to be able to estimate the technology. Secondly is the risk of collusion, although this is at least as big a problem with competition in and for the market. Thirdly is the risk that regulation damages the incentive to invest, which may lead to a need for the regulator to regulate investment as well as price and service quality. And fourthly, is the risk of regulatory capture. The most effective defence against this is for yardstick competition to be practiced by an independent regulator with adequate powers to enforce its decisions.

The Report is composed of eight chapters, plus a short chapter containing the main conclusions.

Chapter 2 discusses the role of benchmarking in EU competition policy. The European Court of Justice acknowledges benchmarking as a tool which can be used both to enforce antitrust rules and in the control of State aid. In antitrust policy, benchmarking may be of value in identifying cases of abusive exploitation of market power, in particular to ascertain whether prices applied by a dominant companies are 'excessive' according to the EU case-law. In the control of State aid, it can be relevant when assessing the existence of State aid, for instance in terms of overcompensation of public service obligations for services of general economic interest. On the other hand, the use of benchmarks by competition authorities and by national courts in the enforcement of competition law requires appropriate methodologies in order to meet the required standard of proof.

Chapter 3 contains an overview of benchmarking methodologies, with a focus on benchmarking as a tool to enhance efficiency. There exists a range of methods used in the academic literature to assess efficiency and productivity performance both between firms and over time. These include productivity indices, the most simple of which is labour productivity; with more advanced indices covering all inputs (total factor productivity).

Non-parametric methods such as data envelopment analysis (DEA) permit the analysis of efficiency as well as productivity analysis. Parametric methods likewise permit a rich analysis. Perhaps a key advantage of the parametric method is that it yields information on the relationship between costs and the regressors (which can be checked and challenged), which data envelopment analysis does not. Whilst parametric methods do require the choice of a functional form – which can be challenging – DEA is not necessarily an antidote to that problem, because it can be seen as a “black box” that “hides” the underlying shape of the frontier. In regulatory applications, our experience is that parametric methods are more used than DEA. In any case, where there are many cost drivers, DEA often requires a parametric second stage, thus reducing the difference between the methods to some extent.

Particularly important aspects of any benchmarking method are that it controls for heterogeneity between firms and over time, and that it deals with the possibility of economies of scale and density that are so prevalent in transport applications particularly (and network industries more widely). It needs to be credible, implying, inter alia, good quality and comparable data; and potentially it also needs to be transparent, which may mean that simplicity could be preferred by regulators in some cases.

Chapters 4-7 discuss the role of benchmarking in the different modes of transport: highways; rail infrastructure and operations; buses and local public transport; ports and airports. For each mode, the authors consider the organization and scope for competition in the market and for the market; the potential role of yardstick competition/benchmarking; particular issues in terms of methodology, data, specification of inputs, outputs, quality and control variables; moreover, the authors provide a brief overview of existing studies and point out further research needs. These chapters emphasise the problems of getting comparable data, especially when the study has to be international because of a lack of comparable firms in a single country (although if there are separate decision making units within one firm, these may be used as a sample if data is available). The authors also illustrate the ways of taking account of heterogeneity and discuss the uses of benchmarking to advise on issues other than price cap regulation, such as the structure of franchises and the problem of whether regulation is needed at all.

Chapter 8 focuses on the relevance of benchmarking in the transport sector to the Italian institutional framework. In Italy, the transport regulator (ART) has extensive responsibilities regarding pricing and also design of franchising schemes for a number of modes of transport, and in performing these duties benchmarking is clearly of great relevance.

For highways, benchmarking methodologies are expected to play an increasing role both in the design of tenders and in the design of contracts, also in case of in house awards, and the role of ART may be useful to promote more uniform efficiency-enhancing regulatory methods. Monitoring the maintenance and investment costs of non-toll highways and roads would also be important to broaden the information basis for the entire sector.

For rail infrastructure, international benchmarking may be useful to ensure proper efficiency-enhancing incentives, provided that it duly takes the quality dimension into account. Such studies could use data for national rail networks or utilise emerging datasets on regions / areas within several different countries.

For passenger rail transport services subject to public service obligations, as for local public transport by bus, benchmarking should be included in a broader efficiency-enhancing strategy. In particular, it may be of value in defining the optimal dimension which should be considered in the award of contracts and in establishing the content of contracts, both for tenders and for direct awards. The current use of 'standard costs' on negotiating contracts is an important step in the right direction, but we believe further work is needed in order to ensure that these are derived using the most appropriate techniques.

Both for Italian seaports and airports, market studies are crucial to understand the market evolution and how it affects the need for regulatory intervention. For seaports, the main challenge for regulation is to promote efficiency-enhancing conditions of access to port areas and infrastructures. Estimates of minimum efficient scale for the different services may support pro-competitive franchise policies in the port areas. Since the traditional fragmentation of regulatory approaches in the port sector has represented a major weakness, ART may play a useful role in promoting a common approach to accounting, admissible costs, etc., so as to improve transparency, comparability and spread of best practices. For airports, benchmarking based on estimates of efficient frontiers can be useful to improve efficiency-enhancing models for airport charges.

Chapter 9 contains the overall conclusions. There is a strong need for the use of benchmarking to support yardstick competition by providing estimates of efficient costs to use in setting price caps and efficiency targets. Benchmarking is also helpful where there is competition for the market, in that it may be used to determine the characteristics of efficient contracts, such as size, length, gross or net cost and conditions attached. Benchmarking may also be valuable where there is competition in the market if there is any doubt as to its effectiveness, as it may suggest where there are areas of inefficiency.

In the past benchmarking often relied on crude comparisons of unit cost, with an attempt to ensure that the decision-making units being compared were broadly comparable. There are now benchmarking techniques available which can deal with issues such as multiple outputs and inputs, economies of scale and density, and heterogeneity of operating conditions which are ignored by simple measures. The biggest problem is the quality and quantity of data. Benchmarking studies using these techniques exist for all the main modes of transport and the technique is already being used in some cases by regulators.

In Italy, there is scope for the use of benchmarking in all modes covered by ART. We believe that there is much more scope for the use of competitive tendering in awarding contracts (for instance for bus and rail services) than is currently the case in Italy, and benchmarking may help determine such factors as the appropriate size and length of concessions. But where competitive tendering is not introduced, or until it is, then benchmarking, as a way of setting targets for efficiency (yardstick competition) and price caps, is a crucial tool.

1. THE USE OF BENCHMARKING IN A PRO-COMPETITIVE REGULATORY STRATEGY

Ginevra Bruzzone, Chris Nash and Marco Ponti

1.1. Three scenarios: competition in the market, competition for the market, lack of competition

Competition in the market, when it is feasible, is a formidable mechanism leading undertakings, in a context of imperfect information, to look for efficiency and better products (through innovation and imitation) and to contain prices to the ultimate advantage of consumers. The process works on a continuous basis. During the process undertakings may make mistakes: they may invest too much, or not enough, or in fruitless directions; they may propose unsuccessful products; or they may set prices at the wrong level, but the mechanisms for the correction of such mistakes are embedded in the system (for instance loss of market shares and reduced profitability).

When competition in the market is not feasible, competition for the market is a partial substitute: if the tender is properly designed, potential competitors will reveal information on the best price/quality combination they are able to offer at the time of the tender (in the language of public procurement law, the most advantageous economic offer). For services of general economic interest and infrastructures, a competitive offer may either entail lower prices for users or reduced taxpayer support, depending on how compensation is structured.

When competition for the market entails obtaining a temporary monopoly, it has a potential weakness compared with a situation of competition in the market: the winner is freed from the efficiency incentives that competition in the market provides on a continuous basis.

Thus, with competition for the market it is up to policy makers to ensure that the conditions of the "contract" stipulated with the company or the regulatory process will provide the proper incentives in terms of efficiency and quality, for instance by allowing them to retain cost savings or revenue increases they achieve for the duration of the contract, and likewise ensuring that they bear any losses resulting from their own inefficiency.[1]

In the absence of either competition in the market or competition for the market, public policy must use other instruments to collect the information needed to ensure, by means of an appropriate contract/regulation, that undertakings will have the right incentives in terms of efficiency, quality and investment. There are different benchmarking methodologies which can be used to this aim.

In a pro-competitive regulatory strategy, there is a presumptive order of preferences with respect to the three scenarios. Competition in the market when feasible is usually better than competition for the market; competition for the market, when feasible, is often better than the alternative of no competition accompanied by regulation/contract with the company. The order of preferences depends on the different effectiveness in extracting information in a dynamic context and in aligning the incentives of companies so as to ensure a good combination of efficiency/quality and proper investment choices, in the ultimate

[1] That may mean ultimately allowing the company to become bankrupt. In such circumstances, the least disruptive immediate response may be to renegotiate the contract, but this risks arousing expectations that future contracts will be renegotiable, and thus encouraging unrealistic bidding whilst damaging incentives for efficiency.

interest of consumers and, in case of public services/infrastructures, also of taxpayers. Thus, in competition advocacy the reduction of unjustified barriers to entry is usually considered the first best strategy.

However, there is also evidence of cases in which competition in the market (e.g. in local public transport in the UK) is not efficient, due for instance to significant economies associated to an integrated supplier. Likewise, competition for the market may not always be better than yardstick competition; in particular, very high transaction costs associated to a change of supplier may suggest to only retender if the incumbent is found not to be performing efficiently.

In light of the above, which of the three scenarios should be pursued in a specific case requires an in-depth analysis of the economic features of the relevant market. Market studies are of the utmost importance to understand what kind of competition or simulated competition is efficient on a market. An analysis based on modes of transport (e.g. rail infrastructure and operations, ports, airports) gives important insights but, from a practical viewpoint, may not be sufficient. As the example of ports and airports clearly demonstrates, ascertaining whether competition in the market can play a role of market discipline requires a case-by-case assessment of the existence, from the point of view of users, of an appreciable degree of substitutability between existing infrastructures.

Benchmarking, broadly understood as a process through which practices are analysed to provide a standard measurement (‘benchmark’) of effective performance within an organisation, can play a role in each of the three scenarios.

In this introductory chapter, we outline the potential use of benchmarking with competition in the market (paragraph 1.2), competition for the market (paragraph 1.3) and in the absence of competition either in or for the market (paragraph 1.4). Paragraph 1.5 focuses on some critical features which must be addressed with respect to the use of benchmarking, and in particular of yardstick competition as a regulatory tool: the main risks that must be managed are the risk that the benchmarking analysis is distorted by collusion between undertakings, the risk of undermining the incentives to invest and the risk of capture, which can make yardstick competition ineffective in reaching its efficiency enhancing goals. Paragraph 1.6 outlines the structure of the Report.

1.2. Benchmarking and competition in the market

With competition in the market, it is in the interest of the companies to take into account what more efficient competitors do, e.g. price/quality combinations, and adapt their business strategy to meet the competitive challenge.

Comparison of the different offers may even be used by companies as a marketing tool (comparative advertising is legitimate pursuant to EU law if it is based on objective criteria). On the other hand, it may be in the interest of purchasers to develop instruments aimed at facilitating the comparisons of the different offers thus reducing transaction costs (e.g. price comparison tools).

Thus, when several companies compete on a market, there is no need for policy makers to promote benchmarking, because the task is carried out by the market (market-led benchmarking).

Within this general framework, there are two exceptions which must be taken into account:

- if companies collude, the virtuous competitive mechanism is blocked;
- if a company, although not a monopolist, has such a market power that it is not significantly constrained by actual or potential competitors (what in competition law is called a dominant position), there may remain a role for public policy.

The prohibition of anticompetitive agreements, enforced by the European Commission and national competition authorities, aims to avoid collusion distorting the market process. In practice, policy makers should always pay particular attention to whether there are signals of collusion in a given market.

As to dominant positions, which are frequently met in the transport sector, there are different tools which may be used to mitigate the problem. Competition advocacy may be used to promote more effective actual or potential competition, for instance reducing barriers to entry and exit from the market. In case of high and lasting market power, since the step to formal regulation is a strong one, monitoring and benchmarking may be used as a soft tool to align the incentives of the incumbent to best practices. However, in the presence of structural obstacles to the competitive process (e.g. a natural monopoly on a lasting basis) ex ante regulation is usually more appropriate than antitrust intervention.

In particular cases in which for any reason regulation is either non-existent or insufficient, competition authorities in the EU are empowered to prohibit the abusive exploitation of market power, such as for instance the application of excessive prices (Art. 102 (a) TFEU). The prohibition of exploitative abuses is seldom used by competition authorities because their main role is not regulating the market, but to protect the market process; however, this power, when needed, is available. Chapter 2 of the Report discusses in more detail how competition enforcers proceed in these cases, with a focus on the use of benchmarking.

1.3. Benchmarking and competition for the market

As to competition for the market, first of all, the analysis of costs, including the assessment of economies of scale, density and scope, is useful to properly design the tender. Concessions need to be large enough to exploit such economies, but there is a need to avoid “oversized” concessions which does not depend only on reasons of static efficiency. The coexistence of several concessionaires helps maintain sufficient experienced bidders for future competitions to work, whilst also making it possible to compare their performance and thus develop forms of competition by comparison.

In addition, public decision makers may use benchmarking in order to set the proper incentives on the company in terms of both price and quality (efficiency and effectiveness), by means of a proper definition of the contract with the supplier of the service.

Collusion may be a problem also with competition for the market, since it may eliminate the incentives to present competitive bids and therefore the information and efficiency enhancing role of tenders. Notably, fighting collusion in public procurement is one of the priorities of competition authorities worldwide.

The advantages in terms of efficiency-enhancing incentives of competition for the market compared to direct award of contracts are acknowledged by EU case-law concerning the control of State aid (Box 1.1.).

BOX 1.1. STATE AID CONTROL OF SGEI: PRESUMPTION THAT COMPETITION FOR THE MARKET LEADS TO COMPETITIVE CONDITIONS

In State aid control, according to the case law of the Court of Justice in the Altmark case (C-280/00) and the interpretation the European Commission gives thereof, in order to ensure that there is no overcompensation of public service obligations and, thus, no State aid, it is sufficient that:

- there is an entrustment act clearly defining the public service obligation;
- the parameters for calculating the compensation are established in advance in an objective and transparent manner;
- the compensation does not exceed the relevant costs and a reasonable profit taking into account the income associated to the provision of the SGEI;
- the provider is chosen through a public procurement procedure aligned to the EU directives on public procurement and concessions (with only two exceptions: when only one bid is submitted; in case of negotiated procedure without transparency).

Thus, pursuant to State aid law, it is presumed that the existence of competition for the market will be able to reveal the conditions which would emerge in a competitive situation, i.e. avoid overcompensation.

On the other hand, it is still necessary to specify in the entrustment act the parameters for calculating, controlling and reviewing the compensation as well as the arrangements for avoiding and recovering any overcompensation: thus, there remains a role for benchmarking in the design of the contract.

1.4. Benchmarking in the absence of competition either in or for the market

In the absence of competition, either in or for the market, for public decision makers the task of collecting information in order to avoid inefficiencies and undue exploitation of market power in terms of poor quality or excessive prices/subsidies is crucial, because no information or incentive to this aim results from the operation of the market.

Public authorities have to collect information so as to write a proper contract/properly regulate the undertaking. This is particularly important for in house contracts and direct award of contracts.

Benchmarking is also relevant for State aid law in the absence of a tender: according to the Court of Justice, when all the other requirements of the Altmark case-law are met except for a public procurement procedure aligned to EU directives, the existence of overcompensation may still be excluded by means of a proper benchmark, i.e. by determining the level of compensation on the basis of an analysis of the costs of "a typical well run and adequately equipped undertaking in the sector concerned" (see Chapter 2).

In the third scenario, i.e. absence of competition, benchmarking methodologies are crucial for policy makers to identify the conditions which might be attained by an efficient company in terms of costs, prices, contribution of taxpayers and take them as a target.

Moreover, in the absence of competition some benchmarking may be needed also to ensure proper incentives as to the provision of quality.[2]

[2] See for instance, CBP Bureau for Economic Policy Analysis (2000), *Yardstick Competition, Theory, Design and Practice*, Working Paper no. 13, December. An example of benchmarking focussing on the output/quality dimension is provided by the system of performance assessment created by the Laboratorio MeS of the Scuola Superiore Sant'Anna in Pisa. In this case, the system of benchmarking works with the participation of a group of regional public administrations on a voluntary basis. See S. Nuti (2012), "Assessment and Improvement of the Italian Healthcare System: First Evidence from a Pilot National Performance Evaluation System*", Journal of Healthcare Management*

Yardstick competition, as designed by Shleifer,[3] is a specific form of regulation which requires collecting information on cost conditions of comparable companies and linking the remuneration of the regulated undertaking to the costs of other companies, so as to incentivise it to become at least as efficient as comparable companies. The more efficient is the company compared to competitors, the higher is its profit. It is just one of the regulatory instruments which are based on benchmarking. Whether some form of yardstick competition or the use of another of the available tools is more appropriate in a given situation should be assessed on a case by case basis, taking all pros and cons into account. Thus, a yardstick competition approach should be compared, for instance, with the mere publishing of the results of a benchmarking process.

1.5. Benchmarking as a regulatory tool: the main challenges

Looking at the potential use of benchmarking as a regulatory tool, there are some main challenges to be considered. In general, it is essential to acknowledge that inefficiency can take different forms depending on the goals of the undertakings involved. In particular, whereas private monopolies usually pursue high profits, the main feature of publicly owned monopolies may be high costs of production.

A first challenge consists in finding comparable situations. The problem arises in all cases of benchmarking; it is particularly serious when benchmarking is used as a basis for regulatory decisions which may be challenged before the courts by companies arguing that the regulator did not make reference to an appropriate benchmark. Some technical tools may be used to deal with the problem of heterogeneity (see Chapter 3 of this Report), but the issue of whether a meaningful benchmark can be found may be a serious problem in some markets.

The second challenge is the risk of collusion, not only within a market but also across markets. If companies collude not to strive for efficiency, any benchmark based on the observation of their conduct will not be effective. Although from an institutional viewpoint the competence to enforce the prohibition of anticompetitive agreements is within competition authorities, an issue to be addressed is how competition authorities can more effectively obtain the relevant information to challenge these practices. This is an area for potential cooperation between sectoral authorities and competition authorities. Another issue is how policy makers may check for collusion before using market information in order to develop benchmarks. In the economic literature, a number of studies focus on how to structure regulation so as to reduce the risk of collusion.[4]

A third challenge which has been pointed out by the literature on yardstick competition is ensuring that the rules do not have an adverse impact on incentives to invest, because companies are incentivized to focus on cost minimization.[5]

In a political economy perspective, a further challenge is represented by the risk of regulatory capture, which may be particularly serious in the absence of competition either in or for the market. The reason is that the links between public administrations and undertakings which may give rise to capture are

57, 3, May-June 2012, 181-199.

[3] A. Shleifer (1985), "A Theory of Yardstick Competition*", Rand Journal of Economics* 16(3), 319-327.

[4] See for instance T. P. Tangeras (2002), "Collusion-proof yardstick competition", *Journal of Public Economics*, vol. 83, 2, February 2002, 231-254.

[5] See, for instance, D. M. Dalen (1998), "Yardstick Competition and Investment Incentives", *Journal of Economics and Management Strategy*, 1998, 7, 1, 105-126; G. Guthrie (2006), "Regulating Infrastructure: the Impact on Risk and Investment", *Journal of Economic Literature*, vol. XLIV, 925-972.

weakened with competition in the market and, at least to some extent, also with competition for the market.

As to the latter, with the periodic tendering of concessions the winners tend to rotate, and the tendering administration will be incentivised to try to extract the best services at a minimum cost, with limited time or opportunity for lobbying to play a role.[6] In the tendering scenario, the main risk of capture arises after the contract is awarded, when the winner may try to renegotiate it.

In the case of yardstick competition the situation is different since the proper incentives depend only on the action of public administrations: the administration either will compare the performance of the existing regulated company with other ones, or will create an "internal" yardstick competition by unbundling the incumbent company, along lines of "minimum efficient dimensions". Two problems arise: information (for instance on costs), that has to be guaranteed by an external agent, and effective sanctions against the non-performing subjects (or, vice versa, prizes). Information, that usually the tendering process allows to get, will not be observable, and sanctions/rewards are to be effective in changing the behaviour of the regulated companies. If capture persists, the administration may have little incentive to sanction, if not formally. Also the incentive to show a strong attitude in defending the public interest against an "external, profit-minded" subject suffers, unless in the presence of a strictly binding budget constraint, and this not overall, but aimed at that specific company.

In light of the above, with yardstick competition the best defence against regulatory capture is entrusting the regulatory tasks to an independent authority with a strong reputation and adequate powers.

1.6. Structure of the Report

This Report aims at discussing the role and potential use of benchmarking in a pro-competitive regulatory strategy in the transport sector.

- Chapter 2 considers the use of benchmarking in EU competition policy, with reference to both antitrust rules and the control of State aid.
- Chapter 3 contains an overview of benchmarking methodologies, with a focus on benchmarking as a tool to enhance efficiency.
- Chapters 4-7 discuss the different modes of transport: highways; rail infrastructure and operations; buses and local public transport; ports and airports. For each mode, the authors consider the organization and scope for competition in the market and for the market; the potential role of yardstick competition/benchmarking; particular issues in terms of methodology, data, specification of inputs, outputs, quality and control variables; moreover the authors provide a brief overview of existing studies and point out further research needs.
- Chapter 8 focuses on the relevance of benchmarking in the transport sector to the Italian institutional framework.
- Chapter 9 contains the overall conclusions.

[6] In principle, also periodic competitive tendering can be twisted to "capture from the incumbent", but this will be ex ante, at a specific point of time, and politically more visible than a sanctions/prizes system that operates over long periods.

Chapter references

CBP Bureau for Economic Policy Analysis, (2000), *Yardstick Competition, Theory, design and practice*, Working Paper no. 13, December.

Dalen, D.M., (1998), "Yardstick Competition and Investment Incentives", *Journal of Economics and Management Strategy*, 7 (1), 105-126;

Guthrie, G., (2006), "Regulating Infrastructure: the Impact on Risk and Investment", *Journal of Economic Literature*, vol. XLIV, 925-972.

Nuti, S., (2012), "Assessment and Improvement of the Italian Healthcare System: First Evidence from a Pilot National Performance Evaluation System", *Journal of Healthcare Management* 57 (3), May-June , 181-199.

Shleifer, A., (1985), "A Theory of Yardstick Competition", *Rand Journal of Economics,* 16 (3), 319-327.

Tangeras, T. P., (2002), "Collusion-proof yardstick competition", *Journal of Public Economics*, 83 (2), February, 231-254.

2. BENCHMARKING IN EU COMPETITION LAW

Ginevra Bruzzone

2.1. Introduction

The European Court of Justice acknowledges benchmarking as a tool which can be used both to enforce antitrust rules and in the control of State aid.

Although competitive benchmarks may be used by competition authorities also in merger control (when the parties, in order to obtain the approval of a notified merger, propose behavioural remedies such as the commitment to grant access to key infrastructures)[8], the most interesting issues in the perspective of regulatory strategies emerge with respect to the enforcement of the prohibition of the abuse of dominance.

In the control of State aid, benchmarking can be relevant when assessing the existence of State aid, for instance in terms of overcompensation of public service obligations for services of general economic interest.

The use of benchmarks by competition authorities and by national courts in the enforcement of competition law requires appropriate methodologies in order to meet the required standard of proof.

2.2. Benchmarking in antitrust policy: exploitative abuses

In the EU, differently from other jurisdictions including the US, competition law includes some provisions which, in principle, may be used as a substitute for regulating the conduct of dominant undertakings. In particular, the list of abuses of dominant position in Article 102 TFEU includes a number of practices which, differently from exclusionary abuses, are not harmful to the structure of competition on the relevant market but may be seen as the 'abusive' exploitation of the market power of the dominant company. In particular, Article 102(a) TFEU prohibits conduct by one or more dominant undertakings "directly or indirectly imposing unfair purchase or selling prices or other unfair trading conditions". These conducts are usually referred to as exploitative abuses of dominance.

A broad consensus exists on the fact that direct control of prices or trading conditions should not be the main task of competition authorities. The best possible intervention against excessive prices or unfair trading conditions is indirect intervention, i.e. enforcement and/or advocacy actions to eliminate or rectify the conditions that cause such prices/practices. In most sectors, high profits will normally attract new entrants and the market will in due course correct itself. Intervention by competition authorities may distort the process. Moreover, in the presence of structural obstacles to the competitive process (e.g. a natural monopoly on a lasting basis) ex ante regulation (either hard regulation or soft regulation entailing continuous monitoring) is more appropriate than antitrust intervention.[9] Importantly, differently from regulation, the application of a prohibition rule, such as the prohibition of the abuse of dominance, entailing sanctions in case of infringement implies a presumption of innocence: it is for the competition authority to prove, according to the standard of proof required by the courts, that a price is "excessive" or a trading condition is "unfair".

[8] See for instance Commission notice on remedies acceptable under Council Regulation (EC) No. 139/2004 and under Commission Regulation (EC) No. 802/2004.
[9] See among others OECD Policy Roundtable (2011), *Excessive prices*, DAF/COMP(2011)18.

Thus, traditionally the prohibition of exploitative abuses, including the prohibition of excessive pricing, has been applied sparingly by competition authorities and courts, although in recent years there has been a kind of renewed interest in cases concerning alleged excessive prices.[10] In practice, competition authorities use Article 102 to prohibit exploitative abuses in specific cases where the protection of users cannot otherwise be adequately ensured. Often, this tool is applied in markets with lasting market power where for some reason regulation is either non-existent or ineffective.[11]

Taking this general framework into account, within a broader study on the use of benchmarking it may be interesting to provide a brief overview of the approach of the European Court of Justice to the assessment of whether the prices applied by a dominant company are "excessive".

As early as 1978, in *United Brands* the Court of Justice indicated that "charging a price which is excessive because it has no reasonable relation to the economic value of the product supplied …is an abuse", but at the same time quashed the decision of the European Commission because it failed to make out a clear case finding an excessive price. The Court suggested that to determine whether a price is abusive, it should be considered "whether the difference between the cost actually incurred and the price is excessive" and, if so, whether the price is either unfair in itself or when compared with other competing products.[12]

Some more precise indications were provided in a recent judgment of the Court of Justice in the *Akka Laa* case.[13]

First, in *Akka Laa* the Court of Justice acknowledges that there are different methods which can be used to ascertain whether a price is excessive, including methods based on a comparison of the prices applied in different Member States. If the price is appreciably higher than those charged in other Member States and the comparison has been made on a consistent basis, that difference may be regarded as indicative of an abuse of a dominant position.[14]

The Court adds that a comparison cannot be considered to be insufficiently representative merely because it takes a limited number of Member States into account, on condition that such States are selected in accordance with objective, appropriate and verifiable criteria. Those criteria may include, inter alia, consumption habits and other economic and socio-cultural factors, such as GDP per capita and cultural and historical heritage. The relevance of the criteria applied in a specific case must be assessed by the competition authority/national court taking into account all the circumstances of the case. For instance, neglecting the different willingness to pay resulting from divergences in living standards and purchasing power may lead to biased results; it may thus be appropriate, when making the comparison, to take the

[10] See for instance the following cases: decision of the Director General of Fair Trading, no. CA98/2/2001, 30 March 2001, *Napp;* Italian Competition Authority, decision of 23 October 2008, case A375, *Airport fees*; *E-on* (Germany, 2008); Competition and Markets Authority, decision of 7 December 2016, case CE/9742-13, *Flynn*; Italian Competition Authority, decision of 29 September 2016, case A480, *Aspen*; Bundeskartellamt, *District Heating Suppliers*, decision with commitments, February 2017; Competition and Markets Authority, *Actavis,* statements of objections of 16 December 2016 and 9 August 2017.

[11] On the application of competition rules in regulated sectors, see *Deutsche Post* and, for an overview, G. Bruzzone (2015), *Rapporti tra l'Autorità di regolazione dei trasporti e altre autorità dipendenti di vigilanza e garanzia*, in "Istituzioni e regolamentazione dei trasporti: temi di riflessione", Report of the Advisory Board of the Autorità di regolazione dei trasporti (ART).

[12] Court of Justice, judgment of 14 February 1978, case 27/76, *United Brands*, § 252.

[13] Court of Justice, judgment of 14 September 2017, case C-177/16, *Akka Laa*, §§ 31-61.

[14] *Akka Laa*, § 38; see also Court of Justice, judgment of 13 July 1989, case 395/87, *Tournier*, § 38; Court of Justice, judgment of 13 July 1989, joined cases 110/88, 241/88 and 242/88, *Lucazeau*, § 25.

PPP index into account. A comparison with a larger set of States may be used to check the robustness of the results.

Moreover, the Court indicates that the comparison may be limited to one or several specific user segments, if there are indications that the possibly excessive nature of the price affects those segments.

In order to assess whether the difference in prices is appreciable, and thus presumptively abusive, the Court observes that it would be improper to refer to fixed minimum thresholds, since the circumstances specific to each case are decisive in this regard. Thus, a difference may be qualified as appreciable if it is both significant and persistent on the facts with respect to the market in question, this being a matter for the competition authorities or courts to verify.

Finally, when on the basis of these criteria the difference with respect to the benchmark is regarded as presumptively abusive, it is for the dominant company to show, when feasible, that its prices are fair because there are objective justifications for such difference.

A more detailed discussion of the use of benchmarking in the application of competition law is found in the opinion of Advocate General Wahl in *Akka Laa* (see Box 2.1.)

BOX 2.1. BENCHMARKING IN THE APPLICATION OF COMPETITION RULES: THE VIEW OF ADVOCATE GENERAL WAHL IN AKKA LAA

Combining different methods

In the absence of an ubiquitous test and given the limitations inherent in all existing methods, it is crucial that in order to avoid (or, more correctly, to minimise) the risk of errors, competition authorities should strive to examine a case by combining several methods among those which are accepted by standard economic thinking and which appear suitable and available in the specific situation. The choice to combine several methods is, in fact, the approach that a number of antitrust authorities have followed worldwide: for example, the UK Office of Fair Trading (OFT) has done so in the *Napp* case.[14] It is also consistent with suggestions made in international discussion fora of those authorities as well as in contemporary economic literature. It is true that such an approach has been criticised on the ground that the combined application of several imprecise methodologies, even where producing mutually consistent results, may not lead to a more reliable conclusion. Admittedly, the weaknesses of one method are not necessarily remedied by applying another equally weak method. Yet, if the methods are applied independently of each other, a given limitation inherent to one of them would not affect the results obtained through the use of other methods. Accordingly, provided that the methodologies used are, in themselves, not flawed, and that they are all applied with rigour and objectivity, the convergence of results may be taken as an indicator of the possible benchmark price in a given case.

[14] In its decision, concerning allegedly excessive prices in the pharmaceutical sector, the OFT has sought to establish a proxy for the competitive price of the product by looking at the prices of competitors and the prices Napp charged elsewhere, and seeing whether those prices would have enabled Napp to earn a reasonable profit. The OFT did not rely on a single comparison, but made six comparisons, all of which were viewed as supporting its conclusions. See decision of the Director General of Fair Trading, no. CA98/2/2001, 30 March 2001, endorsed by the UK Competition Appeal Tribunal, in its judgment of 15 January 2002, *Napp Pharmaceutical Holdings Ltd.* (2002) CAT 1, § 56-69.

Additional indicators

That said, there may be cases in which only one of those methods of determining the benchmark price may be available or suitable. In those cases, it is of the utmost importance that the authority considers other indicators which may corroborate or, conversely, cast doubt on the result of that method

First, a price cannot easily be set significantly above the competitive level where the market is not protected by high barriers to entry or expansion. Otherwise, as mentioned above, the market should, in principle, be able to self-correct in the short to medium term: high prices should normally attract new entrants or encourage existing competitors to expand. That is why unfair prices under Article 102 TFEU can only exist in markets where the scope for free and open competition is reduced, typically in regulated markets. Obviously, the higher and longer-lasting the barriers created by the legislature, the more a dominant undertaking should be able to exercise its market power.

Second, a price significantly in excess of a competitive price is more unlikely to occur in markets where there is a sectoral regulator whose task is, inter alia, to fix or control prices charged by the undertakings active in that sector. Sectoral authorities are clearly better-equipped than competition authorities to oversee prices and, where necessary, act to remedy possible abuses. It would seem, therefore, that antitrust infringements in those situations should be mainly confined to cases of error or, more generally, to regulatory failures: cases where the sectoral authority should have intervened and erroneously failed to do so.

Third, an undertaking with market power is evidently less able to leverage its position when negotiating with powerful buyers. For instance, as to licences for the use of copyrighted musical works, the negotiating position of small shops is likely to be different from that of international platforms (such as Spotify) or groups of large and sophisticated undertakings (such as Hollywood majors). The size and financial strength of an undertaking (or group of undertakings) might indeed have a significant weight in the negotiations. However, the extent to which the licenced products constitute an important (or even indispensable) input for the customers' business may also be of great importance in that context.

Clearly, other factors may also be relevant, depending on the specific circumstances of each case.

Although, as anticipated, there is only a small number of cases in which Article 102 has been applied to alleged exploitative abuses, some of these cases concern the transport sector.

In the early 1990s, there have been some cases in the airport sector in which the Italian competition authority ascertained abuses consisting in requiring from users of the airport infrastructure the payment of services which they had not requested (AGCM, *Ibar- Aeroporti di Roma*, 1993; *Ibar-SEA*, 1994).

In *Scandlines-Port of Helsinborg* (2004), notwithstanding a price difference of 360 per cent, the European Commission rejected a complaint arguing that the prices applied by the Port of Helsinborg were abusive, since the comparison with other ports was not made on a consistent basis and, moreover, did not properly take into account differences in willingness to pay.[15]

In Italy, there have been some cases concerning airport fees, where the competition authority stressed that actual prices were not in line with regulatory obligations (2008)[16]. A case concerning air trasport (*Alitalia-Veraldi*, 2001), entailing a comparison between the prices applied by Alitalia on different routes was dismissed because the alleged abuse was not proven to the required standard.

[15] Commission decision of 23 July 2004, case COMP/A.36568/D3, *Scandlines Sverige AB v. Port of Helsinborg.* For a discussion, see Damien Geradin, Anne Layne Farrar, Nicolas Petit (2012), *EU Competition Law and Economics*, Oxford University Press, 280 ss.
[16] ADR-SEA/Airport fees, 2008.

In some of these cases the issue of whether one should refer to the cost of the dominant company (which may be inefficient) or to efficient costs was discussed. It was considered inappropriate to include in the comparison non profitable situations.[17] If estimates are based on restrictive assumptions, there should be more flexibility in assessing whether a difference in prices is proof of the existence of an abuse.[18]

From a public policy perspective, in the light of the above there are two main issues which should still be debated and clarified.

The first one is what are the circumstances in which the antitrust prohibition against excessive prices *should not* be used. For instance, the checklist provided by the Rome Antitrust Forum in December 2017 includes the following:

1. there are other indirect ways to discipline the dominant firm;
2. the market on which the dominant firm operates is not characterized bu high barriers to entry;
3. the dominant firm is not a monopoly (or quasi-monopoly);
4. the market is supervised by a sectoral regulator and there is no clear regulatory failure;
5. the market is two-sided;
6. the dominant firm is a multiproduct firm with substantial common costs;[19]
7. the customers of the dominant firm do not incur high sunk costs.[20]

The second issue, which arises when enforcement of the prohibition against exploitative abuses is appropriate, is how to apply a proper benchmarking methodology to point out whether observed market conditions may be presumptively considered an abusive exploitation of a dominant position. As stressed by Advocate General Wahl in *Akka Laa*, "a lack of reliable data or the complexity of the operations involved in the calculation of the benchmark price (or in corroborating it) cannot justify an incomplete, superficial or dubious analysis by a competition authority. In other words, difficulties encountered by an authority when carrying out an assessment cannot be to the detriment of the undertaking being investigated. Regardless of the specific situation in a given case, the method(s) applied and the other indicator(s) examined must give the authority a sufficiently complete and reliable set of elements which point in one and the same direction: the existence of a difference between the (hypothetical) benchmark price and the (actual) price charged by the dominant undertaking in question". In this respect, the analysis of benchmarking methodologies used in regulation, contained in Chapter 3 of this Report, may provide useful insights also to competition law enforcers.

In any case, when applying a prohibition rule the possibility for the dominant company to provide a justification for its conduct should be preserved.

[17] Commission, *Port of Helsinborg*; AGCM, *Airport fees*.
[18] Commission, *Port of Helsinborg*.
[19] The allocation of common costs is easier in regulation than in a system based on a prohibition rule, entailing the presumption of innocence.
[20] European University Institute – SNA, Rome Antitrust Forum, 4th meeting, 15 December 2017. For similar checklists, see Geradin *et al.*, cit.

2.3. Benchmarking in State aid control

Looking at the use of benchmarking in State aid control, the Court of Justice in the *Altmark* case (C-280/00) argued that in the absence of a proper tender consistent with EU directives on public procurement and concessions, an alternative way to exclude the existence of State aid, i.e. overcompensation of public service obligations, is to "determine the level of compensation on the basis of an analysis of the costs of a typical well run and adequately equipped undertaking in the sector concerned. Notably, the *Altmark* case considered by the Court of Justice concerned local public transport.

In the guidelines concerning the application of the *Altmark* case-law, the European Commission stresses that, when applying this criterion, taking as benchmark a single undertaking or undertakings with enjoy significant advantages is inappropriate. Moreover, if the undertaking does not meet the qualitative requirements established by the rules of the relevant Member State, it is not considered a well-run undertaking.

Although is seems difficult to find proper benchmarks, in practice EU courts have shown some flexibility in the application of this requirement (C-341/06, *Chronopost*; T-289/03, *Bupa Insurance*).

When the requirement of competition for the market is not met and a proper benchmark cannot be found, the Commission requires further measures so as to limit the risk of overcompensation. In particular, according to the general framework for services of general economic interest (2012/C 8/03) the Commission requires accounting separation and regulatory incentives to the efficient provision of the SGEI. For land transport, however, the looser rules of Regulation no.1370/2007 apply: it is sufficient that compensation should not exceed the costs of the undertaking concerned, including a reasonable profit.

Chapter references

Bruzzone G., *Rapporti tra l'Autorità di regolazione dei trasporti e altre autorità dipendenti di vigilanza e garanzia*, in "Istituzioni e regolamentazione dei trasporti: temi di riflessione", Report of the Advisory Board of the Autorità di regolazione dei trasporti (ART).

Geradin D., Layne Farrar A, Petit N. (2012), *EU Competition Law and Economics*, Oxford University Press.

3. BENCHMARKING METHODOLOGY

Andrew Smith and Giovanni Fraquelli

3.1. Introduction

This section reviews the available parametric and non-parametric methods that can be used for benchmarking purposes and indeed have been used in empirical and regulatory studies, as noted in the other chapters.

It starts by motivating the need for more advanced methods (section 2.1), before setting out the available non-parametric methods in Sections 2.2 and 2.3. Parametric methods are covered in sections 2.4. Section 2.5 concludes.

3.2. Why, we need a benchmarking methodology? [21]

Studies of transport efficiency usually have one of two motivations. Firstly they may aim to identify which transport operators and / or infrastructure managers are efficient and which are not, in order to draw lessons as to the level of improvement that may be required. An example of this is the benchmarking studies conducted on behalf of the British rail regulator in deciding on the financial requirements of Network Rail, the infrastructure manager discussed below (see, for example, Smith *et al.*, 2010). Secondly studies may seek to draw policy conclusions about which policies regarding industry structure, competition and regulation will be most beneficial.

In both cases the importance and motivation of efficiency analysis, and the reforms and policy interventions which may occur based on such analysis, is the delivery of efficiency savings (and productivity gains more generally) with the ultimate aim of delivering either lower prices to users of transport services, or reduced taxpayer support.

In sectors of the economy in which markets are a reasonable approximation to perfectly competitive, a measure of overall efficiency may simply be the profitability of the firm. Under perfect competition, prices are not influenced by the individual firm and therefore the more profitable the firm, the more it has been able to minimise costs of production and to produce the most valuable combination of goods in the eyes of consumers.

However, transport provision is in most cases a long way from being a perfectly competitive industry. In some sectors of rail operations (e.g. coal, commuters), for example, operators still have considerable monopoly power, whilst for this and other social reasons rail prices are often regulated by governments, who also play a key role in specifying passenger sector outputs. If the aim is to examine the efficiency of railway management, these factors must be allowed for. To the extent that railway managers (at least in the European passenger sector) have limited control over their outputs, the key issue is whether they produce them at minimum cost.

Further, in the absence of competition, regulatory pressure is needed to ensure that transport firms operate in an efficient manner. Given the standard problem of asymmetric information – firms know more about their costs than regulators - regulators need to arm themselves with additional information to overcome this asymmetry. Benchmarking – which involves comparing the efficiency performance of the

[21] The material in this section draws partly on Smith and Nash (2014).

regulated firm with other transport firms – is thus an important, indeed crucial part of the regulatory process. The data that is used in benchmarking could be drawn from other domestic transport firms, where possible, international benchmarks (which can be more problematic), or internal benchmarks (for example comparing the performance of different regions of an infrastructure manager).

Economists are used to distinguishing between technical efficiency and allocative efficiency. Technical efficiency is measured by whether output is maximised for a given level of inputs (or conversely inputs are minimised for a given output). The standard economic approach to examine this is to estimate a production function using econometric methods, although non parametric methods – such as data envelopment analysis (DEA) – has also been used. The different approaches are considered further in this chapter. Allocative efficiency considers whether the correct mix of inputs is used to minimise cost for a given level and quality of output.

Cost efficiency is the product of technical and allocative efficiency, and thus takes both technical and allocative efficiency into account. Cost efficiency measures are typically obtained by estimating a cost function or frontier, though such measures can also be obtained via DEA.

Arguably cost efficiency is the most relevant concept from a regulatory perspective since it is through cost reductions that prices can be reduced to users, or the burden of subsidies lowered. That said, in situations where it may be unrealistic for firms to optimise the capital input – as may be the case in regulated network industries – and / or where there is a lack of good and comparable data on costs or input prices, technical efficiency could be a more relevant measure.

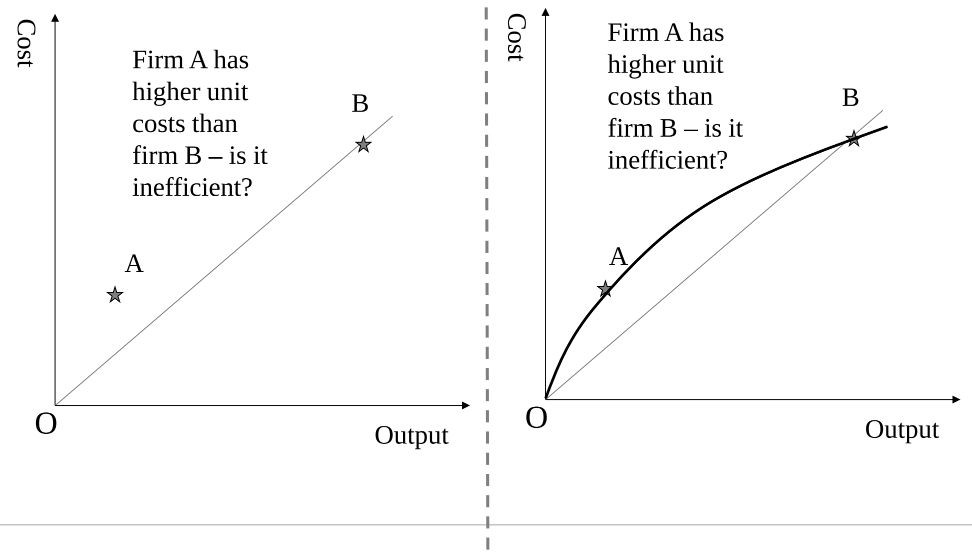
Before turning to discuss advanced techniques for measuring the relative efficiency of transport firms, it is first worth considering why it is important to go beyond simple partial productivity measures - such as (in a rail context) cost per train-km, cost per passenger-km or cost per route or track-km. In principle, benchmarking could simply proceed by collecting comparable data across firms and using these to compare unit costs across companies – thus raising the question as to whether advanced techniques are needed at all.

However, unit cost measures are only partial measures of efficiency performance and raise multiple questions. Firstly, such measures may not cover all costs and thus would not give an overall assessment of performance. For example, staff costs per train-km could be distorted in comparisons between firms if different firms take different out-sourcing decisions. Further, capital substitution possibilities may be ignored. Even if the measure of costs covers all costs, there is then a question as to how to characterise the outputs in multiple output industries where output may be described in terms of passenger-km, train-km or even represented by the size of the network, making cost per track-km another candidate measure. Further disaggregation is also possible as between freight and passenger traffic, or even further traffic-type disaggregation within the generic categories.

Typically, unit cost comparisons with different denominators can give very different rankings. Similar problems emerge in other regulated industries, such as the water sector, where unit costs may be calculated per customer or per unit of water delivered or per length of mains, with very different results potentially. It is thus problematic, since it is not always clear which is the preferred measure, and different rankings and efficiency scores result from using different measures.

An approach that simultaneously takes account of multiple denominators / drivers of costs is therefore needed - this being a key advantage of the econometric approach as multiple cost drivers can be included in the model at the same time. DEA can also handle multiple inputs and outputs, sometimes via a two-stage process. A second key advantage of adopting an econometric framework (and to some extent also a DEA approach) as compared to simple unit cost measures can be illustrated in Figure 1 below.

Figure 1.



Simple unit cost comparisons make an implicit assumption of constant returns to scale, which is very unlikely to hold in transport applications. Thus in the left hand panel of Figure 1 firm A appears to be inefficient because it has higher unit costs. However, once the presence of increasing returns to scale is permitted – as represented by the right hand panel of Figure 1 – it becomes apparent that the reasons for firm A's higher unit costs is that it operates at small scale. Given its scale – which will usually not be under the control of management – firm A turns out to be operating on the cost frontier and is thus an efficient operator.

By using an econometric framework the shape of the cost frontier can be estimated as part of the modelling framework. Indeed the assumption of constant returns to scale can be tested directly. Likewise DEA methods can be adapted to take account of the possibility of non-constant returns. Having controlled for the underlying technology, efficiency comparisons can then be more accurately assessed.

Before proceeding to look at the methods that can be used to operationalise a benchmarking framework, it is worth noting another set of important factors that are needed in a benchmarking framework and which more advanced methods – that is, going beyond simple unit cost measures of key performance indicators – can deal with at least to some extent. These include the need to take account of:

- variations in quality between firms and over time;
- observed (or unobserved) heterogeneity between firms that impacts on costs, but is not related to management performance – these are likely to be persistent factors that do not vary over time, and may relate to factors such as the weather conditions or topography of the country or region being benchmarked;
- the natural cycle of asset replacement and enhancements to assets which can cause expenditure to vary substantially from year to year such that unit cost measures at a point in time may be highly misleading.

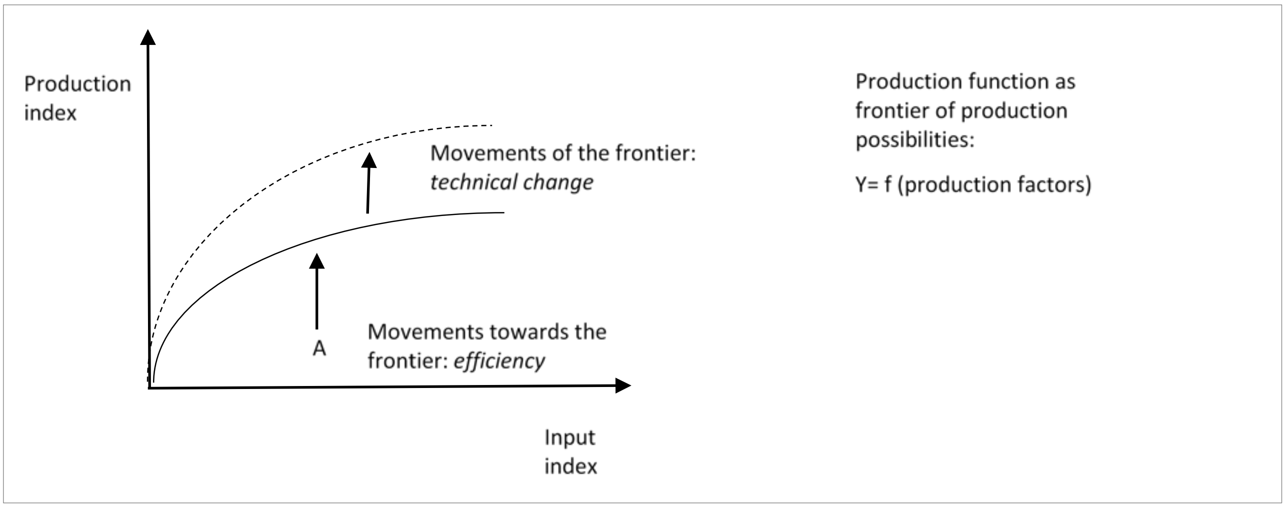
In the subsequent sections of this chapter we document further the methods that have been developed to undertake benchmarking – setting out the essence of the methods whilst also reflecting the state of the art and explaining how the combination of methods, data and regulatory judgment can combine to deliver a powerful benchmarking framework with wide social benefits.

3.3. Benchmarking by productivity indexes

3.3.1. The concept of productivity

The concept of productivity concerns both efficiency and technological change in production. Efficiency relates to the quantity of output obtained from the inputs in the production process, given the technology available (the effect of time is neglected). Technical efficiency improvements imply movements of the observed unit towards "best practises" or the reduction of organizational inefficiencies (Diewert and Lawrence 1999). If we introduce time varying effects, allowing for the change of the productive combination, we can also measure the technological change. Technological change can arise through disembodied forms concerning new scientific results and organizational techniques or through forms embodied in new products or services. In Figure 1, the approach of firm A to the best production function reflects an efficiency improvement, while a technological change allowing to obtain a greater amounts of output with the same set of observable inputs produces a shifting up of the production frontier.

Figure 2.



A correct analysis of productivity trends requires isolating output variations due to the simple increase of inputs from those connected to technological change. It is possible to see the changing in productivity in the Solow (1957) formulation of a production function. Given Hicks neutral technical progress (ratio of marginal product of capital to that of labor remains unchanged at constant capital labor ratio) and a production function:

$$Y_t = A_t f\,(K_t, L_t, M_t)$$

where Y_t is output, $f\,(\cdot)$ is a function of observable inputs capital K_t, labor L_t, materials M_t , and A_t is the factor-neutral index of technical progress (increasing function of t that measures cumulated effects of shift overtime), productivity could be investigated by A_t. The coefficient A_t represents the efficiency, time varying, and $d\,A_t$ represents the technical progress.

The change over time of total product Y is:

$$d\,Y_t = d\,A_t\,f\,(K_t, L_t, M_t\,) + A_t\,(d\,f_t)$$

Therefore, we have:

$d\,A_t\,f\,(K_t, L_t, M_t)$ = the change of output from technical progress

and

$A_t\,(d\,f_t) = A_t\,(\partial Y / \partial K * dK + \partial Y / \partial L * dL + \partial Y / \partial M * dM)$ = the change of output from the variation of use of observable factor inputs, constant A_t.

3.3.2. Different productivity measures

The measures used in the field of productivity analysis are numerous. The choice depends on the objectives of the analysis and on the available data. We can find single factor productivity measures (output compared to a single measure of input) concerning labour or capital and multifactor productivity measures (output compared to a set of inputs) related to labour-capital-energy-materials measures.

Once the productivity measures have been identified, there are several methodological approaches allowing to move to the implementation phase. From this viewpoint, we have two fundamental approaches: parametric or non-parametric estimates.

Relying on econometric techniques, we can estimate the parameters of a production or cost function and in this way we can determine the productivity. When we rely on non-parametric approaches, instead, some properties of the production function are used to identify empirical measures such as index number or efficiency scores such as those obtained through Data Envelopment Analysis ("The growth accounting approach to productivity measurement is a prominent example for non-parametric techniques"; OECD, 2001, p.13).

3.3.3. Accounting productivity indexes [22]

The approach of the American Productivity Center (APC, 1981)

ROI is the most important measure of firm's accounting and financial performance. The ratio (operating profits/ assets) can be usefully decomposed as follows:

ROI = (Operating profits/ Sales)·(Sales/Assets)

The first term of the right hand side is the return on sales (ROS), while the second term is an index of the turnover of assets (asset productivity). The ROS ratio might be expressed as:

ROS = 1- (Sales-operating profits)/sales = $1 - \Pi^{-1}$

Π is the revenue to cost ratio (sales / operating costs), which allows to compare the profitability in different periods. In case of constant input and output prices, Π can be treated as an index of productivity. By comparing Π in two periods, it is possible to examine the variations in productivity (the prices being constant) separately from the effects of changes in prices (the quantities being constant).

[22] The material in this section draws partly on Fraquelli, Vannoni 2000.

APC method: the Revenue to Cost Change Ratio (Π^t/Π^0) disaggregation

Let us define:

- Y_i^t the quantity of output i in period t, $i = 1, \ldots, n$
- p_i^t the price of output i in period t, $i = 1, \ldots, n$
- X_v^t the quantity of the variable input v in period t, v = 1, . . . ,m
- w_v^t the price of the variable input v in period t, v = 1, . . . ,m
- X_f^t the quantity of the fixed input f in period t, f = 1, . . . ,g
- w_f^t the price of the fixed input f in period t, f = 1, . . . ,g
- Π^t the ratio between sales and costs in period t.

The revenue to cost ratio in period t (Π^t) will be:

$$\Pi^t = \Sigma_i Y_i^t p_i^t / (\Sigma_v X_v^t w_v^t + \Sigma_f X_f^t w_f^t)$$

The index of its variation (the revenue to cost change ratio (RCCR)) between t and $t = 0$ will be (Π^t/Π^0):

$$(RCCR) = (\Sigma_i Y_i^t p_i^t / (\Sigma_v X_v^t w_v^t + \Sigma_f X_f^t w_f^t)) / (\Sigma_i Y_i^0 p_i^0 / (\Sigma_v X_v^0 w_v^0 + \Sigma_f X_f^0 w_f^0))$$

The APC index gives the possibility to detect a productivity change ratio and a price recovery ratio:

$$(\Pi^t/\Pi^0) = (\Sigma_i Y_i^t p_i^0 / (\Sigma_v X_v^t w_v^t + \Sigma_f X_f^t w_f^t)) / (\Sigma_i Y_i^0 p_i^0 / (\Sigma_v X_i^0 w_v^t + \Sigma_f X_f^0 w_f^t)) \cdot$$

$$\cdot (\Sigma_i Y_i^t p_i^t / (\Sigma_v X_v^0 w_v^t + \Sigma_f X_f^0 w_f^t)) / (\Sigma_i Y_i^t p_i^0 / (\Sigma_v X_i^0 w_v^0 + \Sigma_f X_f^0 w_i^0))$$

The first term is an index of productivity change, while the second term is an index of price recovery change.

Banker disaggregation of the Revenue to Cost Change Ratio (Π^t/Π^0)

Banker *et al.* (1993) extended the APC method by disaggregating the index in four components (pricing, efficiency, product mix, capacity utilization). The Banker method gives the possibility to separate the pure effects of price and productivity changes from changes in capacity utilisation and in the product mix. A firm might be profitable due to higher prices, or because it is very efficient in activating inputs or in choosing the product mix, or because it has not overinvested in capacity. For applications of the APC method, see Banker *et al.* (1996 a, b); for applications of the Banker model see Fraquelli and Vannoni (2000).

Total factor Productivity indexes

Measuring actual production data

The main economic activity of a firm consists in transforming a set of inputs into one or more outputs. Given a set of inputs, a more efficient technology allows producing a larger quantity of output. As different inputs have different costs, this point turns into a cost minimization problem: a firm is considered more efficient than another one if it is able to obtain the same output by using a cheaper set of inputs.

There are, indeed, different methods that might be used to obtain the same output; for example, in the local public transport industry we can see different inputs per vehicle-kilometre concerning work forces, fuel, maintenance materials and transportation equipment.

The nature and the required quantity of each input depend on several factors, such as the home country economic situation, the relative price of inputs, the particular business links between firms and the different legislation in each country. Regulation may be considered as part of this latter factor and influences the economic decisions of firms.

When we try to construct productivity measures from firm production data, some problem can arise.

The first one relates to the degree of homogeneity among outputs. Given the same level of passenger transported, we can have different quality of service. Quality is expensive and should be valued. At the same time, we have to pay attention to the negative outputs such as the level of pollution linked to the production process: investments, in the bus service sector, can be devoted to reduce harmful emissions moving towards the use of cleaner energy sources. Further, firm data sometime do not contain measures of output quantities. Revenues could be observed from the firm reports, but we need real values referred to the price level of a common year. As inflation rates affect the values of sales, erroneous increases in efficiency might result from analyses based on current values. A correct analysis should consider constant prices by properly choosing a reference year.

The second set of problems concerns the inputs. The measurement of inputs appears delicate because an underestimation of their value results in an untrue improvement in productivity. As regards the labour inputs, we can use the number of employees or the hours worked, or alternative measures such as the annual cost, properly deflated. The last measure has the advantage of highlighting the quality of the resource used in relation to the marginal productivity of the different units of work employed. However, labour quality can be embodied in the productivity measure: “Attempts to capture labour quality differences in labour measures rather than productivity are the impetus behind using the wage bill to measure labour inputs rather than the number of employees or employee-hours. The notion is that market wages reflect variations in workers’ contributions to production; firms with more productive workers will have a higher wage bill per employee. Of course, there are problems with this approach: wage variation might reflect the realities of local labour markets, or causation could be in the other direction, if more productive producers earn rents that are shared with or captured by employees” (Syverson, 2011, pp.340).

The evaluation of the dimension of the physical capital appears even more problematic. Capital could be measured using the firm’s book value of its capital stock or depreciations. There are some doubts about these measures as a correct proxy of the service of capital and about the need to deflate such variables with appropriate methodologies (that we will better discuss later on). About the materials and services, there are problems similar to those highlighted for the output because usually we have the amount of the expenses and not the quantity of inputs.

The third set of problems concerns the nature of the multifactor productivity indexes to be used. The construction of the indexes needs to weight properly the individual outputs and inputs. The weights are closely related to the index used and the nature of the underlying production function. In the next paragraph we will examine the most used and appreciated indexes.

Which indexes? Which periods to compare? Chained or direct comparisons?

The most used index number formulae are the Laspeyres, Paasche, Divisia, Fisher and Törnqvist indexes (Diewert, 1981). The choice is not random. The economic theory can indicate the different capacity of the indexes in correctly approximating the technology outlined by the production or the cost functions. Some indexes, basically, assume a particular production technology. When few observations are available and it is not possible to make econometric estimates, the choice of the index becomes difficult, but the literature often suggests the use of Törnqvist or Fisher index number.

The Fisher index is a geometric average of the Laspeyres and Paasche indexes and the Törnqvist index is a weighted geometric average of outputs and inputs. Törnqvist and Fisher indexes are “superlative” index numbers (Diewert, 1976). “For example, the Törnqvist index is exact for the translog flexible functional form (…it can be directly derived for the translog flexible functional form…) – a widely used specification in empirical economics. Thus, if one accepts a translog form as an approximation to a production function, and uses standard assumptions about producer behaviour, the Törnqvist quantity index provides an exact formulation for inputs and outputs. An index that is exact for a flexible functional form is called "superlative"." (OECD, 2001, p.88). "Törnqvist index is superlative in considerably more general sense than shown by Diewert. We are not aware of other indexes that can be shown to be superlative in this more general sense." (Caves *et al.*, 1982, p. 1411).

When we have decided the specific number index formulae, we have to pay attention to the periods to compare. Comparisons could be made over several periods or concerning two specific periods (for example: period 0 and period 4). We can choose the first or the last observation as the base, or adopt the chain approach. Given a time series, period *t-1* provides the base for period *t*. In the case of fixed-based indices, like Laspeyres (based on weights of the first year) or Paasche (based on weights of the current year), a change in the relative prices of single outputs or inputs gives too much (or too little) weight on goods characterized by a fall (or a rise) in the relative prices. Chain weighted indices are less sensitive to a substitution bias. "The economics literature as well as the SNA 93 are quite unanimous in this respect: for inter-temporal comparisons, changes over longer periods should be obtained by chaining: *i.e.* by linking the year-to-year movements" (OECD, 2001, pp. 83).

Laspeyres and Törnqvist indexes [23]

Laspeyres and Törnqvist indexes, briefly described below, are the most used indexes. The Laspeyres index is represented by a quantity index, which relates outputs and inputs valued at constant prices. The productivity ratio between period t and t=0 by Laspeyres formula is:

$$TFP_L = (\Sigma_i Y_i^t p_i^0 / \Sigma_i Y_i^0 p_i^0) / (\Sigma_h X_h^t w_h^0 / \Sigma_h X_h^0 w_h^0)$$

where *i* indicates the different goods, *h* indicates the different inputs and *t=0* is the time basis. As inputs are simply added up, the model implies perfect substitutability.

Tornqvist index is an approximation in the discrete case of the Divisia indexes (Diewert, 1981). The latter are derived from Solow's production function, which implies constant returns of scale and neutral technological change (par. 2.2.1). In the situation of one output and two inputs (labour and capital) it can be expressed as follows:

$$Y_t = A_t f(L_t, K)$$

where A_t is the neutral technical progress. To model the production function we can start from a Cobb Douglas functional form that is linear in logs, homogenous of degree one and a first-order approximation to any production function:

$Y_t = A_t L_t^a K_t^b$, with *a* indicating the relative weight of labour (labour output elasticity) and *b* indicating the relative weight of capital (capital output elasticity). By taking the logarithm and differentiating with respect to time we get the impact of the variation of the technical progress over time:

$$dA_t / A_t = dY_t / Y_t - (a_{t^*} dL_t / L_t + b_{t^*} dK_t / K_t)$$

We can rewrite the previous equation as: $A_t^* / A_t = Y_t^* / Y_t - (a_{t^*} L^*_t / L_t + b_{t^*} K^*_t / K_t)$, where a dot over a symbol means the derivative respect to time and A_t is the Divisia index of the residual.

Considering the case one good (Y) and g-inputs (X):

[23] The material in this section draws partly on Fraquelli, Vannoni (1996).

$$A_t^* / A_t = Y_t^* / Y_t - (\Sigma_h^g \; s_{ht} * X_{ht}^* / X_{ht})$$

where s_h is the share of the receipts going to the *h*th input. "The apparent shift of the function over time, A^* / A, is the growth of the Divisia total factor productivity. (Star and Hall, 1976, p. 258)".

To get the index of total factor productivity, we can integrate the previous equation:

$$A_T / A_0 = (Y_T / Y_0) \; exp \; (- \Sigma_h^g \int_0^T s_{ht} * X_{ht}^* / X_{ht} \; dt)$$

Normalizing and setting $A_0 = 1$, A_T will be the index of productivity at time T.

In empirical works, the calculation of the Divisia index is not easy because we need a correct value of the share s_h over time. Star and Hall (1976, p.258) suggest a good approximation of the value of share s_h .To calculate the value of the Divisia index at time T, they suggest the use of constant values for the shares ("that give the same index as the true fluctuating shares..") using data only from the periods 0 and T.

First, using a constant share, they demonstrate (p. 258) that:

$$s_h \; log \; (X_{hT} / X_{h0}) = \int_0^T (s_{ht} * X_{ht}^* / X_{ht}) dt$$

And substituting in the previous Divisia index of total factor productivity we have:

$$A_T / A_0 = (Y_T / Y_0) \; exp \; (-\Sigma_h^g \; s_h \; log \; (X_{hT} / X_{h0})) =$$

$$A_T / A_0 = (Y_T / Y_0) \; / \; exp \; (\Sigma_h^g \; log \; (X_{hT} / X_{h0})^{sh}) =$$

$$A_T / A_0 = (Y_T / Y_0) \; / \; exp \; (log \; (\Pi_h^g \; (X_{hT} / X_{h0})^{sh}) =$$

$$A_T = (Y_T / Y_0) \; / \; (\Pi_h^g \; (X_{hT} / X_{h0})^{sh}$$

where the approximation of the true s_h could be the simple average of s_{ht} at the beginning and end of the period (p.259): $s_h = ½ (s_{h0} + s_{hT})$.

The index of inputs of the previous formula "…has been advocated as a quantity index by Törnqvist … "(Star and Hall, 1976, p.259).

The Törnqvist discrete approximation index of total factor productivity takes the following form:

$$\text{TFP}_{Tt,\, t-1} = (Y_{it} / Y_{it-1}) / \Pi_h^g \; (X_{ht} / X_{ht-1})^{1/2(Sht+Sht-1)}$$

The denominator can be considered as a weighted geometric mean of the input ratios between two subsequent years. The weights can be represented by the average expenditure shares. In fact, if we make the assumption of perfect competition and constant returns to scale "….. the elasticities equal the share of revenues paid to each input. This makes constructing the "*Sh,*" simple. Materials' and labour's shares are typically straightforward to collect with the wage bill and materials expenditures data at hand. Capital's share can be constructed as the residual, obviating the need for capital cost measures"…. "cost-share-based TFP index numbers…. are easy to construct and offer the robustness of being a nonparametric first-order approximation to a general production function…". (Syverson, 2011, p.332).

As the Laspeyres indexes are used with fixed base prices and Törnqvist indexes with variable-base prices, the latter take into account relative price changes; this means that they are sensitive to strategies that reduce the weight of those inputs that have become more expensive.

The cost of capital

As to the estimation of the cost of capital, we have to move away from the simple consideration of financial charges since a correct measure of the total cost of capital should include the opportunity cost of equity. We need the value of the amount of capital invested and the cost rate to apply to that value.

Weighted Average Cost of Capital (WACC) is the most used methodology, by regulators, for the determination of the cost of capital. Using productivity as a benchmark of the regulated enterprises, the evaluation of the cost of such input must be consistent with above model. WACC is applied to the estimation of the assets value of the regulated firms (RAB, *Regulatory Asset Base*).

WACC is the return required by equity and debt financing firm investments. The most commonly used formula is:

$$WACC = (1-g) + gRd$$

Where:

Re is the cost of the equity (E);

Rd is the cost of the debt (D);

g is the gearing ratio given by D/(E+D);

(1-g) is the weight of the equity on total funding volume, given by E / (E+D).

Gearing ratio. *Gearing ratio* and the weight of equity on total assets can be obtained by:

- the amount of the value of debts and equity in terms of their book value. If based on historical time series, it is not suitable when we need a ratio for future projections;
- the amount of the market value of debts and equity. In this way we have the actual value of the firm, but there are problems concerning the market volatility;
- the "notional" gearing ratio given by an efficient financial structure of a firm. This ratio, often used by public utilities regulators, is based on a benchmarking approach.

The cost of equity. The cost of equity is usually determined as sum of the risk-free rate (related to an investment with zero risk) and the equity risk premium. The Capital Asset Pricing Model (CAPM), developed by Sharpe (1964), Lindtner (1965) and Mossin (1966) appears the standard way to estimate the cost of equity. The model relates to the return of a share paying attention to the relative volatility compared to the average tendency of the stock market.

The cost of the debts. Also the cost of the debt is composed by two elements, the risk-free rate and the debt risk premium, DRP (the lender compensation for the risks associated to the default probability of the regulated enterprise).

The WACC computed using the formula reported above is often referred to as "vanilla" WACC, since it does not consider the impact of taxes on the average cost of the capital. Adjusting the cost of equity upwards to cover taxes and the expected return of shareholders, the pre-tax WACC will be:

$$WACC = ((1-g) / (1-\tau))+gRd$$

where τ is the company tax rate.

The amount of capital invested and depreciation

The estimation of the total capital invested in the firm could be obtained by subtracting from net assets (at CPP prices) the amount relative to commercial debt and other current liabilities).

The calculation of consumption of fixed capital can be obtained by employing two different approaches.

The first one is a simple revaluation of the book value as it is recorded in the profit and loss accounts. It is important to note that this value does not represent a good measure of physical depreciation and economic obsolescence of installations and machinery, as fiscal reasons and inflation rates have a high influence on it and may lead to biased values that are not economically acceptable.

The second approach results from the application of the perpetual inventory method (Meinen, *et al.*, OECD, 1998, Blades, OECD, 1998) and is based on a subdivision of the capital stock in classes according to the year of acquisition of the assets. It enables to obtain a "real value" of assets for each year. Average depreciation rates drawn from annual reports can be applied to the gross values of fixed assets.

This method takes account of the fact that for each year the value of fixed assets is the result of a stratification process, with investments and disinvestments respectively increasing and decreasing the amount recorded at the beginning of the year. Starting from a base year, whose asset values often need to be expressed in terms of market values, additions and subtractions could be calculated at constant prices following the equation below:

$$K_{t+n} = K_{t+n-1} + I_{t+n} / IP_{t+n} - D_{t+n} / IP_{t+n-z}$$

where K_{t+n} indicates the value of fixed assets for year $t+n$ at the prices of year t and IP_{t+n} is the price index for year $t+n$. The use of the price index of year $t+n-z$ for deflating withdrawals points out the fact that in general disinvestments are not relative to machinery purchased or constructed during the current year but instead they reflect withdrawals and sales of old equipment.

In many cases this approach leads to very high values of gross fixed assets as compared to the book values recorded in the balance-sheets; this reflects the fact that accounting systems based on original costs do not take into account the effects of inflation, that could be very relevant especially with respect to the base year. For applications of Laspeyres and Törnqvist indexes see Fraquelli and Vannoni (1996).

3.4. Benchmarking by DEA approach and bootstrapping [24]

DEA is a benchmarking approach based on non-parametric programming techniques to evaluate the efficiency of homogeneous units in multiple inputs – multiple outputs contexts. The most efficient firms contribute to form a "best-practice frontier" that will be useful to compare the performance of other firms. "DEA's empirical orientation and absence of a priori assumptions have resulted in its use in a number of studies involving efficient or best-practice frontier estimation in the non-profit, regulated, and private sectors. (Cook W. D., Joe Zhu, 2013)".

The measurement of efficiency based on the comparison of a firm (or more generally, decision making unit DMU) with its efficient counterpart lying on the technological frontier dates back to Debreu (1951) and Farrell (1957). The basic idea is that the capacity of each unit to transform inputs into outputs cannot be assessed as efficient or inefficient in itself, but only in connection to an external (optimal) benchmark. A firm can be regarded as efficient if it is not possible to identify a benchmark whose processing capacity is better.

[24] The material in this section draws partly on Erbetta, Fraquelli (2010).

The efficiency of a firm can be measured in terms of minimising the use of the inputs given a certain amount of output (input-oriented approach) or, alternatively, in terms of maximizing the output given certain input level (output-oriented approach). Of course, the choice between these alternatives depends on the assumption about the firm's behaviour, i.e. the degree of control that individual firms may have in terms of input saving or output expansion.

Charnes *et al.* (1978) and Banker *et al.* (1984) subsequently extended the original intuition by Debreu and Farrell by means of a mathematical approach, based on optimization algorithms, aimed at deriving a technological frontier from a set of observed units. This approach takes the name of Data Envelopment Analysis (DEA). For a comprehensive discussion of DEA models see Cook W. D., *et al.* (2014), Coelli *et al.* (2005).

Analytically, let $Z = \{(x_n, y_n) \mid n = 1, ..., N\}$ be a set of N firms for which information is available on the amount of K inputs and M outputs (x_n is a K-dimensional input vector and y_n is a M-dimensional output vector). The input-oriented model in case of variable returns to scale (VRS) for the DMU n_0, consists in the following linear programming problem (Banker *et al.*, 1984):

$$\min_{\theta_{n0}, \lambda_n} \theta_{n_0}$$

s.t.

$$y_{m,n^0} \leq \sum_n \lambda_n y_{m,n}, \qquad m = 1,..., \mathrm{M}$$

$$\theta_{n_0} x_{k,n^0} \geq \sum_n \lambda_n x_{k,n}, \qquad k = 1,...,K$$

$$\lambda_n \geq 0, \qquad n = 1,..., \mathrm{N}$$

$$\sum_n \lambda_n = 1$$

where λ_n is a N-dimensional vector of weights (or "intensity variable") assigned by the linear programming algorithm to each firm in order to determine input/output linear combinations belonging to the efficient frontier. The solution to the linear programming problem provides the optimal value of the scalar θ_{n0}, such that $0 < \theta_{n0} \leq 1$. This represents a measure of radial projection towards the frontier. A value of the latter coefficient equal to 1 indicates full efficiency, while a value below 1 indicates an inefficiency of $1 - \theta_{n0}$. For instance, a score θ_{n0} equal to 0.8 indicates an inefficiency of 0.2, thus suggesting that the firm may reduce its of input use by 20 percent without changing the output. The last equality constraint provides flexibility in terms of scale economies, allowing to compare each unit against a subset of peers characterised by similar size, thus avoiding comparison between units operating at different returns to scale regimes. The efficiency measure calculated in this way is thus net of any effect on average productivity due to economies of scale, and efficiency differential would only reflect managerial factors.

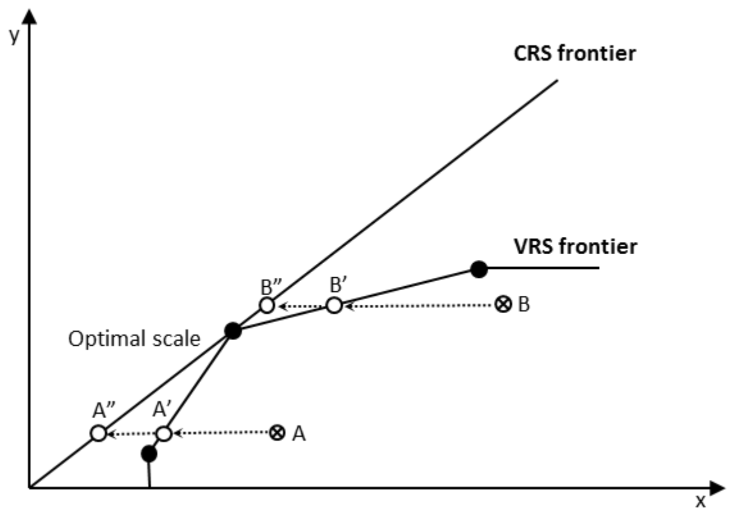
Dropping the last constraint, the efficiency scores are calculated under the hypothesis of constant returns to scale (CRS), in which the scale and pure managerial effects are not separated. Analytically, the scale efficiency (SE) can be measured as the ratio of CRS and VRS efficiency scores:

$$\text{SE} = \theta_{n0CRS} / \theta_{n0VRS}$$

A value of scale efficiency equal to 1 indicates that the firm operates at the most productive scale size. Scale efficiency values lower than unity may be due to under-sizing or over-sizing problems. By comparing observed and optimal scale size, the model allows to identify the presence of local increasing returns to scale (IRS) or decreasing returns to scale (DRS).

Figure 3. illustrates the methodology by means of a simplified one-output one-input technology. The CRS and VRS frontiers are determined by "enveloping" the observed units. The distances B'B and B"B identify respectively the CRS (overall) and VRS (pure managerial) inefficiency. In contrast, the distance B"B' identifies a scale inefficiency associated with over-sizing with respect to the most productive scale size, thus indicating the presence of local decreasing returns to scale. Analogous measures of (in) efficiency may be calculated with respect to the unit A. In this case, the scale inefficiency is due to under-sizing and reveals the presence of local increasing returns to scale.

Figure 3. CRS and VRS efficient frontier



Source: Erbetta, Fraquelli, 2010.

Bootstrapping

DEA efficiency scores are, by construction, biased since they refer to a technology that is *estimated* starting from the sample data, and that is an *approximation* (and also a subsample) of the true one, which is unobserved (Simar and Wilson, 1998). The presence of bias can lead to unreliable efficiency scores and ranking of units. Bootstrap techniques, following Simar and Wilson (1998, 2000a, 2000b), represent a useful approach to correct the bias of DEA results. They are also a way to run inference analysis on the estimates. The bootstrap (Efron, 1979) is an approach (usually computer-based) that relies on repeated resampling with replacement from the original sample, in order to create random replicates of the latter. Statistics can be computed for each replica.

The number of replicates is usually large, thus allowing to obtain a distribution for the relevant statistics (Bogetoft and Otto, 2010). To give a very simple idea of the application of bootstrapping to DEA, it is important to point out that the resampling occurs after a preliminary efficiency estimate, on the obtained efficiency scores. Each new set of scores obtained through resampling is manipulated to simulate a continuous distribution, while avoiding values larger than one. Subsequently, the manipulated scores are applied to the original data to modify them in order to obtain a new frontier. Then, new efficiency scores are computed for the original units with respect to this new frontier. The process is repeated many times (e.g. 2000), allowing to obtain a distribution for each efficiency score, thus allowing inference and bias correction (Bogetoft and Otto, 2010; Simar and Wilson, 1998, 2000a, 2000b).

Table 1. Use of DEA methods in electricity benchmarking in Europe for Distribution Operators (DSO) and Transmission Operators (TSO)

Direct or not direct implementation in regulation	Countries that rely on DEA for benchmarking of DSO	Countries that rely on DEA for benchmarking of TSO
Direct implementation in regulation	Austria, Germany, Norway, Slovenia	Belgium, Denmark, Finland, Germany, Iceland, Netherlands, Norway, Portugal
Not direct implementation in regulation	Belgium, Iceland, Sweden, Switzerland	Austria, Estonia, France, Great Britain, Greece, Italy, Lithuania, Luxemburg, Spain, Sweden

Note: The table reports both exclusive DEA use and DEA combined with other methods.
Source: Our elaboration from Agrell and Bogetoft (2017), p.11.

Undesirable outputs and the directional output distance function approach

Previously, we discussed the nature of the outputs used for analysing the activity of transport companies and relevant for benchmarking. The available measures can be numerous, as demonstrated by the reviews given in Chapters 3 and 4. As component of the outputs, quality is also cited, noting that "quality is expensive and should be valued". Quality can be measured in terms of good or bad outputs, depending on the nature of the available variables and on the issue concerned. A simple example is the maintenance of the road surface. The increase of the vehicle-KM represents a desirable output, but implies more maintenance. The lack of maintenance is made evident by kilometres of rough road and generates delays and interruptions of traffic, hydroplaning, breakage of cars and more accidents. We can reduce such undesirable outputs, but it is expensive. The phenomenon can be regulated by various constraints, but it is known that it is practically impossible to produce complete contracts able to completely avoid any opportunistic behaviour.

A good support to reduce information asymmetries could be represented by a measure of efficiency that includes quality. If the quality level is measured by bad outputs variables, the directional distance function approach can be useful: "…joint production of good and bad outputs has gained attention in the literature on efficiency and productivity. One can distinguish between two lines of inquiry: (i) how to model joint production of good and bad outputs and (ii) how to account for reductions of bad outputs. (Chung and Fare, Discussion paper, 95-24, 1995, pg. 1)". The main idea behind this approach is that, in a certain production process, good and bad outputs are produced together. Usually, the latter are not avoidable (it is not possible to reach a "zero-level" of bad outputs); their proportion, however, can be reduced, but this reduction is costly. A directional (output) distance function model allows to measure efficiency in terms of good output maximization and contemporaneous bad output minimization, for given level of inputs. It relies on a non-parametric framework similar to standard Data Envelopment Analysis. More flexible approaches, oriented to optimizing also (some of the) inputs, can also be considered.

As an example, Martini, Manello and Scotti (2013) apply a directional distance function framework to the airport industry.

3.5. Parametric frontier techniques: stochastic frontier models, cross-sectional or with application to panel data- time-invariant models, time varying models - new proposals

As noted above, the focus by regulators on cost reduction and in turn reductions in user prices or subsidies makes efficiency analysis through the estimation of a cost function advantageous as compared to other possibilities that only address technical efficiency (that is, production or distance functions). There is an additional practical reason, namely that in many contexts cost data is more reliable and better reflects the inputs of production. In particular, physical measures of inputs such as staff numbers are greatly affected by the degree of contracting out, and measures such as the size of network or number of vehicles do not closely reflect the age, condition or amount of investment that has gone or is going into different networks at different times. It is also not possible to represent all inputs in a physical measure, leaving another input category, “other costs”, which either has to be included as a monetary value or in some cases is ignored.

Even for cost data, it should be said, there remain problems of inconsistencies in treatment of costs such as depreciation and interest, particularly in international comparisons. The problem is not simply different assumptions about asset lives. In some cases, where assets are purchased with grants, no depreciation or interest is entered into the accounts. In some cases historic debts have been written off; in other cases interest is still charged on them. Getting consistent data remains a challenge. Where close working with the industry or regulator is possible, rather than relying on published data, there is a greater possibility for better overcoming these challenges.

However, overall, the cost function approach does at least ensure that all inputs are considered and the allocative efficiency (or inefficiency) associated with using different input combinations accounted for; and ensures that the capital input is represented in a form that goes beyond mere consideration of the size of the network.

The remainder of this section focuses on the technical aspects of efficiency analysis conducted through econometric methods in the context of a cost function. The same methods can be applied to other functions such as production and distance functions which are discussed briefly at the end.

Cost function estimation

Derived from assuming cost minimisation in a production process, the cost function relates costs (C) to the level of outputs (Y) and input prices (P) and, where data is available over time, some measure of how costs change over time (t) as a result of technical change. Thus

$$C_{it} = C(Y_{it}, P_{it}, t)$$

It therefore automatically allows for one key issue in comparing costs between railways in different countries, and in a single country over time, namely different input prices.

There have been a vast array of functional forms proposed for the cost function. Notable developments include the constant elasticity of substitution (CES, Arrow et al (1961)) and generalised Leontief (Diewert, 1971). The most widely employed cost function is the Translog (Christensen, Jorgenson and Lau 1971, 1973). The Translog nests the simpler, and widely-used Cobb Douglas function (see, for example Beattie and Taylor, 1985) as a special (restricted) case, however it is not derived from any production function using duality theory.

Instead the Translog cost function is usually presented as a functional form which is a second order approximation to any cost function rather than being derived directly from economic theory. The general form of the Translog cost function for m outputs and n inputs is represented as:

$$\ln C = \alpha_0 + \sum_{i=1}^{m}\beta_i \ln y_i + \sum_{i=1}^{n}\gamma_i \ln w_i + \frac{1}{2}\sum_{i=1}^{m}\sum_{i=1}^{m}\beta_{ij}\ln y_i \ln y_j$$
$$+\frac{1}{2}\sum_{i=1}^{n}\sum_{i=1}^{n}\gamma_{ij}\ln w_i \ln w_j + \frac{1}{2}\sum_{i=1}^{m}\sum_{i=1}^{n}\delta_{ij}\ln y_i \ln w_j$$

The function includes both first and second order terms in all variables. Importantly, the Translog allows for elasticities and marginal costs to vary flexibly with the level of outputs and prices. In this sense the Translog does have appealing economic characteristics, such as the ability to deal with varying degrees of returns to scale and density as firm size varies.

Finally we note that the Translog cost function is often estimated along with the factor share equations. Factor share equations are expressions for the proportion of total cost used by each input and are derived using Shephard's (1953) lemma as the partial derivative of the cost function with respect to each input price. Estimation can then proceed using Zellner's (1962) Seemingly Unrelated Regression (SUR) which is more efficient (in terms of estimation) than single equation ordinary least squares.

When it comes to measuring outputs, the need to distinguish between scale and density effects or the choice between different measures of output is only part of the wider issue of how to account for the heterogeneity of outputs that may occur in transport applications (for example passenger versus bus-km or different types of rail freight traffic). One way to deal with the heterogeneity in outputs is to group outputs (denoted y) into m groups and include a further set of r variables which characterise the outputs (denoted q):

$$C(y_1,\ldots,y_m,q_1,\ldots,q_r,w_1,\ldots,w_n)$$

The move from potentially hundreds or thousands of outputs to a more manageable number of m outputs is obviously a simplification. However the inclusion of output characteristic variables is an attempt to reintroduce heterogeneity in outputs back into the model. Such variables may include revenue measures (such as passenger-km and freight tonnes-hauled) where availability measures such as train-km or vehicle-km are adopted as output and vice-versa. As such it can become a little ambiguous as to what variables represent outputs versus output characteristics versus network size.

Wheat and Smith (2014) is an attempt to introduce heterogeneity by means of a hedonic cost function (Spady and Friedlaender, 1978 and Bitzan and Wilson, 2008). Under this approach there is only one output as opposed to n outputs (this is relaxed in some applications, such as Wheat and Smith, 2014). Second the output and output characteristic variables enter into their own function and this then enters into the general cost function. The benefit of this approach is a more parsimonious model. It is perhaps surprising that there have not been too many applications of hedonic cost functions in transportation operations. In Wheat and Smith (2014), the model included three outputs (train hours, route length and number of stations operated) and many characteristic variables relating to the train hours output. This analysis provided rich insights into the impact of output heterogeneity on economies of scale and density (see section 4), whilst enabling a parsimonious model.

Stochastic frontier methods: introduction

The above discussion has focused on the relationship between costs, outputs, output characteristics and input prices. As noted earlier, a key motivation for policy makers is to understand the relative cost efficiency of transport operators. The cost function relationships discussed above can be augmented to allow the relative efficiency of companies to vary and for this degree of variation to be estimated.

The efficiency measurement literature cites three functions which may be estimated, depending on the appropriate behavioural assumption: cost functions, production functions or distance functions (the latter two are focused on technical efficiency as noted above). Cost functions are often used in highly regulated environments such as transport applications (and particularly railway infrastructure), where it may be seen as appropriate to view transport firms as seeking to minimise cost for a given level of output (where the latter is more or less determined by government). In this section we focus on cost function relationships or, more precisely now that we are introducing inefficiency into the approach, cost frontier relationships.

The simplest econometric approach is to use the method of corrected ordinary least squares (COLS). This method proceeds by ordinary least squares (OLS), but then shifts the regression line down by the amount of the largest negative residual (for the cost function case), thus translating an "average" cost line into a cost frontier. However, like DEA, the COLS method is a deterministic approach which does not distinguish between genuine inefficiency and statistical noise when looking at deviations from the frontier. It is however, with suitable adjustments, widely used by UK economic regulators, in part due to its simplicity.

The alternative and more widely used method in the academic literature (and increasingly by economic regulators) is stochastic frontier analysis (SFA); see the equation below. The stochastic cost frontier model can be represented as:

$$C_{it} = f(Y_{it}, P_{it}, N_{it}, \tau_t; \beta) + v_{it} + u_{it}$$

where the first term ($f(Y_{it}, P_{it}, Q_{it}, \tau_t; \beta)$) is the deterministic component, and Y_{it} is a vector of output measures, P_{it} is a vector of input prices, N_{it} is a vector of exogenous network characteristic variables, τ_t is a vector of time variables which represent technical change and β is a vector of parameters to be estimated. C_{it} represents the cost variable to be explained. The i and t subscripts refer to the number of firms and time periods respectively. Whilst some applications may use only cross sectional data, most transport applications utilise panel data, and this type of data greatly expands the possibilities for increasing the richness of the analysis in a number of ways as discussed further below. The v_{it} term is a random component representing unobservable factors that affect the firm's operating environment. This term is distributed symmetrically around zero (more specifically assumed to be normally distributed with zero mean and constant variance). A further one sided random component is then added to capture inefficiency (u_{it}).

For cross-sectional data, it is necessary to make distributional assumptions concerning the one-side inefficiency term, and the estimation proceeds via maximum likelihood. This is a significant limitation as these assumptions may not be valid. For panel data, there are additional estimation possibilities. Before turning to the panel data approaches it is worth summarising the benefits of the econometric methods for studying the structure of costs and relative efficiency performance.

Compared to cost function (or average response function estimation) it is clear that frontier methods are a significant development since they explicitly allow for the possibility of variation in efficiency performance between firms and over time. Compared with the DEA approach, econometric methods provide estimates of the underlying structure of production / costs, for example, the elasticity of costs with respect to different cost drivers, such as traffic volumes - which DEA does not. As noted, the study of these elasticities allows us to say something about the scale and density characteristics of the industry. Whilst DEA can be adapted to allow for non-constant returns to scale, it does not specifically produce estimates of the extent of returns to scale or its variation with firm size (as with a translog); further it is more difficult for DEA to deal with economies of scale versus density (though this can be addressed to some extent through using a two-stage approach).

In addition, through the development of stochastic frontier analysis, econometric techniques are also able to distinguish between random noise and underlying inefficiency effects. However, econometric approaches do require the choice of an appropriate functional form, and the more flexible forms (such as the translog) are not always straightforward to implement due to the large number of parameters to be estimated. In addition, the choice of distribution for the inefficiency term in stochastic frontier analysis is arbitrary. The precise method that researchers should use will therefore depend on a range of factors, and in many academic papers more than one method is used in order to provide a cross-check against the other approaches.

Stochastic frontier methods: panel data approaches

The existence of panel data offers a number of important benefits. First of all, by combining cross-sectional and time series observations it provides additional degrees of freedom for estimation. This may be very important, particularly if the number of companies for which data exists is small as it often is for economic regulators. Second, it provides an opportunity to simultaneously investigate inter-firm efficiency disparities, changes in firm efficiency performance over time, as well as industry-wide technological change over the period of the study. Third it can, for some models, permit the estimation of firm efficiency without recourse to potentially restrictive distributional assumptions. Finally, it offers the prospect of disentangling inefficiency from unobserved factors. This latter benefit may be particularly important for transport applications, where substantial differences exist between transport infrastructure and operations both within and between countries, but where it is hard to capture these differences in a set of variables to be included in the model.

One way of dealing with a panel is to treat each data point as a separate firm. In this case, each observation, including observations for the same firm over multiple time periods, is given a separate efficiency score. In the case of econometric estimation this assumption may not be appropriate, since it assumes that inefficiency is independently distributed across observations, even though it might be expected that an inefficient firm in one period is likely to retain at least some of that inefficiency in the next period. It is however a widely used method by economic regulators in the UK owing to its simplicity and thus transparency in terms of how the regulator communicates with regulated firms.

The alternative and more usual approach is explicitly to recognise the panel nature of the data set. Within this alternative, there are two further options. Firstly, to estimate the model using traditional panel data methods (fixed effects or random effects (GLS)); see Schmidt and Sickles (1984). Alternatively, Pitt and Lee (1981) offer a maximum likelihood version of the same approach. In both cases, inefficiency is assumed to be "time-invariant" and each firm is given one efficiency score for the whole period, rather than one score per firm for each period as in the simple pooled approach. The advantage of the traditional panel approach (fixed and random effects) is that it does not require distributional assumptions concerning the inefficiency term as in the maximum likelihood equivalent. This benefit does come at a cost though, as it requires the assumption that inefficiency does not vary over time, which is restrictive.

For long time periods, the assumption of time invariant inefficiency is clearly problematic, and a number of approaches which allow for inefficiency to vary, whilst retaining some structure to the variation, have been developed. Time varying models have been developed for both the traditional panel data methods (e.g. Cornwell, Schmidt and Sickles, 1990, and the maximum likelihood approach (e.g. Battese and Coelli, 1992; Cuesta, 2000. Kumbhakar and Lovell (2000) describe these approaches in detail. A key distinction in the literature is between those models which make the assumption of independence in inefficiency over time (e.g. pooled SFA; Battese and Coelli, 1995) and those which permit firm inefficiency to change in a structured and not random way over time (Cuesta, 2000). The latter seem to have advantages from a regulatory and economic perspective.

An important and relatively recent development in the literature has revolved around the problem of disentangling inefficiency from unobserved heterogeneity. In the standard panel literature, fixed and random effects are assumed to represent unobserved, time invariant factors that vary between firms. As noted, in the efficiency literature, these models have been applied as efficiency estimation approaches, with the firm effects re-interpreted as inefficiency. This approach risks badging unobserved factors – genuine heterogeneity between railways – as inefficiency. Methods have therefore been developed in the literature to address this (Greene, 2005; Farsi *et al.*, 2005; Kumbhakar *et al.*, 2014; Colombi *et al.*, 2014). One version of Greene's approach includes a firm-specific dummy, to capture unobserved heterogeneity between firms, which is assumed to be time invariant (e.g. environmental factors, such as topography or climate) as well as the one-side inefficiency term (which varies over time). The decomposition therefore relies on the assumption that inefficiency varies randomly over time whereas unobserved heterogeneity is time invariant (as well as on the distributional assumptions of the model). The model is then estimated via maximum likelihood. This is one of the so-called "true" models, and there is also a random effects version of this approach. The Farsi *et al.* (2005) approach separates inefficiency from unobserved heterogeneity by making the assumption that the former is assumed not to be correlated with the regressors whilst the latter may be (inefficiency being a function of the ability of management to control costs given the exogenous set of output requirements and input prices that it faces – hence this would not be expected to be correlated with the regressors). Finally, the approaches set out by Kumbhakar *et al.*, 2014; Colombi *et al.*, 2014 seek to go further and separate the model residual into four components: random noise, time varying inefficiency, time invariant inefficiency and time invariant unobserved heterogeneity. This model relies entirely on distributional assumptions to make this separation, which is a limitation. It further assumes that unobserved heterogeneity is uncorrelated with the regressors, which may not be valid. It is worth noting that these are relatively new approaches with relatively few applications in transport.

A note on production and distance functions

As noted above we focus on cost function estimation in this section, since ultimately it is cost that regulators care about. Thus technical and allocative efficiency are both important and analysis of the cost function permits this. Production functions – or their multi-input, multi-output counterparts may be used in circumstances where cost data or input prices are not available. These can provide useful insights, though they do not permit the analysis of allocative efficiency. That said, in railways and network industries it may not be possible for firms to adjust all inputs optimally, so technical efficiency analysis may be pertinent. Kennedy and Smith (2004) use a distance function approach that also aims to include bad outputs (see section 2.3 above). This paper is important from a regulatory perspective as it comprises an internal (within-country) comparison on rail infrastructure regions within Great Britain – such an approach is often the most pragmatic in transport applications, where there is only one infrastructure provider within a given country, and international comparisons are hard to achieve.

3.6. Concluding remarks

There exists a range of methods used in the academic literature to assess efficiency and productivity performance both between firms and over time. These include productivity indices, the most simple of which is labour productivity; with more advanced indices covering all inputs (total factor productivity). Non-parametric methods such as data envelopment analysis (DEA) permit the analysis of efficiency as well as productivity analysis, and parametric methods likewise permit a rich analysis.

Perhaps a key advantage of the parametric method is that it yields information on the relationship between costs and the regressors (which can be checked and challenged), which data envelopment analysis does not. Whilst parametric methods do require the choice of a functional form – which can be challenging – DEA is not necessarily an anti-dote to that problem, because it can be seen as a “black box” that “hides” the underlying shape of the frontier. In regulatory applications, our experience is that parametric methods are more used than DEA. In any case, where there are many cost drivers, DEA often requires a parametric second stage, thus reducing the difference between the methods to some extent.

Particularly important aspects of any benchmarking method are that it controls for heterogeneity between firms and over time, and that it deals with the possibility of economies of scale and density that are so prevalent in transport applications particularly (and network industries more widely). It needs to be credible; and potentially it also needs to be transparent, which may mean that simplicity could be preferred by regulators in some cases.

There exists a wide range of methods for estimating cost inefficiency in the literature, some of them relatively simple and widely used, particularly by regulators, and others more complex. However, some of the more complex methods are now entering the economic regulation sphere. In the UK, for example, the British rail (and road) regulator, i.e. the Office of Rail and Road (ORR), has adopted a range of advanced methods, and others, such as OFWAT, have considered these approaches at least, though to date have fallen back on simpler methods, given the data and results obtained.

In the regulatory context in the UK, the Competition and Markets Authority (CMA) which, inter alia, acts as a referee in disputes between regulators and regulated firms has also tended to prefer simpler methods, partly from a transparency perspective. Panel methods offer much more scope for a rich analysis of the cost structure (economies of scale and density) and inefficiency and these are the most widely used in railways. The question of dealing with random noise and heterogeneity (observed and unobserved) remains a key issue for all regulators in railways and other sectors. Ultimately, the choice of technique will depend on a number of factors, including the number of data points, availability of cost driver data, model performance, economic theory and practical considerations. Usually, it is appropriate to run a range of approaches and compare the results and in some cases it will not be possible easily to choose between them. Economic regulators in that case tend to average the efficiency results across a range of models.

Chapter references

Agrell, P. J., Bogetoft, P., (2017), "Regulatory benchmarking: Models, analyses and applications", Data Envelopment Analysis Journal, Vol. 3 (1–2). Version Retrieved from: https://www.researchgate.net/publication/305490664_Regulatory_Benchmarking_Models_Analyses_and_Applications.

American Productivity Center APC, (1981), *Total Productivity Measurement*, APC, Houston, TX.

Arrow, K.J., Chenery, H.B., Minhas, B.S. and Solow, R.M., (1961), "Capital-Labor Substitution and Economic Efficiency", *Review of Economics and Statistics*, August, pp. 225—250.

Banker, R.D., Chang, H.H., Majumdar, S.K., (1993), "Analyzing the underlying dimensions of firm profitability", *Managerial and Decision Economics*, 14, 25–36.

Banker, R.D., Chang, H.H., Majumdar, S.K., (1996a) "A framework for analyzing changes in strategic performance", *Strategic Management Journal,* 17, 693–712.

Banker, R.D., Chang, H.H., Majumdar, S.K., (1996b), "Profitability, productivity and price recovery patterns in the US telecommunications industry", *Review of Industrial Organisation*, 11 (1), 1–17.

Banker, R.D., Charnes, A., Cooper, W.W., (1984), "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis", *Management Science*, 30(9), 1078-1092.

Battese, G.E. and Coelli, T.J. (1992), "Frontier Production Functions and the Efficiencies of Indian Farms Using Panel Data from ICRISAT's Village Level Studies", *Journal of Quantitative Economics*, vol. 5, pp. 327-348.

Battese, G. E. and Coelli, T. J., (1995), "A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data", *Empirical Economics*, 20, 325-332.

Beattie, B.R. and Taylor, C.R. (1985), *The Economics of Production*, New York, Wiley.

Bitzan, J., (2003), "Railroad costs and competition", *Journal of Transport Economics and Policy*, 37 (2) 201-225.

Blades, D. (1998), *Measuring depreciation,* OECD.

Bogetoft, P., Otto, L. (2010), *Benchmarking with Dea, Sfa, and R,* Springer Science & Business Media, Vol. 157.

Caves, D.W., Christensen L.R., Diewert W.E., (1982), "The Economic Theory of Index Numbers and the Measurement of Input, Output and Productivity", *Econometrica*, Vol. 50, 6, 1393- 1414.

Charnes, A., Cooper W.W., Rhodes E., (1978), "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, 2(6), 429-444.

Christensen, L., Jorgenson, D., Lau, L., (1971), "Conjugate Duality and the Transcendental Logarithmic Production Function", *Econometrica*, 39, pp 255-256.

Coelli, T., Prasada Rao D.S., O'Donnell C.J., Battese G.E., (2005), *An Introduction to Efficiency and Productivity Analysis*, Springer, New York.

Colombi, R., Kumbhakar, S.C., Martini, G. and Vittadini, G., (2014), "Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency", *Journal of Productivity Analysis*, 42, pp. 123-136.

Cook, W.D., Zhu, J., (2013), "Data Envelopment Analysis: Balanced Benchmarking", *Create Space Independent Publishing Platform*.

Cook, W.D., Zhu, J., (2014), *Data Envelopment Analysis: A Handbook on Modelling of Internal Structures and Networks*, Springer Science.

Cornwell, C., Schmidt, P., Sickles, R.C. (1990), "Production Frontiers With Cross-Sectional And Time-Series Variation in Efficiency Levels", *Journal of Econometrics*, vol. 46, pp 185-200.

Cuesta, R. A. (2000), "A production model with firm-specific temporal variation in technical inefficiency: with Application to Spanish Dairy Farms", *Journal of Productivity Analysis*, 13 (2), 139-152.

Debreu G., (1951), "The Coefficient of Resource Utilization", *Econometrica*, 19(3), 273-292.

Diewert, W.E. (1981), "The Theory of Total Factor Productivity Measurement in Regulated Industries", in Cowing, T.G. and Stevenson, R.E. (eds.), *Productivity Measurement and Regulated Industries*, Academic Press Inc., New York.

Diewert, E. W., Lawrence, D., (1999), "Measuring New Zealand's Productivity", *Treasury Working Paper* 99/5, http://www.treasury.govt.nz/workingpapers/99-5.htm.

Diewert, E.W., (1976), "Exact and Superlative Index Numbers", *Journal of Econometrics*, 115-45.

Diewert, W. (1971), "An Application of Shephard Duality Theorem: A Generalised Leontief Production Function", *Journal of Political Economy*, 79, 481–507.

Efron, B., (1979), "Bootstrap Methods: Another Look at the Jackknife", *The Annals of Statistics*, vol. 7(1), 1-26.

Erbetta, F., Fraquelli G., (2012), "Managerial efficiency and size: an assessment of strategic potential of Italian manufacturing firms", *European Review of Industrial Economics and Policy ERIEP*, 4.

Farrell, M.J., (1957), "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society*, 120(3), 253-281.

Farsi, M., Filippini, M., Kuenzle, M., (2005), "Unobserved heterogeneity in stochastic cost frontier models: an application to Swiss nursing homes", *Applied Economics*, 37(18): 2127-2141.

Fraquelli, G., Vannoni D., (1996), "Regulation and Total Productivity in Electricity: A Comparison between Italy, Germany and France", *Vierteljahrshefte zur Wirtschaftsforschung*, Berlin, Vol. 65, Iss. 4, 512-524.

Fraquelli, G., Vannoni D., (2000), "Multidimensional performance in Ttelecommunications, regulation and competition: analysing the European major players", *Information Economics and Policy,* 12, 27–46.

Greene, W., (2005), "Reconsidering heterogeneity in panel data estimators of the stochastic frontier model", *Journal of Econometrics*, vol. 126, pp. 269-303.

Kennedy, J., Smith, A.S.J., (2004), "Assessing the Efficient Cost of Sustaining Britain's Rail Network: Perspectives Based on Zonal Comparisons", *Journal of Transport Economics and Policy*, 38, pp.157-190.

Kumbhakar, S.C., Lovell, C.A.K., (2000), *Stochastic Frontier Analysis*, Cambridge University Press, Cambridge UK.

Kumbhakar, S.C., Lien, G., Hardaker, J.B., (2014), "Technical efficiency in competing panel data models: a study of Norwegian grain farming", *Journal of Productivity Analysis*, 41, 321-37.

Lintner, J., (1965), "The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets", *The review of economics and statistics*, 13-37.

Meinen G., Verbiest P., de Wolf P.P., (1998), *Perpetual Inventory Method. Service lives Discard patterns and Depreciation methods*, OECD.

Martini, G., Manello, A., Scotti, D., (2013) "The influence of fleet mix, ownership and LCCs on airports' technical/environmental efficiency", *Transportation Research Part E: Logistics and Transportation Review*, *50*, 37-52.

Mossin, J., (1966), "Equilibrium in a capital asset market", *Econometrica*, 768-783.

OECD, (2001), *Measuring Productivity, Measurement of aggregate and industry-level productivity growth*, OECD, Paris.

Pitt, M.M., Lee, L-F., (1981), "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry", *Journal of Development Economics*, vol. 9, pp 43-64.

Schmidt, P., Sickles, R.C., (1984), "Production Frontiers and Panel Data", *Journal of Business & Economic Statistics*, vol. 2 (4), pp 367-374.

Simar L., Wilson WP., (1998), "Sensitivity analysis of efficiency scores: how to bootstrap in nonparametric frontier models", *Management Science* 44(1) 49-61.

Simar, L., Wilson, PW., (2000a), "Statistical inference in nonparametric frontier models: the state of the art", *Journal of Productivity Analysis*, 13 49-78.

Simar, L., Wilson PW., (2000b), "A general methodology for bootstrapping in nonparametric frontier models", *Journal of Applied Statistics* 27(6) 779-802.

Shephard, R.W., (1953), *Cost and Production Functions,* Princeton: Princeton University Press.

Sharpe, W.F., (1964), "Capital asset prices: A theory of market equilibrium under conditions of risk", *The journal of finance*, 19(3), 425-442.

Smith, A.S.J., Wheat, P. Smith, G., (2010), "The role of international benchmarking in developing rail infrastructure efficiency estimates", *Utilities Policy*, 18, pp.86-93.

Smith, A.S.J., Nash, C.A., (2014), "Rail Efficiency: Cost Research and its Implications for Policy", *International Transport Forum Discussion Paper*, 2014: 22, OECD.

Spady, R.H., Friedlaender, A.F., (1978), "Hedonic cost functions for the regulated trucking industry", *The Bell Journal of Economics*, 9 (1), 159–179.

Syverson, C., (2011), "What Determines Productivity?", *Journal of Economic Literature*, 49(2), 326–365

Solow, R., (1957), "Technical Change and the Aggregate Production Function", *Review of Economics and Statistics*, Vol. 39, 312-320.

Star, S., Hall, R.E., (1976), "An approximate Divisia index of total factor productivity", *Econometrica,* Vol.44 (2), 257-263.

Wheat, P.E., Smith, A.S.J., (2014), "Do the usual results of railway economies of scale and density hold in the case of heterogeneity in outputs: A hedonic cost function approach", *Journal of Transport Economics and Policy* (online first, January 2014).

Zellner, A., (1962), "An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias", *Journal of American Statistical Association*, 57, pp. 348-368.

4. BENCHMARKING HIGHWAYS

Chris Nash

4.1. Introduction

Highways in Europe are for the most part provided by public authorities (although both construction and maintenance may be contracted out) and free at the point of use. A small number of cities now have road pricing schemes, and several countries have toll roads, bridges or tunnels. In some cases these are provided as concessions, let by a process of competitive tendering. Despite not in most cases having road tolls, Britain has a PPP scheme under which certain roads are essentially contracted out long term, including construction or upgrading and maintenance, following a competitive tendering process (Initially, the British approach remunerated the private contractor by a system of shadow tolls, but it was realised that it was not sensible to give an incentive to generate more car traffic, so the shadow tolls for cars have been replaced by availability payments, but with incentive payments for reliability of journey times and safety). In principle, competitive tendering should give a strong incentive for efficiency provided that there is a genuine transfer of cost risk to the contractor. Even where roads are concessioned, however, renegotiation or the extension of concession periods without competitive tendering is common. In short, roads are an area where competition is relatively weak, and where the use of benchmarking and yardstick competition may therefore be seen as important.

There is a long history of benchmarking highways, but most of this literature simply compares key performance indicators with no formal way of allowing for differences in environment, highway quality, user satisfaction or input prices (Wheat, 2017). There is a second strain of analysis of highway costs, but from the point of view of assessing marginal costs for use in pricing decisions with no allowance for inefficiency. Much of this work formed part of EC projects in infrastructure costs and charges. Since it usually utilises data on short sections of road which do not form decision making units, it is not able to shed light on issues of economies of scale in road maintenance and renewal. It does however suggest strong economies of density, with elasticities for renewal cost between 0.5 and 0.8 and for maintenance cost between 0.4 and 0.7, while the elasticity for operations cost appears to be more or less zero (Link et al, 2014). Thus, comparing costs per unit of traffic between road systems without allowing for traffic density would be very misleading, but since the elasticities for maintenance and renewals are far from zero, comparing systems on the basis of costs per road or lane kilometre without allowing for the costs related to traffic levels would also be misleading.

Neither of these strands of the literature will be further reviewed here; instead, we will concentrate on benchmarking studies using formal statistical methods, of which there appear to be relatively few.

4.2. Benchmarking roads using formal statistical techniques - discussion

Use of formal benchmarking methods requires a sample of different decision making units whose efficiency may be compared. Where major highways are run by a single national body, with no allocation of responsibility to lower levels of disaggregation, then such studies can only be done with a cross section of countries, involving all the problems of international comparability such studies encounter. Where there are separate bodies responsible for highways for different regions, or where there is significant delegation of authority to regions within a single national body, such studies may use regional data. In most countries local roads are the responsibility of states or local authorities so benchmarking studies may use within country data. Finally, as noted above, some countries award concessions for construction and maintenance of roads within an area, usually toll roads. This again permits within- country benchmarking.

In England, the Office of Rail and Road (ORR) is responsible for assessing the efficiency of the national provider of main roads (Highways England) and has commissioned a review of experience and possibilities for road benchmarking (KPMG, 2016). They conclude that the scope and detail of the publicly available information is high relative to that found in most other regulated sectors. Information can be sourced from international datasets including those produced by the OECD, World Bank, International Road Federation, European Road Federation and World Economic Forum, but there is uncertainty about comparability.

Several international benchmarking studies have been undertaken for roads, but most, such as Directors of Roads in Europe (2010), do not use formal analytical methods. An exception is Braconier et al, (2013), which uses DEA analysis, but considers inputs and outputs of the road system as a whole, including number of vehicles, accidents and carbon emissions, and does not consider infrastructure cost, noting that this is a small part of the cost of the road system as a whole. In other words, this is a sector in which quality of service is very important given its impact on vehicle operating costs, which constitute the majority of the costs of the system.

ORR is now developing methods for benchmarking costs of Highways England regions using a stochastic cost frontier approach; an initial presentation on this work was given at the European Transport Conference in 2017 (Spencer-Bickle, 2017).

There is debate in the literature on how the outputs of highways agencies should be measured (e.g. Massiani and Ragazzi, 2008). It is generally agreed that costs depend both on the capacity provided (usually measured as lane kilometres, although the costs of 100km of 2 lane roads may differ from those of 50 km of 4 lane roads) and traffic volumes. Vehicle km may be relevant for toll collection costs, but in terms of damage done to road pavements, standard axle km would be a more appropriate measure. However, this data is rarely available; the best that can be done usually is to consider whether the proportion of traffic made up by heavy goods vehicles (the main determinant of standard axle km) varies across the sample. Most studies examine operating, maintenance and renewals cost rather than capital costs. Costs may also vary significantly with the characteristics of the network, such as the presence of bridges and tunnels, and the charging mechanism used. Quality of service, including the smoothness of the road surface and waiting times at manual toll booths, may also be relevant variables although ones on which data is scarce.

4.3. Benchmarking roads using formal statistical techniques - results

In practice, the main source of econometric studies of road infrastructure costs is within country comparisons. The earliest study we have found is Sikow and Talvitie (1996). This used a translog cost function to examine economies of scale and scope for road construction, rehabilitation and maintenance for 20 districts and 100 construction projects in Finland. Outputs consisted of kilometres of each of three types of roads, in terms of traffic density, and other explanatory variables included traffic volumes, numbers of bridges and interchanges and kilometres of cycleway and walkway. This was not really a benchmarking study, but did produce important conclusions for the efficient design of contracts in the industry. It found economies of scale in both construction and maintenance at existing organisation size in Finland; the most efficient size of maintenance district was found to cover 800-1000km. There were no economies of scope, but surprisingly contracting out was found to raise costs. It was speculated that this was because in the short run in house organisations were unable to reduce their staff in line with reductions in workload.

More recent studies have used data envelopment analysis. For instance, Ozbek et al (2012) use a variable returns to scale DEA model to examine the relative efficiency of road maintenance in 8 counties in Virginia, USA. Their input is expenditure and outputs road condition and area of road maintained (i.e. the sum of length times width of the roads in the system). They start with a long list of uncontrollable variables representing traffic levels, climate, terrain, subsurface conditions and type of road, but use earlier research to group regional factors into a categorisation of low, medium and severe in terms of their impact on maintenance costs. The variation in efficiency scores is very large, indeed implausibly so. The same authors have applied a similar model to examine the efficiency of bridge maintenance in Virginia (Ozbek et al, 2010). A similar study with some of the same authors (Fallah-Fini et al, 2012) examined the efficiency of maintenance of the Interstate Highway system in Virginia. A bootstrapping method was used to correct for bias, and this reduced the variability of efficiency scores. The main interest in the results is that Virginia at the time used a mix of traditional contracts which specified the work to be done and performance based contracts which specified the target road condition sought but left the method to the contractor. Surprisingly the traditional contracts were found to be the more efficient approach, although the authors suggest that this may be partly because performance based contracts were still in their infancy, and a hybrid of the two might prove to be best.

Wheat (2017) estimates a stochastic cost frontier model using maintenance costs for 51 British local authorities for 5 years with road length, traffic levels, mix of road types, quality (no of defects) and user satisfaction as explanatory variables. He finds there are strong economies of scale for small authorities, but that these are exhausted by road network sizes of 4-8k km after which diseconomies of scale set in. He finds that poor road quality raises costs. At low and high levels of public satisfaction, there is a strong positive relationship between satisfaction and cost, but this relationship is weak in the middle range.

A number of econometric benchmarking studies relate to toll road concessions in Italy and Norway. Of the Italian toll concession studies, Massiani and Ragazzi (2008) use a cross section of 18 concessions for the year 2006. They estimate a stochastic cost frontier model in which operating costs are regressed on output measured as weighted road km (where the weights are the number of lanes) and vehicle km (it is argued that the proportion of heavy goods vehicles varies little in the sample). Since this is a simple cross section within a single country, input prices are assumed to be constant across the sample. Results using a translog cost function are found to be very sensitive to the exact specification of the error term of the model; therefore the authors prefer to use a simple linear model. For this reason no clear results on economies of scale are obtained. However, the authors do find evidence of considerable inefficiency and suggest that yardstick competition could help drive these out.

In a second study, Benfratello, Iozzi and Valbonesi (2009) estimate a translog cost model for 20 Italian motorways concessionaires over the 1992–2004. This is not a frontier model as the aim was to examine economies of scale and density and technical change rather than efficiency. So strictly it is outside the scope of this review, but is included because of the interest of its results for Italian highway concessions. They use length of the road network and volume of traffic as output measures and input prices for labour, maintenance cost per vehicle km and other cost per road km. They also use the number of viaducts or tunnels and the percentage of the system that is three lane to control for other characteristics of the road network and a time trend for technical progress.

They find considerable technical progress, strong economies of density (with an elasticity of around 0.5) and economies of scale at least up to networks of 300km. Beyond this length, it is estimated that economies of scale are exhausted by 600km, but the authors warn that the estimates of economies of scale are dependent on inclusion of a single large operator, so no clear conclusion can be reached. Privatisation reduces costs by 3%, whilst introduction of price cap rather than cost plus regulation appears to have had no discernible effect. In further work reported in the ART 2017 Annual Report (ART, 2017) this work has been further developed but using cost frontier methods and obtained similar results.

It is noticeable that for what it is worth this suggests that economies of scale are exhausted at a very much lower level of road kilometres than found by Wheat for Britain. Of course, his results are for local roads, where spending levels are much lower, so it may be expected that economies of scale persist at much higher levels of road kilometres than for inter urban roads. Nevertheless, the difference is pronounced.

The Norwegian studies are of less interest, as in Norway highway construction and maintenance is the responsibility of the Norwegian Public Roads Administration, and thus studies of toll road company costs only include the cost of tolling and administrative costs, and not road maintenance and renewals. Of the Norwegian studies, Amdal et al (2007) estimate a translog cost function for a panel of 26 concessions for the years 1998–2004. They only use traffic, and not network length, as an output variable but introduce as other variables the number of lanes in the toll station(s), total debt (in billion NKr) by the end of each year, share of vehicles using on board units, a dummy variable for toll cordon, a dummy variable for passenger charging and a dummy variable for competitive tendering, as well as a full set of time dummies included to control for any effects of aggregate factors common to all companies. The reason for including financial debt as a variable is that toll companies report that a considerable part of their operating cost is related to financial management, whilst number of lanes, use of on board units, presence of a cordon and passenger charging all affect the cost of toll collection. Economies of scale are found up to a traffic volume of approximately 190 million paying vehicles per year, but as network length is not included nothing can be deduced about the appropriate size of a concession. Competitive tendering is found to reduce costs by 24%.

Welde and Odeck (2011) apply both DEA and stochastic frontier analysis to a sample of 20 companies over the period 2003-2008. As output they use the number of vehicles handled per lane, and as inputs operational and administrative costs. This is a curious formulation and it is not clear to the current reviewer why the output is expressed per vehicle lane but apparently the costs are not. They also use several exogenous variables related to the form the tolls take. including the extent to which the system for collecting tolls is based on electronic toll collection, whether there is a charge per passenger as well as the charge per vehicle and whether the company being assessed is a cordon toll company. They find considerable variation in efficiency, although this varies between the DEA and cost frontier methods, but no evidence of economies of scale.

4.4. Concluding remarks

In conclusion, it seems that although there are many sources of data on roads, data availability is still a significant problem. Ideally roads cost benchmarking would use both measures of traffic, including vehicle and standard axle kilometres, and of capacity, such as lane km, as outputs and also network characteristics, form of tolling and quality of service as control variables. Often these are not available and proxies have to be used. There is strong evidence of economies of density in roads, but the position regarding economies of scale is less clear. For local roads, Wheat finds economies of scale persist up to road lengths of 4-8k km, whilst according to the limited evidence available, economies of scale for Italian toll roads are exhausted at a much lower network size. However, these studies are limited by the fact that the evidence is heavily dependent on a single larger company. Finally all studies have found considerable levels of inefficiency, suggesting a role for benchmarking and yardstick competition where other forms of competition are not feasible, as well as competitive tendering for any new concessions.

Chapter references

Amdal, E., Bardsen, G., Johansen, K., Welde, M., (2007), "Operating costs in Norwegian toll companies: a panel data analysis of 26 concessions for the years 1998–2004", *Transportation,* 34.

Autorità di Regolazione dei Trasporti, (2017), *Fourth Annual Report to the Parliament Chamber of Deputies*, ART, Turin.

Benfratello, L., Iozzi, A., Valbonesi, P., (2009), "Technology and incentive regulation in the Italian motorways industry", *Journal of Regulatory Economics*, 35, 201–221.

Braconier, H., Pisu, M., Bloch, D., (2013), "The Performance of Road Transport Infrastructure and its Links to Policies", *OECD Economics Department Working Papers*, No. 1016, OECD Publishing, Paris.

Conference of European Directors of Roads (CEDR) (2010) BEXPRAC, Brussels.

Fallah-Fini, S., Triantis, K., Jesus, M., Seaver, W.L., (2012), "Measuring the efficiency of highway maintenance contracting strategies: A bootstrapped non-parametric meta-frontier approach", *European Journal of Operational Research*, 219(1), 134-145.

KPMG, (2016), *Benchmarking Highways England: A report to the office of rail and road*, KPMG, London.

Link H., Nash, C.A., Ricci, A., Shires, J.D., (2014), "A generalised approach for measuring the marginal social costs of road transport in Europe", *International Journal of Sustainable Transport*.

Massiani, J., Ragazzi, G., (2008), "Costs and efficiency of highway concessionaires: a survey of Italian operators", *European Transport,* 38, 85-106.

Ozbek, M.E., de la Garza, J.M., Triantis, K., (2010b), "Efficiency measurement of bridge maintenance using data envelopment analysis", *ASCE Journal of Infrastructure Systems*, 16(1), 31–9.

Ozbek, M.E., de la Garza, J.M., Triantis, K., (2012), "Efficiency measurement of the maintenance of paved lanes using data envelopment analysis", *Construction Management and Economics*, 30, 995–1009.

Spencer-Bickle, A., (2017), "Initial benchmarking of Highways England's regional maintenance spending", *European Transport Conference*, Barcelona.

Welde, M., Odeck, J., (2011), *Utilities Policy*, Volume 19, Issue 3, 162-171.

Wheat, P., (2017), "Scale, quality and efficiency in road maintenance: Evidence for English local authorities",*Transport Policy*, 59, 46–53.

5. RAIL BENCHMARKING

Andrew Smith and Chris Nash

5.1. Introduction

Traditionally European rail systems have been structured as government owned vertically integrated monopolists. In that situation, one would have expected governments to be very interested in benchmarking as a way of determining the efficiency of their own rail systems. Some benchmarking studies did take place (for instance BR/University of Leeds, 1979, although this made no use of formal statistical techniques) but perhaps it is surprising governments did not make more use of this technique, although of course most of the methodology in use today had not then been developed.

Starting with Directive 91/440, the European Commission encouraged the division of rail companies into what it saw as the natural monopoly element (the infrastructure) and the potentially competitive element (train operations). Subsequent legislation has required provision of an independent regulator and of incentives for efficiency, either as part of the regulatory process or as part of a multiannual contract between the infrastructure manager and the ministry. Track access charges may be based on the costs of an efficient infrastructure manager. That left an obvious need for benchmarking of the costs of infrastructure managers, although again surprisingly little use has been made to date of benchmarking techniques. The British regulator has however used these in its determinations as described below.

For train operations, new entry has been permitted into all freight and international passenger sectors. In the coming years new entry will be permitted into commercial passenger operations and as competition for public service contracts; this has already happened in some countries. Thus it may reasonably be argued that actual competition in or for the market may render yardstick competition redundant in train operations. However, the stage when competition plays this role has not yet been reached. Even in the freight market, the incumbent still has more than half of the market, whilst in the passenger sector often there is little or no competition in or for the market. Thus we would see benchmarking as having an important role in train operations as well, at least for some years to come.

Experience of benchmarking in the two sectors will be considered in turn, followed by an examination of the evidence relating to rail systems as a whole.

5.2. Benchmarking infrastructure managers

As noted above, the only case where we are aware of a regulator making formal use of benchmarking techniques is in Britain.

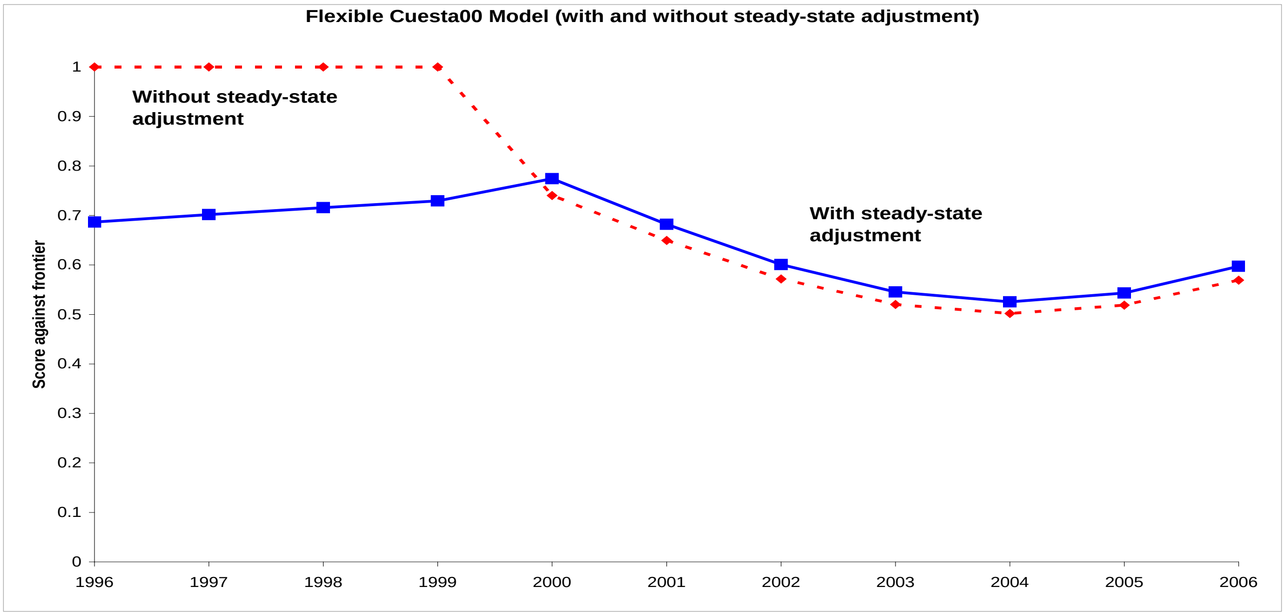
ORR carried out a review of Network Rail's efficiency performance, in the light of large cost increases, and commissioned a wide range of studies for its 2003 Interim Review of the company's finances. A key weakness of the 2003 review though was that the ORR's efficiency determination was ultimately based on two bottom-up consultant reviews of Network Rail's business plan (LEK, TTCI and Halcrow, 2003 and Accenture, 2003). These results were supplemented by internal benchmarking, which indicated the kind of savings that could be achieved if Network Rail implemented its own best practice consistently across the network.

Ultimately then, the 2003 Interim Review was unable to provide a clear, empirically based assessment of Network Rail's relative efficiency position based on hard data from external sources. ORR nevertheless set a tough efficiency target of 31% over 5 years (2004-2009). However, costs were starting from a very high base. Thus, although costs then started to fall as Network Rail set about delivering its efficiency targets, by the time of the next periodic review in 2008 the scene was set to take the benchmarking approach a step forward by attempting international comparisons.

Two approaches were adopted during the 2008 review. The first used a panel of thirteen European infrastructure managers over an 11 year period. The data was provided from the Lasting Infrastructure Cost Benchmarking (LICB) project undertaken by UIC. The dataset included data on costs (adjusted based on PPP exchange rates), traffic volumes (by type), measured in both train-km and tonne-km, network length, and a range of other variables characterising differences between the companies (for example, extent of electrification, network density). Ideally, both train km (as determinants of capacity requirements) and gross tonne km (as determinants of maintenance and renewals costs) would be included.

A structured inefficiency model (Cuesta, 2000) was used that permits inefficiency to vary by firm over time, but in a structured way that recognises the panel nature of the dataset. The results are shown for Network Rail in Figure 1 (other companies cannot be shown for confidentiality reasons); see Smith (2012). Results are shown for maintenance only, and for maintenance and renewals, with the additional model variant to allow for Network Rail's renewals costs to be reduced downwards prior to modelling to allow for the fact that the company was renewing at above steady state levels in terms of renewal volumes. The overall message of Figure 1 is that Network Rail's efficiency deteriorated sharply after 2000, compared to its European comparators, leaving the company with an efficiency gap of around 40% by the end of the period. The analysis was carried out by the University of Leeds, with ORR and in conjunction with Network Rail and UIC.

Figure 1. Profile of Network Rail Efficiency Scores: Preferred Model



Source: Reproduced from Smith (2008); see also Smith (2012).

In a separate, supporting study, ORR and University of Leeds collected a new dataset comprising five other rail infrastructure managers in Europe and North America. This includes data on costs, outputs, and network characteristics at the regional level within each country. Thus, although the number of companies included was smaller than in the LICB dataset, the sample size was expanded via the use of regional data within companies (sub-company data structure). The dataset also allowed ORR to study within-country variations in inefficiency. The results broadly confirmed the results of the main study using LICB data (see Smith *et al.*, 2010; Smith and Wheat, 2012b).

It is further worth noting that the ORR carried out a range of other studies, principally based on bottom-up evidence. These confirmed the existence of a substantive gap, supported by examples of best practice in other countries (see Table 1).

Although ORR carried out / commissioned a wide range of studies – all of which pointed in the direction of a large efficiency gap – it was the output of the LICB-based econometric model which was used to set Network Rail's efficiency targets. ORR chose to compare Network Rail against the upper quartile of the peer group, rather than the frontier, thus meaning that the starting efficiency gap for its analysis – based on the preferred econometric model from the analysis of the LICB data - was 37% rather than 40%. ORR also gave the company ten years to close the gap, with only two thirds of the gap targeted to be closed during the immediate control period (control period 4 (CP4); 2009-2014).

In the next periodic review (PR13), ORR shifted the emphasis of its approach to bottom-up methods. This was driven by a number of factors, but in part reflected increased doubts after 2008 about the quality of the LICB data and the commitment of the different companies to providing accurate information. A re-run of the sub-company approach was also attempted, but again it was considered that there was insufficient time to get enough certainty about the quality and comparability of the data received. Therefore, although Network Rail acknowledged the size of the efficiency gap resulting from the PR08 econometric modelling, emphasis switched in the PR13 review to bottom-up analysis. Whilst new econometric modelling with an updated LICB dataset was carried out and reported, in the process also applying more advanced techniques, the econometric modelling played a supporting role to the bottom-up analysis (thus reversing the approach taken in PR08; see ORR, 2013).

Table 1. Examples of European best practice

Asset inspection and asset management	In general best practice European railways undertake fewer track inspections but inspections are generally of higher quality. It is estimated that similar techniques applied in Britain could reduce foot patrolling inspection costs by around 75% and tamping expenditure by 20%
Recycling components	This is common European practice. In Switzerland, for example, rail, point motors, sleepers and signal heads are regularly refurbished then cascaded from higher to lower category routes. Cascaded rail on lines re-laid with steel sleepers could lead to savings. Additionally ballast cleaning (partial renewal) as opposed to traxcavation (complete renewal) could reduce ballast renewal cost in Britain by 40%
High output rail stressing	Stressing continuously welded rail by heating it rather than physically stretching it is a process discontinued in Britain in the 1960s and 1970s. Some European networks (using modern equipment) have re-introduced this method which doubles on-site productivity and, if applied to the renewals re-railing workbank in CP4, could lead to significant annual savings for Network Rail
Formation rehabilitation trains	Modern high output European plant is regularly used to undertake formation and also ballast renewals. If applied to Network Rail's CP4 category 7 and 12 track renewals RailKonsult estimate that it could reduce unit costs for both activities by around 40%
Lightweight station platforms	The use of modular construction polystyrene station platforms in the Netherlands could provide opportunities in Britain, given the substantial CP4 platform extension workbank. Analysis suggests a unit cost saving of around 25% in Britain
Efficient European re-railing techniques.	This particular study brought together many themes from the previous RailKonsult work by focussing upon the Swiss re-railing method. Bespoke plant, high output welding techniques and dedicated teams are applied routinely. Put together for basic re-railing work alone this method is around 40% more efficient than current Network Rail practice
Use of dedicated teams	Contractors are widely used by most continental railways, as they are in Britain. However there is generally a greater degree of specialisation by activity in Europe (such as S&C renewal or tamping). This ensures a highly skilled and productive workforce dedicated to particular tasks in contrast to the situation in Britain where contractors are often not even dedicated to rail.

Source: Taken from Smith *et al.*, 2010.

Perhaps one of the lessons that may be learned here is that international benchmarking is problematic because it takes considerable time and commitment from a group of countries to make the analysis credible and usable. In PR08 ORR had the advantage of a ready-to-go dataset, produced by UIC, and this enabled top down, econometric international benchmarking to play a more significant role than it has in other regulated sectors. If international benchmarking is to work, then it may require concerted efforts by regulators / governments across Europe working together to establish a common benchmarking framework against which all companies can be compared, thus also implying that data can be requested and audited by regulators and policy makers.

Finally, a further opportunity for benchmarking remains the notion of internal benchmarking. Whilst not without its problems it remains a useful part of a regulators toolkit as it establishes the savings that could be achieved if best practice (within-country) is consistently applied. The existence of disaggregation into units that have managerial autonomy (at least to some degree), as with Network Rail's routes, is of course a pre-requisite for such an approach, but these groupings / disaggregations do also exist in other railways. It should be noted that ORR have developed this approach as part of the 2018 Periodic Review, benchmarking Network Rail's 10 routes and 37 maintenance delivery units against each other (using panel data); see ORR (2018). A relatively comparable and rich dataset has been assembled; though the analysis is still in its early stages and has only been used as supporting evidence by the regulator. An early implementation of this kind of approach can also be found in Kennedy and Smith (2004).

5.3. Train operations

Again the evidence on train operating costs comes mainly from Britain and from the passenger sector. The reason is that Britain has 20-25 passenger franchises (the number has changed a little over time) each of which is constituted as a separate company publishing its own accounts. In most countries such data for a number of individual train operators is not available.

Both DEA and cost function (including stochastic frontier) methods have been used. However, the purpose of the studies has typically been to examine issues such as franchising policy or economies of scale and density rather than benchmarking *per se*.

Whatever approach is taken to measurement of efficiency, there is an immediate need to determine what are the outputs and inputs of a train operator and that issue will be considered first, before we look at actual studies.

For outputs, there is typically a choice of considering passenger km, train km or vehicle km. Passenger km may be considered the ultimate output of passenger train operators, although their costs may vary sharply with the type and quality of service (a passenger km seated in a luxury coach of a long distance train is both more expensive and more valuable than a passenger km standing in a crowded commuter train). However, neither commercial nor social considerations would suggest that train operators minimise costs per passenger km by operating long infrequent and overcrowded trains. Frequency of service is an important quality of service characteristic and is often stipulated in public service contracts. This suggests train km may be a better output measure, although obviously costs vary with the characteristics of the train, including its speed and number of stops. Since both staff and rolling stock are paid for per unit of time rather than distance, costs per train km tend to be reduced by increases in speed, although at high speeds this may be partly offset by increased energy and maintenance costs. Obviously costs also increase with length of train, although not proportionately, as longer trains still only require one locomotive (for locomotive hauled stock) and one driver.

This suggests that passenger km, train km, vehicle km and average speeds may all be relevant measures, although only the study by Smith and Wheat (2014) was able to use all these measures or derivatives of them; usually not all are available. Ideally some measure of the peakedness of traffic would also be valuable, as it is peak demand that determines the size of the vehicle fleet required.

On the input side, providing a rail service requires locomotives, passenger coaches or freight wagons (or self-powered vehicles), track, signalling, terminals and a variety of types of staff (train crew, signalling, track and rolling stock maintenance, terminals and administration). While ultimately all may be regarded as forms of labour and capital, the length of life of the assets and government intervention over employment and investment will often mean that at a particular point in time an undertaking will not have an optimal configuration of assets and staff. This renders attempts to measure inputs simply as labour and capital difficult, as measures of the value of capital stock will need to allow for excess capacity and inappropriate investment. An alternative is to simply look at physical measures of assets (e.g. kilometres of track, numbers of locomotives, carriages and wagons for railways), but this obviously makes no allowance for the quality of the assets. Further, and importantly, physical measures such as staff numbers can be greatly affected by contracting out and, unless a balancing item such as other costs is included as an input, substantially distort measures of productivity and efficiency.

A key problem in measuring technical efficiency is that of joint costs and economies of scale and density. For instance, a single-track railway may carry both passenger and freight traffic, a passenger train first- and second-class passengers, and a freight train a variety of commodities. In this situation, only some of the costs can be specifically attributed to one of the forms of traffic; the remaining costs are joint. The result is that railways typically are characterized by economies of scope; i.e., the costs of a single railway handling a variety of types of traffic are less than if each distinct product were to be handled by a different railway. Moreover, most evidence suggests that railways are subject to economies of traffic density. Putting more traffic on the same route generally reduces unit costs and raises measures of total factor productivity, unless the route is already heavily congested.

The result is that apparent rises in productivity may be caused by diversification into new products or by increased traffic density rather than being relevant to the measurement of performance. Of course, under conditions of economies of density, running more services (and possibly different types of service) on the network does lead to a genuine improvement in productivity. The argument here, however, is that the improvement in productivity arises naturally as a result of the shape of the cost function, and not because of any improvement in working practices. Of course, to the extent that the method used contains relevant measures of outputs and output characteristics such that it can capture some of these features of the technology (e.g. scale and density effects; quality; network complexity), then it should be possible to obtain measures of technical efficiency after having taken account of these effects.

As noted above, a variety of methods have been in studies of British passenger train operating costs, including non-parametric DEA (Affuso *et al.*, 2003; Cowie, 2009, Merkert et al, 2009) and index number approaches (Cowie, 2002a; Smith *et al.*, 2009), as well as parametric estimation of cost functions (Cowie, 2002b; Smith and Wheat, 2012a, Wheat and Smith, 2014), production functions (Cowie, 2005) and distance functions (Affuso *et al.*, 2003). Clearly the former methods (DEA and index number approaches) can only consider cost or technical efficiency and produce no estimates regarding the actual cost structure.

An important issue is whether to include an infrastructure input in any analysis of train operating costs. Clearly the infrastructure input may be an important part of the transformation function and so should be considered for inclusion in any analysis. The four papers by Cowie all include some measure of infrastructure input in the analysis which is some combination of route-km and access charges paid by operators to the infrastructure manager (to form a price if applicable).

This in turn raises two important and related problems: (1) the infrastructure input is hard to measure (this is particularly the case in Britain where access charges change significantly from year to year depending on the degree of subsidy); and (2) the inclusion of this input turns the analysis into an assessment of rail industry costs/production, rather than being targeted on the TOCs. In their study, Affuso *et al.* (2003) produce two models: one including the infrastructure input, and one not. The results differ as a consequence, although this problem is less severe during the early period after privatisation (which the study covers), since access charges and infrastructure costs were fairly stable during that period. Whilst there are good reasons for capturing the infrastructure input in a study of TOC performance, to capture the possibility that this input affects the TOC transformation function, Smith and Wheat (2012a) and Wheat and Smith (2014) argue that, given the measurement problems noted above, infrastructure inputs are best left out of the analysis. The dependent variable in their paper is thus defined as TOC costs, excluding fixed access charges. Route-km is also included as an explanatory variable in their model, not as a measure of the infrastructure input, but to distinguish between scale and density effects.

The focus of Smith and Wheat (2012a) was on the impact on cost efficiency of contract regimes after several renegotiations and temporary contracts being introduced following franchise failure. It used a panel data stochastic frontier framework which allowed efficiency to evolve over time (based on Cuesta, 2000; see section 3). They also included dummy variables in the cost function to allow the extent of cost effects of different contract types to be directly estimated.

The focus of the Wheat and Smith (2014) work, in contrast, was how to best model the cost structure of the industry. This work utilised a hedonic cost function (see section 3.2) and the description of the data used is given in Table 2. In particular, they defined three generic outputs (Route-km, Train-hours and number of stations operated) and nine characteristics of train services which go into the Train-hours output function. These characteristics control for the heterogeneity in train service provision. They also defined two inputs and associated prices.

Table 2. Data used in Wheat and Smith (2014)

Symbol	Name	Description	Data Source
y_1	Route - km	Length of the line-km operated by the TOC. A measure of the geographical coverage of the TOC	National Rail Trends
y_2	Train Hours	Primary driver of train operating cost	National Modelling Framework Timetabling Module
q_{12}	Average vehicle length of trains	Vehicle-km / Train-km	Network Rail
q_{22}	Average speed	Train-km / Train Hours	National Modelling Framework Timetabling Module
q_{32}	Passenger Load Factor	Passenger-km / Train km	Passenger-km data from National Rail Trends. Train-km data from Network Rail.
q_{42}	Intercity TOC	Proportion of train services intercity in nature	National Rail Trends for the categorisation of TOCs into intercity, LSE and regional. Where TOCs have merged across sectors a proportion allocation is made on an approximate basis with
q_{52}	London South Eastern indicator	Proportion of train services into and around London (in general commuting services)	
q_{62}	$q_{42}q_{52}$	Interaction between Intercity and LSE proportions	

q_{72}	$q_{42}(1-q_{42}-q_{52})$	Interaction between intercity and regional (non-intercity and non-LSE services) proportions	reference to the relative size of train-km by each pre-merged TOC
q_{82}	$q_{52}(1-q_{42}-q_{52})$	Interaction between LSE and regional proportions	
q_{92}	Number of rolling stock types operated	Number of "generic" rolling stock types operated	National Modelling Framework Rolling Stock Classifications
y_3	Stations operated	Number of stations that the TOC operates	National Rail Trends
Prices			
P_1	Non-payroll cost per unit rolling stock		TOC accounts for cost, Platform 5 and TAS Rail Industry Monitor for rolling stock numbers
P_2	Staff costs (on payroll)		TOC accounts (both costs and staff numbers)

Source: Reproduced from Table 1 in Wheat and Smith (2014).

Findings on the cost structure of passenger train operations

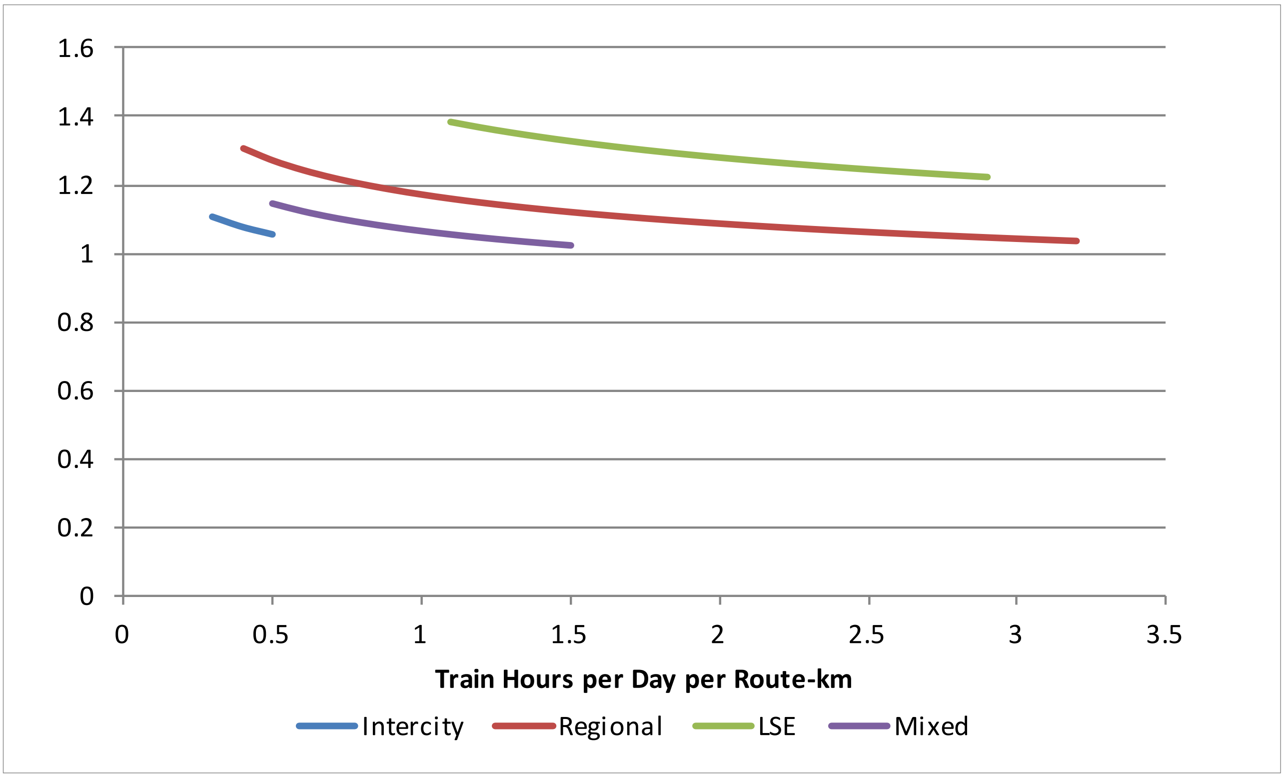
In this sub-section we present some results of the work undertaken to date to illustrate the richness and usefulness of the methods employed. Returns to scale (RtS) and returns to density (RtD) can be defined specifically for operations (as distinct from infrastructure). RtS measure how costs change when a firm grows in terms of geographical size. RtD measures how costs change when a firm grows by running more services on a fixed network.

The DEA analysis yields few results with relation to economies of scale or density. Indeed the paper by Cowie imposes constant returns to scale. Merkert et al (2009) did estimate a variable returns to scale model and found that British and Swedish TOCs were below minimum efficient scale, while the largest German operators were above. Of the parametric papers, Cowie (2002b) estimates a cost model which provides evidence on economies of scale. Cowie finds evidence for increasing returns to scale and that these are increasing with scale. There is no attempt to differentiate between scale and density economies in the analysis.

Smith and Wheat (2012) put forward a model which yields estimates of the extent of both economies of scale and economies of density, where the primary usage output is train-km. They found constant returns to scale and increasing returns to train density. The policy conclusion of this finding is that whilst there would not be scale benefits from merging franchises, such mergers may reduce costs by allowing greater exploitation of economies of density (a single operator running trains more intensively down a given route); though see further discussion on economies of density and heterogeneity below. One limitation of the Smith and Wheat (2012) work was the inability to estimate a plausible Translog function. Instead, a restricted variant was estimated, selected on the basis of general to specific testing and on whether key elasticities were of the expected sign. This implicitly restricts the variation in economies of scale and economies of density.

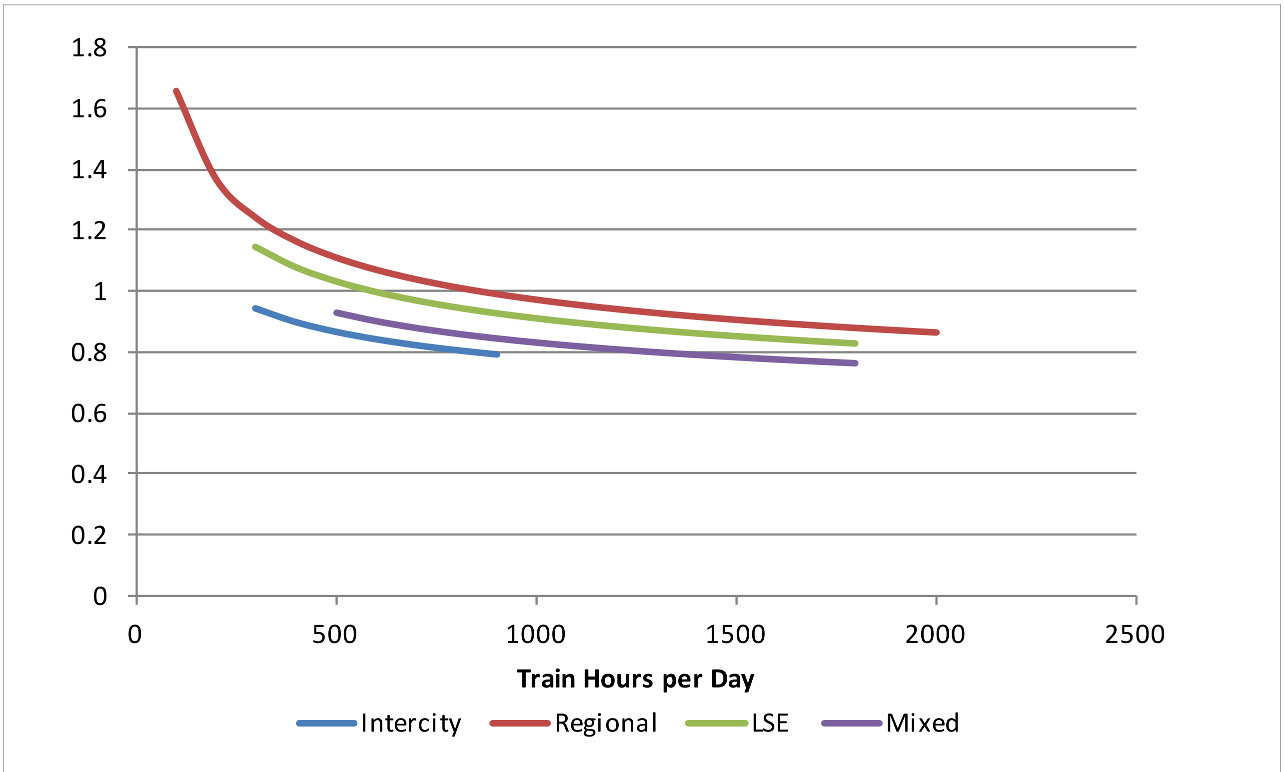
Further work by Wheat and Smith (2014) estimated a Translog simultaneously with the cost share equations and adopt a hedonic representation of the train operations output in order to include characteristics of output in a parsimonious manner. This work provides new insights into the scale and density properties of train operations, since it allows RtS and RtD to vary with the heterogeneity characteristics of output. Figures 2 and 3 summarise the findings on RtD and RtS.

Figure 2. Returns to density for different TOC types holding other variables constant



Source: Reproduced from Wheat and Smith (2014).

Figure 3. Returns to scale for different TOC types holding other variables constant



Source: Reproduced from Wheat and Smith (2014).

Figure 2 shows that all TOC types exhibit increasing RtD and that this does fall with density, although RtD are never exhausted within the middle 80% of the sample. At any given train hours per route km level, intercity TOCs exhibit the lowest RtD, while LSE (commuter services into London) exhibit the strongest (and indeed even at the 90th percentile density in sample the RtD estimate is in excess of 1.2). Intuitively, the curve for mixed TOCs is somewhere in-between the curves for intercity and regional. The policy conclusion from the analysis of RtD is that most TOCs should be able to reduce unit costs if there is further growth in train hours (on a fixed network) in response to future increases in passenger demand.

Figure 3 provides a similar plot for RtS. This shows that for all of the central 80th percent of the train hours distribution, intercity (and mixed) TOCs exhibit decreasing RtS. LSE TOCs exhibit increasing returns to scale only for the very smallest in sample, whilst regional TOCs are the only TOC type to have an appreciable range exhibiting increasing returns to scale. The results are consistent with a u-shaped average cost curve, although it would appear that most TOCs are operating at or beyond the minimum unit cost point. This finding has important implications for examining the optimal size of TOCs and is relevant to the recent franchise policy change that has resulted in substantial franchise re-mapping, and in turn larger franchises.

The overall conclusion from this section is that modelling returns to scale and density in passenger train operations potentially requires a rich model to fully capture the effects. The initial work published in Smith and Wheat (2012) based on a restricted translog model suggested broadly constant returns to scale combined with fairly strong economies of density. This may suggest that there could be a case for making franchises smaller, which could additionally help in reducing the risk of franchise failure, which has been a key problem in Britain. Britain's franchises are already considerably larger, in general, than those elsewhere in Europe. However, if reducing the size of franchises also increases the degree of franchise overlap, then important economies of density may be lost in the process, so it is not a clear policy conclusion. Turning the argument the other way round, larger franchises, that result in reductions in franchise overlap and the exploitation of economies of density, may reduce costs.

That said, Wheat and Smith (2014) develop a richer model, which takes account of service heterogeneity (in particular, in terms of train speed and TOC type) in relation to returns to scale and density. In that later paper it is found that the ability to exploit economies of density may be constrained by service heterogeneity. Likewise, the losses of economies of density from reducing franchise size might be smaller than indicated above. It is further found that some franchises in Britain are operating at decreasing returns to scale, and may therefore be too large.

What the above research suggests is that it is possible to shed new light on the structure of costs of passenger train operations, and draw broad conclusions about the economies of scale and density of those operations. The most recent work suggests that there could be cost savings from reducing franchise size (because of scale diseconomies) and that losses in economies of density might be reduced by service heterogeneity.

5.4. European rail system benchmarking

Whereas separate studies of infrastructure and operations are largely confined to Britain, there are many studies of European rail systems (infrastructure and operations) as a whole. These studies add together infrastructure and train operating costs. However, there remain serious data problems and these are getting worse. Most past studies that concentrate on the impact of reforms at the European level have applied data envelopment analysis to physical data; our problems with that approach have been outlined above. Moreover they have usually used the data published by the Union International des Chemins de Fer

(UIC)[25], data which has been shown to contain inconsistencies (van de Velde *et al.*, 2012). Moreover this source only contains data on UIC members, generally the incumbent but not new entrants, and in some cases covers their activities in a number of countries rather than just their home country. A rare example of the estimation of cost functions to study the impact of European reforms is Asmild et al (2008); they also went to considerable efforts to clean up and supplement the UIC data. They found that competitive tendering for passenger, open access for freight services and accounting separation of infrastructure from operations all improved efficiency, but could find no further effect of complete separation of infrastructure from operations. However, their data series ended in 2001 before many reforms took place.

A number of studies have been undertaken using DEA. In one of the more recent studies, Cantos et al (2010) uses data envelopment analysis on physical data for 16 railways for the period 1985-2005. In a first stage they use two outputs (passenger km and freight tonne km) and four inputs - employment, number of passenger carrying vehicles, number of freight wagons and route km – to compile measures of efficiency. They then regress these measures on variables reflecting the operating conditions of the railway concerned (percentage of train km that are passenger, traffic density in terms of train km per route km, mean passenger train load and mean freight train load) and on variables reflecting vertical separation and introduction of competition.

They find separate beneficial effects of vertical separation and introduction of competition in the freight market, whereas passenger franchising has no such effect. However, of the four countries in their sample in which passenger franchising has been introduced only in Sweden and Germany has it covered a significant proportion of regional services and in none has it covered commercial services. That vertical separation has improved efficiency over and above the impact of competition is surprising; we worry that this may be because vertically separated systems undertake more subcontracting and that this has not been picked up in the physical data, but we have no evidence on this issue.

Cantos (2011) uses a greater sample of 23 countries and a more up to date period of 2001-8. It repeats the data envelopment analysis approach of the earlier paper, but also applies a stochastic frontier approach, showing that this leads to much lower efficiency scores although the ranking of countries in terms of efficiency is similar. The results of the second stage analysis are rather different however. Vertical separation has no significant effect on efficiency, whilst the strongest positive effect on efficiency comes from passenger tendering. Freight open access has a positive effect on efficiency that is significant in one of the models. Unfortunately, they seem to have a less adequate set of variables as controls in the second stage analysis in this paper. They only include population density and rail route length. The results could be considerably biased by the lack of data on passenger and freight train loads, which tend to be heavily influenced by the geography of the country and government policy, but which are major determinants of efficiency.

A recent example of the use of a translog cost function with panel data from a large number of countries is Mizutani and Uranishi (2013). They used data for 30 railway companies from 23 OECD countries for the years 1994 to 2007, giving 420 observations. Whilst most of the observations were from Europe, they included the vertically integrated passenger railways of Japan, and also South Korea and Turkey. Where vertical separation had been implemented, they added together the infrastructure manager and the train operating company to form a single observation. The basic source of data was UIC, but this was supplemented as necessary by data from company annual reports.

[25] Note this is published data as distinct from the confidential data from the LICB project described earlier.

Two separate models were estimated, one using passenger kilometres and freight tonne kilometres as outputs and the other total train kilometres, with the share of passenger revenue to total revenue, passenger load factors and length of haul and freight number of cars per train to reflect differences in the characteristics and therefore costs of the train kilometres. Factor prices for labour, track, rolling stock and other materials were estimated. Finally route kilometres, train kilometres per route kilometre and the percentage of line electrified were included as descriptors of the network.

The rail reform variables were dummies reflecting complete vertical separation (the holding company model being regarded as integrated) and horizontal separation of passenger and freight operations. It was found that whilst horizontal separation unequivocally reduces costs, vertical integration only reduced costs for densely trafficked railways; for most European railways it increased them. Given that there are no separate variables representing the degree to which competition is permitted or actually takes place, it must be assumed that these impacts are the net effect of any additional costs directly caused by vertical separation and of the impact of competition which in most of Europe must presumably be sufficient to outweigh these costs. The explanation given for the impact on costs varying with density is that given above, that the transactions costs caused by vertical separation will be much greater in densely trafficked networks than in less densely trafficked ones.

Van de Velde et al (2012) takes this work further by updating and improving the data set and introducing separate dummy variables for holding companies and complete vertical separation (this work also later published in the academic literature, see Nash *et al.* forthcoming and Mizutani et. al, 2014). They also added in Britain, the country in which the most radical reforms had taken place, but which had been excluded from most previous studies due to lack of data. Finally, they introduce dummy variables representing passenger and freight market competition.

They confirm the previous finding that, compared with complete vertical integration, vertical separation reduces costs at low levels of density but increases them at high; at mean European density levels costs are not affected by the change. This effect is not likely to be one of pure transactions costs (negotiating and enforcing contracts), which have been shown to be a relatively small proportion of total systems costs (Merkert et al, 2012) but is more likely a problem of misalignment of incentives leading to poor integration of infrastructure and operations in circumstances (dense traffic) when this is particularly important. They find weak evidence (significant at 10% only) that the holding company model reduces costs compared with vertical integration, but this does not vary with density, so the holding company would be preferred to vertical separation at high levels of density but not at low.

Within the range of the data, the introduction of competition seems to have had no effect on costs. Horizontal separation of freight and passenger undertakings seems to have sharply reduced costs (perhaps because this has typically been associated with preparation of the freight undertaking for privatization), whilst a high proportion of revenue coming from freight rather than passenger tends to increase the costs of vertical separation (perhaps because planning freight services efficiently requires closer day to day working than passenger, since freight services vary from day to day whereas passenger services are generally fixed for the duration of the timetable). The paper also provides qualitative evidence on the issue of how misalignment of incentives may raise costs and show how, whilst efficiently set track access charges and performance regimes are important, they do not provide incentives for railway undertakings to assist infrastructure managers in seeking the minimum cost solution to infrastructure provision. Only a mechanism for sharing of changes in costs and revenues, as provided for in some of the alliances now being negotiated in Britain, will achieve that.

The conclusion of the studies in this section is that there is no one size fits all policy for European railways. Based on a mixture of qualitative and quantitative research the evidence suggests that vertical separation may perform less well than the Holding company model for intensely used networks, whilst being the structure of choice for less dense networks. Whilst this is in part intuitive, it is not totally clear why separation reduces costs for lightly used networks, particularly if there is little competition. It is further disconcerting that it has not been possible to find clear competition effects in the data. To date no research has yet been published to consider the impact of regulation on costs, though research is ongoing at University of Leeds in this respect. A final note must be that although we consider the cost function approach to be the best, and with the van de Velde *et al.* (2012) study incorporating new data from CER members to supplement published data, there nevertheless remains work to be done on the data side to improve its comparability.

5.5. Concluding remarks

The literature contains a rich set of studies focused on analysis of rail costs for efficiency (and other) purposes, covering infrastructure, train operations and the vexed question as to what extent infrastructure and operations should be separated or combined. Key issues in all cases are to develop a technology that reflects the possibility of economies of scale and density, and to deal with heterogeneity through specifying a rich set of explanatory variables. This often then leads to a parametric approach as the most readily able to deal with this issues in a transparent way. Where data does not exist on key variables, it is accepted that the residual contains both unobserved factors, and inefficiency; and decomposition of the residual has been dealt with through either advanced econometric techniques or through regulatory judgement (e.g. through the application of upper quartile techniques).

In addition to questions of model specification and efficiency decomposition, the general question of whether the data is of good quality and comparable across firms (and particularly across countries) is a key challenge (perhaps the greatest challenge) for economic regulators doing benchmarking. Measurement of capital can be a particular challenge – e.g. lumpy renewal volumes and costs – and a number of methods have been applied to deal with this problem. Past studies have also made good use of internal benchmarking to overcome the issues associated with international studies and sometimes combined these two approaches, joining together regional data within countries from several different countries. This latter point enables the study of what might be the optimal size of a infrastructure region. In the train operations literature this leads to a similar question – what is the optimal size of a rail franchise? Indeed these questions of scale (and density), though not formally captured as inefficiency, could lead to large cost savings if answered correctly. Broadening the discussion, it can be difficult to view rail infrastructure in isolation from train operations, and much study has been done to look at whether it is cheaper to integrate the two. A key finding in a literature with somewhat mixed results is that the answer may depend on the density of usage; with more intensely used networks more amenable perhaps to vertical integration.

Finally, is important to link benchmarking analysis with underlying business and engineering understanding. The latter can come in two forms. First in helping to interpret the elasticities on cost drivers in a cost function – do they make sense? If so this implies that the model has greater credibility. Second, bottom-up analysis can help understand why there is an efficiency gap. This approach was used most notably by ORR in 2008, when the econometric model indicated a gap between Network Rail and the best infrastructure managers in Europe of up to 40%. Bottom-up analysis of best practice provided supporting evidence to explain that gap. More recent policy in Great Britain has focused on bottom-up studies – it is accepted that a gap exists, and the focus has been on getting costs down (though this has not been altogether successful yet in the British context, though for a complex set of reasons).

Chapter references

Affuso, L., Angeriz, A., Pollitt, M.G., (2003), *Measuring the Efficiency of Britain's Privatised Train Operating Companies*, mimeo (unpublished version provided by the authors).

Asmild, M., Holvad, T., Hougaard, J.L., Kronborg, D., (2009), "Railway reforms: Do they influence operating efficiency?", *Transportation*, 36 (5), 617-638.

British Rail/University of Leeds (1979), *A Comparative Study of European Rail Performance*, British Rail, London.

Cantos, P., Pastor, J.M., Serrano, L., (2010), "Vertical and horizontal separation in the European railway sector and its effects on productivity', *Journal of Transport Economics and Policy*, 44 (2), 139-160.

Cantos, P., Pastor, J.M., Serrano, L., (2012), "Evaluating European railway deregulation using different approaches", *Transport Policy*, 24, November, 67-72.

Cowie, J., (2002a), "Subsidy and Productivity in the Privatised British Passenger Railway", *Economic Issues*, 7 (1), 25-37.

Cowie, J., (2002b), "The Production Economics of a Vertically Separated Railway – The Case of the British Train Operating Companies", *Trasporti Europei*, August, 96-103.

Cowie, J., (2005), "Technical Efficiency versus Technical Change – The British Passenger Train Operators". In Hensher, D. A., (2005). *Competition and ownership in land passenger transport: selected refereed papers from the 8th International Conference* (Thredbo 8) Rio de Janeiro, September 2003. Amsterdam; London: Elsevier, 2005.

Cowie, J., (2009), "The British Passenger Rail Privatisation: Conclusions on Subsidy and Efficiency from the First Round of Franchises*", Journal of Transport Economics and Policy*, 43 (1), 85-104.

Cuesta, R.A., (2000), "A production model with firm-specific temporal variation in technical inefficiency: with Application to Spanish Dairy Farms", *Journal of Productivity Analysis*, 13 (2), 139-152.

Kennedy, J., Smith, A.S.J., (2004), "Assessing the Efficient Cost of Sustaining Britain's Rail Network: Perspectives Based on Zonal Comparisons", *Journal of Transport Economics and Policy*, 38, 157-190. LEK, TTCI, Halcrow, (2003), *Bottom-up review of Network Rail's business plan: 2003/04-2005/06*, Report to ORR, London.

Merkert, R., Smith, A.S.J., Nash, C.A., (2009), "Benchmarking of train operating firms - A transaction cost efficiency analysis", *Journal of Transportation Planning and Technology*, 33 (1), 35-53.

Mizutani, F., Uranishi, S., (2013), "Does vertical separation reduce cost? An empirical analysis of the rail industry in OECD countries", *Journal of Regulatory Economics*, 48 (1), 31-59.

Mizutani, F., Smith, A.S.J., Nash, C.A., Uranishi, S., (2014), "Comparing the Costs of Vertical Separation, Integration, and Intermediate Organisational Structures in European and East Asian OECD Railways", mimeo (available from the authors on request).

Nash, C.A., Smith, A. S. J., van de Velde, D., Mizutani, F., Uranishi, S., (forthcoming), "Structural reforms in the railways: incentive misalignment and cost implications, Research in Transportation Economics".

ORR (2018), PR18 Econometric top-down benchmarking of Network Rail, London.

Smith, A.S.J., (2008), International Benchmarking of Network Rail's Maintenance and Renewal Costs: An Econometric Study Based on the LICB Dataset (1996-2006), *Report for the Office of Rail Regulation*, October 2008. Available at: http://www.rail-reg.gov.uk/server/show/nav.2001

Smith, A.S.J., (2012), "The application of stochastic frontier panel models in economic regulation: Experience from the European rail sector", *Transportation Research Part E*, 48, 503–515.

Smith, A.S.J., Wheat, P.E., (2012a), "Evaluating Alternative Policy Responses to Franchise Failure: Evidence From the Passenger Rail Sector in Britain", *Journal of Transport Economics and Policy*, 46 (1), 25-49.

Smith, A.S.J., Wheat, P.E., (2012b), "Estimation of Cost Inefficiency in Panel Data Models with Firm Specific and Sub-Company Specific Effects", *Journal of Productivity Analysis*, 37, 27-40.

Smith, A.S.J, Nash, C., Wheat, P., (2009), "Passenger Rail Franchising in Britain – has it been a success?", *International Journal of Transport Economics*, 36 (1), 33-62.

Smith, A.S.J., Wheat, P., Nash, C.A., (2010), "Exploring the Effects of Passenger Rail Franchising in Britain: Evidence from the First Two Rounds of Franchising (1997-2008)", *Research in Transportation Economics*, 29 (1) 72-79.

Smith, A.S.J., Wheat, P.E., Smith, G., (2010), "The role of international benchmarking in developing rail infrastructure efficiency estimates", *Utilities Policy*, 18, 86-93.

van de Velde, D., Nash, C., Smith, A., Mizutani, F., Uranishi, S., Lijesen, M., Zschoche, F., (2012),*EVES-Rail - Economic effects of Vertical Separation in the railway sector*, Report for CER - Community of European Railway and Infrastructure Companies, by inno-V (Amsterdam) in cooperation with University of Leeds – ITS, Kobe University, VU Amsterdam University and Civity management consultants, Amsterdam/Brussels.

Wheat, P.E., Smith, A.S.J., (2014), "Do the usual results of railway economies of scale and density hold in the case of heterogeneity in outputs: A hedonic cost function approach", *Journal of Transport Economics and Policy* (online first, January 2014).

6. BUS BENCHMARKING

Giovanni Fraquelli and Marco Ponti

6.1. Benchmarking the cost structure of bus industry

The operation of public transport services has a significant impact on the budget of territorial bodies, since in most cases the revenues from end users tickets and subscriptions are not sufficient to recover the cost of providing the service. Cost-based compensation instead of incentive-based mechanisms predominated in the last decades. In recent years, in order to induce more efficiency, enhance productivity and reduce huge deficits, many countries have put in place reforms in which the institutional reorganization of the industry is combined with the design of new regulatory measures that foresee specific incentives to increase efficiency.

An effective regulation of bus industry cannot be implemented without a detailed analysis of the cost structure of the urban and suburban transport modes (such as motorbus, tramway, streetcar, trolleybus, metro and coach renting). The study of the economic and technological characteristics of the sector is a fundamental step, since it allows the identification of the proper configuration of the network. It can provide guidance in favour of regulatory interventions, while it may suggest the opportunity of opening to competition in the market, or to the competition for the market if the empirical investigations exhibit natural monopoly characteristics.

6.2. Benchmarking efficiency

The study of the comparative efficiency of public transport companies contributes to decrease the information asymmetry between policy makers and public transport firms and represents a good input to the debate on public transport regulation. Results of empirical analyses on efficiency and productivity of single operators are sensitive to the context of the study, to the methodological approaches adopted to measure efficiency and to the nature of inputs and outputs. Efficiency performance in the local public transport is the object of many empirical analyses because the industry attracts the interest of the schools of economics, management and engineering. Within the wide available literature, we will refer to two reviews that adopt different visions and summarize the most significant results. The first one has mainly an economic perspective, while the second one combines economic and engineering approaches.

6.3. Benchmarking efficiency: an economic perspective

The review by Jarboui *et al.* (2012) considers 24 works relating to the period 2001-2011 and therefore represents a useful benchmarking tool for the first decade of the 2000s. The works are classified paying attention to the nature of the paper (theoretical or empirical), to the countries of the study, to the methodological approach and to the nature of outputs and inputs.

Table 1 summarises the results. We can see that Norway (5 studies) and US (4 studies) are the countries analysed most frequently. However, other 11 countries (Canada, France, Germany, Great Britain, Greece, India, Italy, Portugal, Spain, Switzerland, Taiwan) are involved in the reviewed works. Only the studies concerning Brazil versus 7 European countries (Sampaio *et al.*, 2008) and Norway and France (De Borger *et al.*, 2008) adopt an international comparative perspective. As to the methodological approaches, the largest part of the works has adopted the DEA technique (21 studies); other studies adopt the Stochastic Frontier Analysis approach (SFA, 4 studies), Free Disposal Hull (FDH, 1 study) and three of them combine the different methods, i.e. DEA-SFA, DEA-SFA-FHD and DEA-MEA (Multi directional Efficiency Analysis).

Table 1. Main results of the review

Years and Countries of the study	Methodological approach (Number uses)	Outputs	Inputs	Efficiency scores
2001Spain, Norway	Dea (21)	Vehicle-KM	N. of employees	Nonparametric and parametric approaches results in different efficiency scores.
2002 Great Britain,	SFA (4)	Seat – Km	N. staff employed	
Theory	FDH (1)	Passenger-Km	Driving hours	Technical inefficiency occurs in many countries (USA, Canada, Italy, India) and the average value is about 20-25%.
2003 USA, Norway	DEA – SFA (1)	N. passengers	Fuel used	
2004 Norway, Canada,	DEA- SFA-FDH (1)	Annual ridership	Energy	
USA	DEA- MEA (1)	Bus utilisation	Vehicle operated	Comparative studies in different countries (Brazil versus 7 European countries and Norway-France) indicate a great variability, both between and among countries.
2006 Switzerland		Load factor	Vehicle capacity	
2007 Greece, Italy, USA		N.Pass./dist.travell.	Fleet size	
2008 Norway,		Revenue	Tot. n. bus-km	AS far as economies of scale: decreasing returns in USA 2004; Brazil vs 7 European countries (2008) 56% transport system increasing R.S., 29% decreasing R.S., India (2011) increasing return to scale are dominant.
Brazil vs 7 E.Countries,		Average trav.time	Network length	
Norway- France		Frequency	N. of intersections	
2009 Spain, Taiwan		Bus-Km/inhabit.	Priority lines	
2010 Portugal		Accessibility	Service duration	
2011 India, USA,		Comfort, safety	Driver costs	Small operators – increasing R.S.; Average-sized firms: constant or increasing R.S.; Large systems: decreasing R.S.
India, India		Acc. rate populat.	Fuel cost	
			T. operating ex.	

With regard to datasets, in the first years of the decade, we find some cross sections or time series, but starting from the year 2007, usually studies use panel data in their applications.

The output variables, in most of the works, consist of supply-oriented measures such as vehicle-km and seat-km, or demand oriented measures such as passenger-Km, bus utilization and ridership; moreover, revenues and sales are used in several studies. The output specification represents a very important issue and the debate in literature is broad:

> *"…. The main arguments have been summarised by Berechman (1993). First, inputs do not vary systematically with demand-oriented output measures so that they do not allow an adequate description of transport technology. Second, supply-oriented output indicators are to a larger extent more under the control of transport companies than demand-oriented outputs[…] Third, independently of the achievement of frontier aims defined in terms of passenger transport services really consumed, supplying the transport services in the least costly way may be considered a best requirement for transport operators. Therefore, the focus should be on pure supply indices when measuring productivity and efficiency."* (Jarboui *et al.*, 2012, p.122).

When the focus is not only on pure efficiency, but also includes quality of service, the outputs encompass several other variables, such as average travelling time, frequency, accessibility, comfort, safety, the rate of accidents in the reference population.

As input variables, most studies use labour, capital and energy, although measured by different variables. For example, labour can be expressed by the number of staff employed, by the driving hours, or in terms of driver costs. Capital includes prevalently vehicles operated, fleet size, vehicle capacity. Among the energy variables, we find the fuel used, the fuel cost and the lubricant cost. In some cases, the quantities of inputs are missing and are replaced by the total operating expenses. Many studies include environmental variables to better account for the configuration and the quality of the network: we find the length of service, the service duration, the number of intersections, the priority lanes, the size of the park, the number of accidents per kilometre.

As for the empirical outcomes, technical inefficiency is reported with respect to many countries (USA, Canada, Italy, India) and the average value is about 20-25%. We know that deterministic-nonparametric and stochastic-parametric approaches may generate relatively different efficiency scores. For this reason, it is impossible to make direct comparisons between the different works. Strong inefficiencies are highlighted in almost all countries, even if the comparative studies (Brazil versus seven European countries and Norway – France) indicate great variability of the technical inefficiency, both within and among countries. De Borger *et al.* (2008) attribute such variability to the heterogeneity of the service management in the different countries and in particular to the differences in the regulatory framework, population density, quality management and operating environment. The works cited are not specifically oriented to the evaluation of the economies of scale but provide useful information on this subject. In general, small firms are characterised by increasing returns to scale, while average-sized operators evidence increasing or constant returns to scale, and we find decreasing returns to scale for large companies.

A great number of variables can affect efficiency, but many authors agree about the working and operations practises, the quality of the management and the presence of low financial constraints. Technical investments seem to be an important source of efficiency recovery. Sometime regulation (or deregulation) and privatization are good solutions to improve efficiency, although this topic is strongly debated. Studies of the different operating contexts based on different approaches and representative indicators are lacking.

6.4. Benchmarking efficiency: an engineering approach

The review by Daraio *et al.* (2016) provides a classification of the existing literature on economic efficiency in the urban public transport sector. The analysis concerns 124 paper starting from 1970. The authors identify five areas of research dealing with efficiency. The study aims to deepen different issues:

- technical efficiency (tools from the efficiency/productivity literature, focusing on cost minimisation and/or profit maximisation);
- the factors affecting technical efficiency. By non-parametric frontier analysis (NFA) and parametric frontier analysis (PFA), the studies try to identify a number of possible explanatory variables;
- the effect of alternative regulatory regimes on the efficiency of the operation;
- the relative advantages of public versus private ownership and/or operation;
- the economies of density, scope and scale (space and time dimension of service provision and characterisation of transportation technology).

The authors propose a general evaluation framework, summarising the different approaches of research with respect to six issues, performance measures and goals: economic goals; operational performance of the transport system; role in relieving road congestion; environmental impacts in terms of sustainability; social inclusion issues; territorial accessibility), reported in Table 2:

> *" [...]The first one summarises the economic perspective, the second one the key interest of the travellers, while the last four represent the main benefits that are typically expected by the community when implementing public transports systems…"* (Daraio *et al.*, 2016, p. 3).

As reported in table 2, the analysis gives evidence that, in all the reviewed works, profit and cost aspects related to the operation of the service are considered. Service performance is examined mainly in terms of travel and waiting times, equally addressing the viewpoint of the operator and the customer. Differently from the economic literature, the engineering literature gives less attention to the other four evaluation perspectives (the last four rows) listed in table 2. However, congestion issues have an important role in the analysis of competing transport means, while sustainability issues, related to external environmental costs, represent an essential component in every process of design of public transport.

Table 2. Matching evaluation aspects with producer viewpoint and with literature on efficiency

Evaluation aspect	Relevance for the producer viewpoint (efficiency)	Relevance for the evaluation of efficiency and productivity	Relevance for the evaluation of determinants of technical efficiency
Profit/cost analysis	***	***	***
Service performance	**	***	***
Road congestion	**	*	**
Sustainability	*	**	**
Social inclusion	*	*	*
Accessibility	*	*	*

* = little relevant; ** = relevant to some extent; *** = strongly relevant

Source: Our synthesis of table n. 1 p.3, Daraio *et al.*, 2016.

The set of input and output variables adopted gives the same evidence previously seen within the first review. *"…Input variables, these normally fall in two main categories: "physical" production factors with their own measurement units (number of employees, hours of work etc.) on one side, and costs in monetary units on the other, that are further split into capital expenses and operating expenses[…]. The number of employees (or hours of work), fuel consumption, number of vehicles in the fleet are largely the most considered variables since they represent the main inputs in the production process. […] As regards the other clusters of input variables we analogously find that the prices of labour, capital and fuel are the far more utilised while only a limited number of papers […]consider a broader set of cost oriented input variables such as maintenance and overhead costs[...]"* (Daraio *et al.*, 2016, p. 6).

On the output side, the authors categorise three main groups: variables that measure the efficiency of the service from the supply point of view (vehicles or seats by travelled Km are the ones most frequently used); variables related to the effectiveness of the production process with respect to the demand (number of passengers and passengers by travelled km); monetary variables related to the service revenues.

An additional set of variables that can be considered as inputs or outputs refers to quality and characteristics of the service, such as service performances (commercial speed, punctuality), and accessibility (length of network, number of stops). Some works consider the average age of the fleet (this variable refers contemporarily to efficiency and effectiveness). Variables describing the socioeconomic characteristics of the service and the management approach are also typically considered. In synthesis, the input-output variables emerging from the engineering literature are related to the following aspects: operational efficiency, intensity of use of the service by passengers, service use related to input (expenses/passenger-KM), relative service dimension (fleet dimension/population of the service area), service coverage (lines length/population of the service area), dimension of the potential market (passengers/population of the service area), revenues generation, externalities (number of accidents/(vehicle∗kilometre)).

As for the methodological approach, the review identifies two groups of methods that are prevailing in LPT studies. On the one hand, we find the dominance of the parametric methods (Cobb-Douglas costs functions and trans-log cost functions). These studies use Stochastic Frontier Analysis (SFA) and multistage analyses of cost functions in relation to total cost estimates. To investigate costs, however, a smaller group of studies adopt DEA and total factor productivity indexes (Malmquist indexes). On the other hand, nonparametric methods based on DEA and indexes dominate the empirical analysis of factor productivity. They "are more flexible and are characterised by a wider range of used variables…" (p. 9). The effects of external environmental variables on efficiency are widely studied in the field of nonparametric efficiency analysis: the efficiency scores, estimated in a first stage, are regressed on environmental variables using a "two-stages approach".

The authors suggest potential improvements of the research by promoting and developing international comparable data at the national and regional level.

6.5. Benchmarking economies of scale and scope

Strong evidence about the optimal dimension of the local transport network is necessary to have an efficient policy making, since it allows planning and designing the provision of the service. In particular, the design concerns the extension of the service area, the frequencies of buses and the number of bus lines. On the other hand, the knowledge of the underlying production technology gives the possibility to set pricing policies. "*[…]A finding of diseconomies of scale can imply that, for example, a city can have different parts of its system operated by separated companies at a lower unit cost of output*" (Karlaftis and McCarthy, 2002, p.…).

Moreover, the operators are often multi-service firms. They operate in regulated market such as urban and intercity transport and in non-regulated markets, such as a long distance express coach and long distance hired coach services. Coach services can represent an important complement of regular transit system (Talley 2007). They can be easily interconnected with other modes of transport. Hired coach travel are typically characterized by non-scheduled times and non-fixed routes. For this reason this service is mainly addressed to occasional users, such as the tourist sector. Conversely, after the process of transport liberalization (2009), long distance coach transport is growing exponentially so as to directly compete with railways and airlines services. If multi-service firms are exploiting scope economies, it is desirable to let them run integrated production in regulated and unregulated services.

Early studies on the analysis of costs in transportation were mainly focused on the effects of diversification among different transit modes within the same urban area.

Colburn and Talley (1992), by analysing different modes of transport in urban systems, find evidence of limited cost complementarities. Viton (1993), by investigating the processes of aggregation between different suppliers, shows that cost savings resulting from mergers depend on the transport modes of the companies as well as on the number of firms involved in the merger. Farsi *et al.* (2007), exploring a Swiss multi-modal transport system, show the presence of economies of scale and scope and support integrated multi-mode operations as opposed to unbundling. As these authors argue, "*when transport modes are legally unbundled, bidding can be opened to both single-mode operators and multi-mode companies. Whereas in the present situation the competition is difficult for companies specialized in a single transit service because of the comparative advantage of the incumbent multi-mode companies. In this case, due to fewer potential bidders, the benefits from competition for the market would be lower*" (p. 2) […] "*In the case* *of multi-modal systems the regulator has to decide to open the competitive tendering procedure for supplying the entire transport services or to unbundle the multi-modal systems and open separate tenders for different modes of transport"* (p. 1)[…] *"Therefore, it is relevant for the local authorities to know if and how much a multi-mode supplier could use the scope and scale economies to reduce their costs in comparison to a group of single-mode operators"* (p. 2).

In order to estimate scale and scope economies, the most popular method is to use a multi-output specification of the cost function.

As for scale economies, Gagnepain *et al.* (2011) report that a significant number of empirical studies are in line with a U-shaped average cost curve, exhibiting increasing returns to scale for smaller operators and decreasing returns beyond a certain output level. As an example, Cowie and Asenova (1999) estimate that small companies (with a bus fleet of less than 200 vehicles) experience some economies of scale.

6.6. Efficiency and economies of scale and scope: some findings about Italy

Fraquelli *et al.* (2004) investigate the cost behavior and the existence of scope economies by relying, for the estimation, on a set of dummy variables to distinguish between specialized operators (in urban or intercity service) and integrated companies, and find evidence of lower costs for integrated bus transport firms.

Cambini *et al.* (2007), focusing on a set of medium and large Italian municipalities, find evidence of economies of scale in most cases, suggesting that companies should operate on the entire system of urban network, without fragmentation of the service. They also argue that mergers between operators of neighboring cities or between suppliers of urban and intercity transit services would be desirable in order to reduce operating costs.

Di Giacomo and Ottoz (2010) estimate a total cost function for multi-service LPT companies. They highlight very mild scope economies (around 2%) between urban and intercity services. The extent of scope economies tends to decrease as the firm size increases, and modest scale economies are observed for the median firm. Ottoz and Di Giacomo (2012), analyzing the LPT system of a specific Italian region (Piedmont), provide empirical evidence of the impact on costs of different diversification strategies. The results show the presence of scope economies for the median firm ranging between 16% and 30%, depending on the cost function specification as well as on the number of outputs. Lower global scope economies are found for publicly owned firms and, more in general, for large operators.

A recent analysis including coach operators (Abrate *et al.*, 2016) investigates the presence and the magnitude of scale and scope economies in the provision of passenger transport services, using a sample of Italian bus and coach operators. They estimate a multi-product cost function including hired coach, urban and intercity passenger services as three separate outputs.

The dataset is an unbalanced panel of 47 firms observed during the years 2008–2012, for a total of 147 observations. Of these, 30 observations refer to specialized firms, while 9 observations refer to fully integrated firms. The vast majority, however, is represented by firms performing a couple of services (in particular intercity and for-hire services, or intercity and urban). Data related to costs, output quantities and input prices were obtained by integrating the information available in the annual reports of each company with additional information drawn from questionnaires directly sent to managers. The long-run total cost (C) is the sum of the expenditure for fuel and other raw materials and services, labor and capital costs of each firm. All monetary values are expressed in 2010 constant prices. The three output categories (urban transit, intercity transit and for-hire transit), are measured by vehicle-kilometres.

As to the methodology, a Composite Cost Function econometric model (Pulley and Braunstein, 1992) was used. It allows to disentangle potential synergies emerging when firms provide different combinations of the three types of transit services.

The research finds evidence that the average firm exhibits constant aggregate returns to scale and is characterized by the absence of global scope economies. Small multi-service firms, however, may benefit from cost reductions with respect to specialized operators. As the size of the firm increases, the cost savings survive only for the intercity bus service. For firms larger than the average, the presence of decreasing returns to scale counterbalances the effect of scope economies. The results suggest that firms for which the core business is urban transport can benefit from diversifying into intercity services, while firms specialized in intercity services can exploit scope economies by diversifying into coach renting services. A diversification strategy involving all three activities is not beneficial (except for very small operators).

6.7. Further research needs

A particularly important aspect of public transport concerns the degree of use of the service; the analysis carried out above (Table 6.1) highlights, among the outputs, variables such as: passenger-Km, number of passengers, annual ridership, bus utilisation. In fact, in some circumstances the bus transport demand is considered, but further insights are necessary. For instance, concerning the operational and environmental costs, often the main issue is not cost reduction: it is, above all, increasing the number of passengers. This aspect suggests the need of further insights with respect to the variables related to modal transfer, such as the switch from private car to bus or to other public transport options.

As for scale and scope economies, especially in the urban case, there is the need of examining the possibility to expand outputs by diversifying towards other similar activities. Some works give evidence that the diversification towards intercity services could be a valid option, when the size of the urban area does not allow a minimum efficient dimension of the firm. On the other hand, it could be useful to verify if the demand for mobility in large metropolitan areas creates the conditions for having separate operators providing urban and intercity services and which is their optimal size. As for hire services, we need further research to investigate if their peculiar characteristics related to non-scheduled times and non-fixed routes favour or not their integration with other transit services.

Chapter references

Berechman, J., (1993), *Public transit economics and deregulation policy*. North-Holland, Amsterdam.

Cambini, C., Piacenza, M.,Vannoni, D., (2007), "Restructuring public transit systems: evidence on cost properties and optimal network configuration from medium and large-sized companies", *Review of Industrial Organaisation*, 3, 183-203.

Cowie, J., Asenova, D., (1999), "Organisation form, scale effects and efficiency in the British bus industry", *Transportation* 26, 231–248.

Daraio, C., Diana, M., Di Costa, F., Leporelli, C., Matteucci, G., Nastasi, A., (2016), "Efficiency and effectiveness in the urban public transport sector: A critical review with directions for future research", *European Journal of Operational Research*, 248, 1-20.

De Borger, B., Kerstens, K., Staat, M., (2008), "Transit costs and cost efficiency: bootstrapping non-parametric frontiers", *Research in Transportation Economics*, 23(1), 53–64.

Farsi, M., Fetz, A., Filippini, M., (2007), "Economies of scale and scope in local public transportation", *Journal of Transport Economics and Policy*, 41 (3), 345-361.

Gagnepain, P., Ivaldi, M., Muller, C., (2011), "The industrial organization of competition in local bus services", in de Palma, A., Lindsey, R., Quinet, E., Vickerman, R., (Eds.), *A Handbook of Transport Economics*, Edward Elgar, Cheltenham (UK), 744–762.

Jarboui, S., Forget, P., Boujelbene, Y., (2012), "Public road transport efficiency: a literature review via the classification scheme", *Public Transport*, 4 (2), 101-128.

Karlaftis, M. G., McCarthy, P., (2002), "Cost structures of public transit systems: a panel data analysis", *Transportation Research,* Part E, 38, 1-18.

Sampaio, BR., Neto, OL., Sampaio, Y., (2008), "Efficiency analysis of public transport systems: lessons for institutional planning", *Transportation Research*, Part A, Policy and Practice, 42 (3), 445–454.

Talley, W.K., (2007), "Classifying urban passenger transportation services", in Coto-Millán, V., Inglada, V. (Eds.), *Essay son Transport Economics*, Springer, New York,. 65-77.

Viton, P., (1993), "How big should transit be? Evidence from San Francisco Bay Area", *Transportation*, 20, 35-57.

7. SEAPORTS AND AIRPORTS

Eddy Van de Voorde and Chris Nash

7.1. Introduction

Ports and airports are important parts of the transport system. This is due to the large volume of goods and passengers they handle, but also to derived effects in terms of employment and investment. Both ports and airports are crucial nodes in complex logistics chains and as such a welfare generating actor for a region and/or country.

Competition between and within ports and airports is therefore fierce. Many actors are involved in that competition. Consequently, port and airport competition is influenced by a multitude of related and sometimes even conflicting interests. For a port three types of competition may be discerned, i.e. intra-port competition at operator level (competition between port undertakings within a single port), external port competition at operator level (competition between port undertakings from different ports), and inter-port competition at port authority level.[26] The same types of competition applies on airports. A daily fight for throughput and passengers!

But the question arises whether all ports and airports do face strong competition. There may be a case in which one or more forms of market failure lead to allocative inefficiency. Possible causes of market failure include market power, asymmetric information, externalities and public goods. At that moment one starts thinking about regulation. A major argument in favour of regulation is based on the natural monopoly character of some industries. Production at the minimum possible long-run average cost will happen only if one firm controls the industry's total production, i.e. the minimum efficient scale (MES) is more or less equivalent to, or larger than the total market size. Whilst some market areas may be able to support several ports and airports of minimum efficient scale, in smaller geographical markets this may not be possible. Even where it is possible, planning constraints and environmental considerations may prevent the development of competing facilities, giving existing ports and airports monopoly power.

So in the case of seaports and airports, one can raise the question whether they always work in a competitive environment, or not. In the latter case there might be a need for regulation. If so, should regulation be applied to all ports and/or airports? If we regulate, do we opt for structural regulation focusing on market structure, or conduct regulation seeking to influence the behaviour of firms?

In what follows we start from the heterogeneous character of seaports and airports, stressing the relationship between a high number of different actors. No two seaports or airports are physically and economically the same. We refer to already existing forms of regulation, the market evolution towards less and bigger market players and/or alliances and the consequences for more regulation.

[26] E. Van de Voorde and W. Winkelmans, (2002), "A general introduction to port competition and management", in M. Huybrechts, H. Meersman, E. Van de Voorde, E. Van Hooydonk, A. Verbeke and W. Winkelmans (eds.), *Port Competitiveness. An economic and legal analysis of the factors determining the competitiveness of seaports*, De Boeck Ltd., Antwerp, pp. 1-16.

7.2. The seaport: a heterogeneous mix of actors [27]

Major port activities deal with the physical throughput of goods and passengers. In the course of time, this throughput function has become separated with the emergence of new, specialized functions such as among others forwarding and agency, and a distribution function. Each port in itself became a chain of consecutive links, while the port as a whole also constitutes a link within a global logistics chain. Over time, the relative importance of the various links has clearly changed, due to significant efficiency-increasing technological developments.

The port that contributes to the cheapest logistics chain is, in theory at least, most likely to be called at. The ultimate decision process of the port user is whether the port considered does offer advantages compared to other ports serving the same hinterland. Or does the port offer sufficient advantages in order to be considered as an additional port of call for an existing or yet-to-be-established liner or feeder service.

Port operations involve a great many players, both at management level and at operational level. These actors may be united in a single company, as is the case in some privately owned ports in the UK, or they may constitute a mixture of firms and authorities within the port. A typical European port consists of more than a port authority and a terminal operating company. Many different players and decision makers are active within ports, including shipping companies, port authorities, goods handlers, agents, industrial and production companies, etc. Moreover, activities of actors are changing, as do the mutual relationships between their companies.

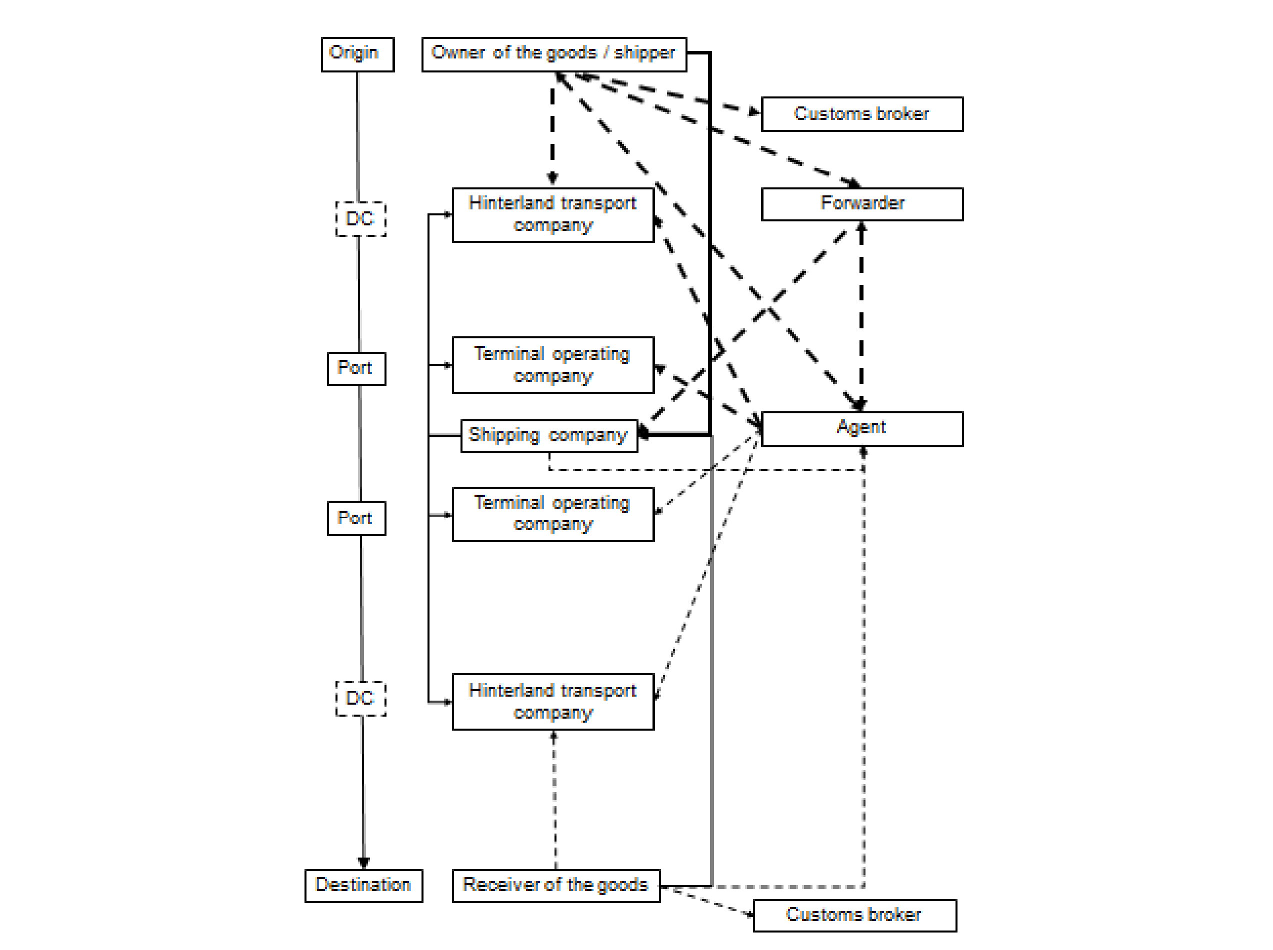
The question arises which player in the chain takes which decisions. Earlier studies[28] have shown very clearly that certain players are particularly influential, namely the owner and/or shipper of the goods, the forwarder and the shipping companies. Terminal Operating Companies (TOCs), on the other hand, are highly dependent upon the decisions taken by these three other parties, even though TOCs themselves are also required to make long-term commitments: investment in superstructure (e.g. storage capacity) and terminal infrastructure (e.g. cranes, straddle carriers,…).

Figure 1 gives the structure of a typical maritime supply chain, including the potential relationships between actors. The possibility of monopoly power, and of the need for regulation, may arise at any point in the chain. Cost minimization, for example, is important for every actor in that chain, but clearly a shipping company has greater scope than some other actors for restricting costs while being able to maintain a price level that guarantees a wide profit margin. Ultimately, the (new) owner of the goods will have to pay the bill.

[27] For this section we explicitly refer to H. Meersman, E. Van de Voorde and T. Vanelslander, (2009),” The Economic Fabric of Ports”, in H. Meersman, E. Van de Voorde and T. Vanelslander (eds.), *Future Challenges for the Port and Shipping Sector*, Informa, London, pp. 89-107; and to H. Meersman, E. Van de Voorde and T. Vanelslander, (2010), “Port Competition Revisited”, *Review of Business and Economics*, 55(2), 210-232.

[28] See e.g. F. Coppens, F. Lagneaux, H. Meersman, N. Sellekaerts, E. Van de Voorde, G. Van Gastel, T. Van Elslander and A. Verhetsel, (2007), *Economic impact of port activity: a disaggregate analysis. The case of Antwerp*, Working paper document n° 110, National Bank of Belgium, Brussels, p 79.

Figure 1. Structure of a maritime supply chain



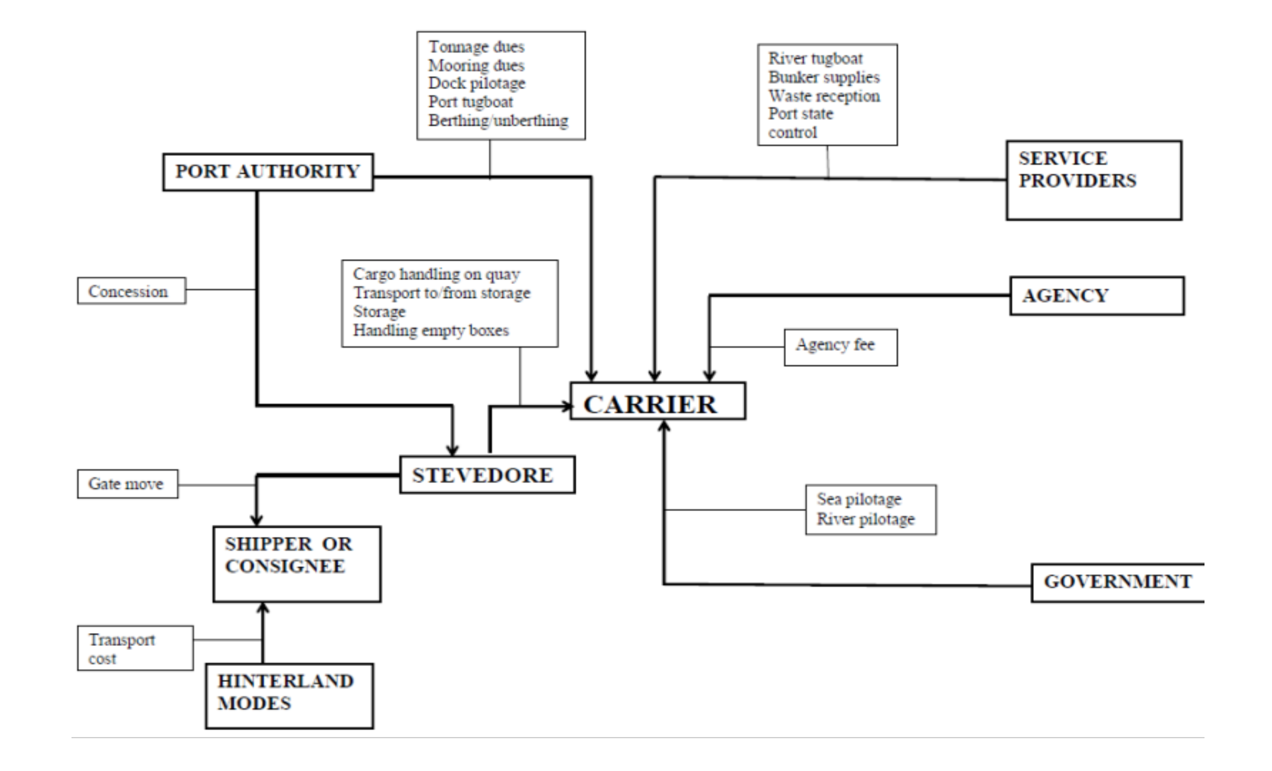
Source: H. Meersman, E. Van de Voorde and T. Vanelslander, 2010.[29]

Within that maritime logistics chain we concentrate in a next step on the pricing and payment of port bills (see figure 2). By quantifying these relationships between actors one gets an idea about the relative power between actors.[30] This helps to investigate whether some actors and/or activities are abusing their (potential) dominant position. Therefore, there is a need to understand the relative importance and the negotiating and market power of the various port actors. One needs to know what their mutual relationships are, whether some actors are dominant, and whether or not they have financial stakes in one another. In a situation where a TOC depend on only one big client (ship owner), negotiations concerning prices and quality of services can/will be different from a situation where that same TOC has several clients, without a dominant one. At the same time the driving forces behind the various forms of co-operation that manifest themselves (mergers, alliances, participation, etc.) should made clear.

[29] Meersman, Hilde, Van de Voorde, Eddy and Thierry Vanelslander, 2010, Port Competition Revisited, Review of Business and Economics, 55(2), 221.

[30] This has been done for the port of Antwerp. Drawing on empirical data, a quantified understanding has been reached of the mutual relationships among the port actors, and their mutual dependency. The influence of a port actor (Antwerp in this case) on its customer (in this case another port actor) is measured by means of forward linkages. The influence an Antwerp port actor has on its suppliers – in this case an Antwerp port actor – is defined by decomposed backward linkages. For Antwerp one thing was very clear: the importance of forwarders. See Meersman, Hilde, Van de Voorde, Eddy and Thierry Vanelslander, 2009, The Economic Fabric of Ports, in: Meersman, Hilde, Van de Voorde, Eddy and Thierry Vanelslander (eds.), Future Challenges for the Port and Shipping Sector, Informa, London, pp. 89-107.

Figure 2. Pricing and payment of port bills



Source: Meersman, Van de Voorde and Vanelslander.

The average European port as a physical entity is managed by a port authority (PA) which, in turn, is usually partly or wholly controlled and/or regulated by a higher (public or administrative) authority. The PA's major remaining trump card is the allocation of concessions within the port perimeter to terminal operators and industrial and service companies. Normal auction procedures are applied and published. If this process is handled effectively and there is real competition for the concessions then that should ensure efficiency of the various actors within the port. Complaints can result in court procedures.

The major clients of a port are the ship-owners and their vessels that call at the port. But the commercial relations are not between the ship-owner and the port authority, but between the ship-owner (or its agent) at one side, the terminal operating company and the forwarder at the other. So, we concentrate first of all on the port authority, the ship-owners and the terminal operating company as three of the major port actors.

7.2.1. Port authorities

Port authorities decide about the port dues, the price a vessel has to pay to call at the port.[31] Shipping regulations, contributing to the safety, apply in any port. The harbour master enforces international, national and local laws and regulations. International regulations are issued by IMO, the International Maritime Organization (e.g. SOLAS).

[31] See H. Meersman, S. Pettersen Strandenes and E. Van de Voorde, (2015), "Port Pricing", in C. Nash (editor), *Handbook of Research Methods and Applications in Transport Economics and Policy*, Edward Elgar Publishing, Cheltenham, pp. 329-341.

In some countries or regions, a port commissioner supervises whether port authority decisions do correspond with the formal regulation.

Recently we see port authorities deciding to co-operate with other port authorities. An example is the merger of the ports of Ghent (Belgium) and Sealand port (The Netherlands).

7.2.2. Ship owners

In the past two decades, the container port industry has seen three waves of consolidation. The last one being the most significant. Eight of the top 20 players have disappeared from the market in the last two years. This seems to be the consequence of supply and demand out of balance, resulting in low rates and in a market-share rate war. Table 1 shows some of those deals, all of them subject to regulating approval in various jurisdictions.

Table 1. Recent scale increase in the container liner business

Year	Buyer	Company acquired	Amount
2016	CMA CGM	Neptune Orient Lines (NOL)	2.4 billion $
2017	Maersk	Hamburg Süd	4 billion $
2017	Cosco	Orient Overseas Limited (OOIL)	6.3 billion $
2017	Ocean Network Express (ONE)	Merger of 3 Japanese companies: NYK, MOL and 'K' Line	n.a.

Source: *own composition, based on the websites of the companies*.

After this move, there will be just seven global carriers with a total capacity of over 1 million TEU. The container shipping industry consolidated into three huge alliances: Ocean Alliance, The Alliance, and 2M. Industry consolidation should raise concerns among the customers of the liners, e.g. the cargo shippers. Can we be sure that the market will remain highly competitive? Anti-trust regulators do not intervene yet. No individual carrier will commend more than 20% of overall market volumes. Consolidation probably will support more stable rate levels. Container liners hope to get more profitability and promise their customers less volatile freight rates and better schedule reliability.

7.2.3. Terminal operating companies

Container handling evolved over time from small stevedores, competing with each other at local port level, towards integrated global container terminal operating companies, TOCs. This as the result of the fact that local stevedores could not follow the need for capital anymore in order to cope with capital investments in new facilities and container cranes. At the same time, the local companies had to negotiate with liner companies working under the umbrella of strategic alliances and/or companies recently consolidated or merged.

Table 2 gives a forecast which terminal operator will be number one by 2020 as measured by capacity. The table shows that the merger of China Shipping and Cosco is likely to see it become the clear world leader in capacity terms. But it is clear that we get global and international terminal operators, focusing more on M&A opportunities and less in greenfield projects.

Table 2. Forecast global/international terminal operator capacity ranking, 2020

Operator	Capacity rank	
	2020	Current
Cosco-China Shipping	1st	4th and 8th
APM Terminals	2nd	2nd
PSA International	3rd	3rd
Hutchison Port Holdings	4th	1st
DP World	5th	5th
Terminal Investment Ltd	6th	6th
CMA CGM	7th	9th

Source: Drewry Maritime Research.

The consolidation of ownership in the terminal operating business should be considered a natural response to the increasing size of liner alliances. The consequence can be that future negotiations between liner companies and terminal operators should be treated as a situation of bilateral oligopolies. Anyway, markets are still contestable. Entry barriers seem to be low, knowing that with Yilport Holdings, a new entrant joined business (and purchased the Portuguese group Tertir).

7.2.4. Pilots and towage companies

For some ports not located at sea (e.g. Hamburg, Antwerp,…), assistance by pilots and towage is needed. We often see that those professions and companies do have a quasi-monopoly concerning the supply of their services.

7.3. The changing port game

Liner shipping consolidation, resulting in fewer ship calls with bigger vessels and higher peak handling numbers, also reduces the number of individual customers. However, those customers increase their scale of cargo and become more powerful. They expect for instance that terminals are able to turn round an 18,000 TEU ship as quickly as a 12,000 TEU vessel. That means that container ports and their local actors are being confronted with downwards pressure on tariffs and upward pressure on costs. At the level of port authorities, this can lead to more co-operation, for instance collaboration between neighbouring ports.

That means that the competition game is changing. It starts with less container liners, with lower profit margins, deciding to order bigger vessels. As a reaction, port authorities need to invest in lengthening of quays and deepening of berths and other investments to meet the requirements of ultra-large containerships. Terminal operators have to invest in wider and faster cranes. A much higher peak demand means that there are more periods when staff is inactive. Moreover, a lot more should be done on the landside, such as automatic gates and trucking processes.

As a result, funding becomes a lot more difficult. Governments tried already earlier to diminish their involvement in port infrastructure investments. But now also infrastructure funds and pension funds raise questions whether potential returns in port terminals will be as sure as they used to be. Also in a port environment some actors may make mistakes, for instance by investing too much, or not enough. Or one proposes unsuccessful products. Or prices may be fixed at wrong levels. Till now the port system believes that a correction of such mistakes will follow immediately, for instance due to a loss of market share.

Is that reasoning correct? Or does this evolution leads to a need for regulation? Competition rules concerning the container liners were drawn up in a very different era from now. Container liners were far less powerful than today. Co-operation agreements just as for instance vessel-sharing agreements were likely to be limited to a specific geographical area. In the meantime, the container shipping industry consolidated. Should they be regulated differently from more traditional consortia agreements?

The current EU regulation concerning consortia agreements is not suitable for the global alliances. Rules need to be updated to make sure global alliances do not abuse their dominant positions. The European Commission will start in 2018 to review the rules governing consortia (including the consortia block exemption). In the U.S., one also considers whether now is the time to adjust the maritime antitrust regime, in view of the current round of mergers and acquisition activity, the collapse of Hanjing Shipping, and the concentration of the global trades into the hands of three alliances.

This could become a process that may be the catalyst for regulatory reform, in the sense that the potential consequences of dominant groups negotiating with other actors like port authorities and terminal operating companies, should be monitored on a continuous basis. At the same time, one should make clear that another customer being the owner of the goods and/or his forwarder, does not become the victim of this changed power play among port actors and users.

However, these developments may suggest a need to regulate the shipping lines themselves rather than ports. Rising power of the shipping lines should increase the pressure on ports to be efficient.

7.4. The actual use of port benchmarking

What is the actual use of benchmarking by regulators? Are there any benchmarking studies available that can be used by regulators? Not only regulators, but also port authorities and other port actors do have a need to benchmark their activities and performance. A benchmarking process delivers additional information, also about 'peers', allows learning by example and adopting best practices, and gives an incentive to make performance step changes. At the end it is all learning from others. Indeed a continuous comparison of own performance with that of peers identifies gaps in performance and creates opportunities to improve that performance and to generate additional benefits.

As usual in benchmarking, questions arise as to comparability. Should studies focus on ports with similar markets, or can differences in traffic be allowed for by the techniques used?

It is clear that benchmarking a port as a whole is of limited value. One should benchmark at actor's level. Some of those port actors can have and/or use different benchmarks and indicators, but at some level they also have a joint interest in performance improvements. Even if the focus of the benchmarking activity is on costs, it is clear that a number of quality aspects must also be taken into account.

Table 3 gives the example of the relevant concerns of two major clients of a port authority, being a shipping line and a Terminal Operating Company (TOC).

Table 3. Potential benchmark indicators for a shipping line and a TOC

Shipping line	Terminal operator
• Total time to service a vessel • Level of port dues • Level of handling charges • Level of storage charges • Flexibility of berthing (and handling) windows • Availability of feeder services • Access to hinterland infrastructure and modes	• Profitability • Utilization degree of facilities • Dwell time • Flexibility of the operating system • Planning reliability • Distance/time cargo moves to the stacking zone

Source: own composition.

It is clear that, even while using a different set of benchmarking indicators, both major port actors do have common interest and targets, being a high productivity of the berth, the yard and the loading/unloading and stacking equipment used. A high productivity will depend of a high reliability. That high reliability can for instance be reached by better forecasting the vessel arrival times, the hinterland mode (train, barge, truck,…) arrivals, the resource and yard planning. As important is the way both actors co-operate to handle exceptional incidents, for instance when a vessel is delayed due to weather conditions.

The port literature deals with benchmarking port activity, while often making the distinction between port authorities (PAs) and port Terminal Operating Companies (TOCs). The question remains how and in what way regulators are actually using benchmarking as an instrument. Do they benchmark themselves? Or do they use the results of published benchmarking studies?

7.4.1. Benchmarking Port authorities

Zahran e.a. (2017)[32] analyze port authority efficiency while generating revenues, using Data Envelopment Analysis (DEA). According to the authors there has been no study to focus on PAs and compare, in terms of efficiency, their revenues generation mechanisms. Two different modelling approaches are presented: one that relates revenues to the basic throughput and activity of the port and another that takes into consideration land, infrastructure and labor indicators (cf table 4).

Table 4. Models used by Zahran et al (2017): inputs and output used

Model (1)	Model (2)
Inputs • Number of vessels called • Total throughput (x000 tons) • Number of passengers (x1000) Output • Total revenues (million $)	Inputs • Area of open yards (hectares) • Number of berths • Number of cargo handling equipment Output • Total revenues (million $)

Source: Zahran et al (2017), p. 526.

A sample set of 18 international landlord, multi-activity ports has been used, including Houston (US), Le Havre (France) and Barcelona (Spain). The analysis first discriminates between efficient and non-efficient ports and for the latter it identifies their benchmarks and provides target values for their revenues. In a post DEA-stage the authors examine the relation between efficiency and the size of a port.

A comment to be made is the fact that a landlord port is not in charge of the direct management of the port's facilities. Concessions are one of the few remaining powerful tools, in order to realize the most

[32] S.Z. Zaharan, J. Bin Alam, A.H. Al-Zahrani,Y. Smirlis, S. Papadimitriou and V. Tsioumas, (2017), "Analysis of port authority efficiency using data envelopment analysis", *Maritime Economics & Logistics*, 19(3), 518-537.

efficient utilization of its terminals and land. Also, Goss (1990)[33] stated that revenues by themselves are not a sufficient indicator of efficiency.

Another study of Tovar and Rodriguez-Déniz (2015)[34] states that while the actual estimation of port efficiency has been extensively studied, the existing literature has paid little attention to developing robust methodologies for port classification. The problem of distinguishing between heterogeneity and inefficiency is widely acknowledged in benchmarking, and aggravated when international datasets are used. Out of the literature survey follows that authors using DEA (the vast majority) have tried to solve this problem by splitting the sample into homogeneous groups before the frontier estimation.

The United Nations Conference on Trade and Development (UNCTAD) published a document on port performance, including a section on port benchmarks.[35] An interesting exercise concerns the so-called 'port performance scorecard'. Table 5 gives an overview of the indicators used for the port entity only.

Table 5. Indicators used by UNCTAD

Port entity only	Indicators
Finance	EBITDA/revenue (operating margin)
	Vessel dues/revenue
	Cargo dues/revenue
	Rents/revenue
	Labour/revenue
	Fees and the like/revenue
Human resources	Tons/employee
	Revenue/employee
	EBITDA/employee
	Labour cost/employee
	Training costs/wages
Vessel operations	Average waiting time (hours)
	Average overall vessel length per vessel (m)
	Average draft per vessel (m)
	Average gross tonnage per vessel
Cargo operations	Average tonnage per arrival - all
	Tons per working hour, dry or solid bulk
	Box per hour, containers
	Twenty-foot equivalent unit dwell time (days)
	Tons per hour, liquid bulk
	Tons per hectare - all
	Tons per berth metre - all

Source: UNCTAD (2016).

This type of indicators can be collected on a yearly basis. Comparisons can be made for one port over time, and between ports.

[33] R. Goss, (1990), "Economic policies and seaports: Are port authorities necessary?", *Maritime Policy and Management*, 17(4), 257-271.

[34] B. Tovar and H. Rodriguez-Déniz, (2015), "Classifying ports and efficiency benchmarking: a review and a frontier-based clustering approach", *Transport Reviews*, 35(3), 378-400. This paper contains an excellent and extensive review of the literature on classification methods for port efficiency. Moreover an approach has been presented that combines stochastic frontier analysis, clustering and self-organized maps. The methodology is applied to a database orf Spanish port authorities.

[35] UNCTAD, (2016), *Port Management Series*, Volume 4: Port Performance. Linking performance indicators to strategic objectives, Unitd Nations Conference on Trade and Development, p 45.

7.4.2. Benchmarking Terminal Operating Companies

Pinto *et al.* (2017)[36] deal with benchmarking operational efficiency of port terminals. In their opinion the operational data of port terminals is often evaluated by operators and scholars, with the purpose of finding characteristics that lead to superior performance, or identifying the most efficient terminal in a sample. Often methodologies are used that do not allow distinguishing between manageable and unmanageable (exogenous) factors, thus often leading to ambiguous results. An example to illustrate: standard high-level indicators, such as throughput and berth occupancy, are also influenced by factors that cannot be manipulated by port managers, such as weather conditions, maritime access channel conditions, external norms and regulations (e.g. work hours) and storage factors of bulk cargoes.[37] The authors propose a methodology to overcome this limitation, based on the breakdown of the high-level Overall Equipment Effectiveness (OEE) indicator into a set of indicators, each addressing either manageable or unmanageable factors. Based on this set it is possible to define achievable efficiency targets for each terminal.

Within the literature most port terminal efficiency analyses use efficiency frontier methods, using either parametric or non-parametric models.[38] A parametric model relates inputs (time, equipment, costs,…) and output. After defining the parameters of the function, the performance of each terminal can be evaluated and compared to its expected value to assess efficiency.[39] Non-parametric models identify the most efficient decision making units of a sample, and evaluate the efficiency gap between top performers and others.[40] Most authors here apply the Data Envelopment Analysis (DEA) model. Pinto (2017)[41] states that sophisticated analyses, using parametric and non-parametric models for defining efficiency frontiers, often fall into the same pitfall, being the contamination of exogenous effects in efficiency calculations.

Port performance and efficiency can also be measured by using a set of performance indicators that can be benchmarked or compared to theoretical values. It looks simple, but this approach allows the analysis of specific processes of a system. Table 6 gives a list of the operational performance indicators commonly used in the port industry.

The indicators listed in Table 6 are applicable to the vast majority of ports and terminals, regardless of the types of cargoes and facilities used.

---

[36] M.M.O. Pinto, D.J.K. Goldberg and J.S.L. Cardoso, (2017), "Benchmarking operational efficiency of port terminals using the OEE indicator", *Maritime Economics & Logistics*, 19(3), 504-517.
[37] Pinto et al, 2017, p. 305.
[38] See for an excellent literature review Pinto et al (2017). See also K. Bichou, (2006), "Review of port performance approaches and a supply chain framework to port performance benchmarking", *Research in Transportation Economics*, 17(6), 567-598; and M.M. Gonzales and L. Trujillo, (2009), "Efficiency measurement in the port industry: a survey of the empirical evidence", *Journal of Transport economics and Policy*, 43(2), pp. 157-192.
[39] See for instance C.P. Barros, (2005), "Decomposing growth in Portuguese seaports: a frontier cost approach", *Maritime Economics & Logistics*, 7(4), 297-315.
[40] See for instance K. Cullinane, D.-W. Song and T.-F. Wang, (2004), "An application of DEA windows analysis to container port production efficiency", *Review of Network Economics*, 3(2), 184-206.
[41] Pinto et al, 2017, p. 515.

Table 6. Operational performance indicators commonly used in the port industry

Indicator	Description
Throughput	Total amount of cargo handled in tons (or containers) in a given period
Load size/Tonnage per ship	Average amount of cargo handled per ship call
Ships arrival rate	Average number of ship calls per day
Berth occupancy rate	Percentage of total time with ship alongside berth
Storage occupancy rate	Ratio between total cargo stored (in average) and the rated capacity of the facility
Total time in port	Average time spent by the ships in the port area
Queue time	Average time spent by the ships in the queue
Gross (net) loading/unloading rate	Average amount of cargo loaded/unloaded between the arrival and the departure of ships (excluding non-operational times)

Source: Pinto et al, 2017, p. 507.

7.5. Ports: a need for yardstick competition?

The former sections clearly show a potential danger that in some specific relations an actor starts to abuse its dominant position. In such a market environment yardstick competition can offer a solution. It is a specific form of regulation which requires collecting information on cost conditions of comparable companies.[42] Any remuneration of the regulated undertaking will be linked to the results of the other companies, in order to give incentives to become much more efficient.

In order to prepare the application of yardstick competition, one can gather information following a checklist:

1. is there enough competition, at each level, from shipping companies to TOCs, pilots, towage companies and other service providers? That means the need for information concerning negotiation processes and results, besides quantitative information concerning market shares, for all actors involved;
2. is there any risk of collusion between actors, both horizontally (e.g. between TOCs) and vertically (between a shipping line and a TOC)? What is the effect of more concentration in the liner business on port competition, e.g. in the case where one liner company dominates the port throughput?
3. are there any barriers to entry and/or exit in the port market (with a reference to contestability theory)? Does the concentration in the TOC-business create barriers (cf. those ports with only one TOC)?

Within the maritime logistics chain we more and more get forms of co-operation, from alliances between shipping companies to take-overs and mergers in a port context.[43] [44] In some cases one risks to get a situation where companies start to collude, horizontally and/or vertically. In some ports there exist joint ventures between the major shipping line and a TOC. That can lead to so-called ‘dedicated terminals’ of ‘virtual dedicated terminals’. In what way is such an agreement blocking normal competitive mechanisms? Collusion between companies can eliminate the incentive to present competitive bids and therefore the

[42] G. Bruzzone, (2017), *The role of benchmarking in competition policy and how to use yardstick competition in a procompetitive regulatory strategy*, mimeo, Meeting of the Advisory Board, Turin.
[43] T. Heaver, H. Meersman, F. Moglia and E. Van de Voorde, (2000), “Do Mergers and Alliances Influence European Shipping and Port Competition?”, *Maritime Policy and Management*, 27 (4), 363-373.
[44] T. Heaver, H. Meersman and E. Van de Voorde, (2001), “Co-operation and Competition in International Container Transport: Strategies for Ports”, *Maritime Policy and Management*, 28 (3), 293-305.

information and efficiency enhancing role of tenders. Fighting collusion remains a crucial task, to be started by checking and comparing the tenders.

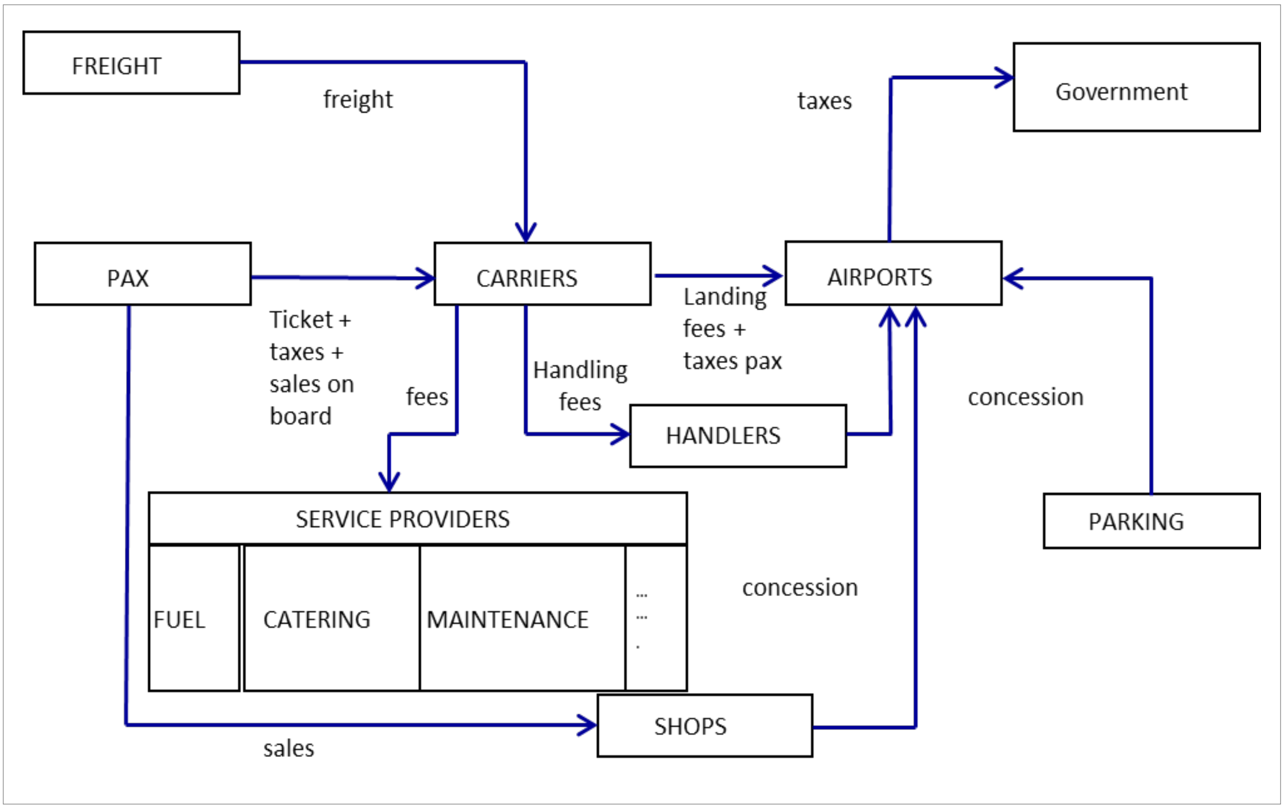
Benchmarking also requires the analysis of costs, including the assessment of economies of scale and scope. This is important at each level of port activities, for instance while tendering a new concession for a TOC. A port authority of a big port, being aware that oversized concessions should be avoided, will always try to opt for several concessionaires. That creates the possibility to compare the performance of all concessionaires and to develop forms of competition by comparison. An excellent example are the concessions within the port of Antwerp to PSA and DP World at the Deurganckdock. Those concession agreements included performance targets. If those contracted targets are not reached, the port authority applies a penalty.

Another issue is that some port authorities are fully or partially controlled by public authorities, e.g. a national or regional government, or a city council. One then may use benchmarking in order to set the proper incentives on the port authority, both in terms of prices (e.g. 'port dues') and/or quality.

7.6. The case of airports

There is quite a lot of similarity between seaports and airports, be it that seaports in the majority of cases deal with freight, while most airports are active in both the passenger and freight market. Also airports can be considered a chain of consecutive links, while the airport as a whole also constitutes a link within a global passenger and/or logistics chain. Within the airport, a large number of actors are involved, each of which pursues its own objectives, which gives rise to a considerable degree of heterogeneity. Figure 2 gives an overview of these relationships, based on the pricing and payment of airport bills.

Figure 2. Pricing and payment of airport bills



Source: Rosario Macàrio and Eddy Van de Voorde.

There seems to be an important difference between the seaport and airport sector. The potential power of seaport authorities seems to be much lower than the potential power of airport authorities. Although airlines also work together in alliances, their market power seems to be less than the one of liner conferences. The reason could be 'slot allocation'.

In some countries, the airport authority is regulated, whilst the other airport actors are not. In what follows, we will highlight the most important conclusions of a report on the economic regulation of airports, with an application to the Brussels Airport Company.[45] The study considered the following aspects: the length of the regulation period; (adjusted) single till versus dual till; tuning of tariffs to those applied at a set of reference airports versus the use of a financial model; the distinction between regulated and non-regulated activities (i.e. the level of economic regulation) and the role of the economic regulator. The purpose of the study was to investigate which type of regulation would be most appropriate for *Brussels Airport Company*, taking into account the concepts of natural monopoly and market power.

What arguments might justify economic regulation in the airport business? The customary answer, as mentioned before, is the occurrence of market failure. A regulatory intervention for the purpose of introducing a correction is generally regarded as efficient if the costs of intervening will be lower than the cost incurred through market failure. Again, though, no generalisation is possible. The market power of an airport will, for example, vary according to the market segment. Larger airports, most of them acting as hubs, may be expected to have some market power over so-called networked airlines, which aim expressly at achieving maximum benefits of scale and scope. There is minimum or even no market power vis-à-vis low-cost airlines, companies specialising in point-to-point services, and charter companies. Invariably, though, the market power of an airport will be largely determined by the presence of nearby competing airports. The likelihood of substitution is crucial, so that a case-by-case assessment is called for.

All this obviously entirely re-contextualises the notion of 'regulation' in an airport environment. There can be no doubt that some form of monitoring of airports remains necessary. After all, European Directive 2009/12/EC regarding airport charges (in effect since 15 March 2009) does prescribe explicitly that there should be an independent supervisory authority that safeguards the principles underlying the setting of airport charges (Article 11) and, more generally, to supervise the airports concerned (Article 12). In this context, price cap regulation is a commonly used instrument to discipline private companies with market power, like airport authorities.

7.7. The actual use of airport benchmarking

From a technical point of view one can expect airport benchmarking to be comparable to seaport benchmarking. However, there is one big difference. Quite a lot of airports are privately owned, or in the process to get privatized. The result is that financial indicators get emphasized, together with indicators measuring capacity utilization (e. g. runway slots, parking gates,…) and efficiency.

The literature provides quite a number of benchmarking studies. A very important one is the ATRS Global Airport Benchmarking Report[46], measuring and comparing the performance of several important aspects of airport operations: productivity and efficiency, unit costs and cost competitiveness, financial results and airport charges. Moreover the report also examines the relationships between various performance

[45] This is based on F. Kupfer, H. Meersman, T. Pauwels, E. Struyf, E. Van de Voorde, T. Vanelslander, (2013), "Economic regulation of airports: The case of Brussels Airport Company", *Case Studies on Transport Policy*, 1 (1), 27-34.

[46] ATRS, (2017), *The ATRS Global Airport Benchmarking Report*, Air Transport Research Society, Embry-Riddle Aeronautical University, Daytona Beach, Florida, USA. This benchmarking project was initiated in 2000 at the University of British Columbia. It is now hosted at the College of Business at Embry-Riddle Aeronautical University.

measures and airport characteristics as well as management strategies in order to provide a better understanding of observed differences in airport performance. Being a yearly report, the 2017 version includes 206 airports and 24 airport groups of various sizes and ownership forms in Asia Pacific, Europe and North America.

The Civil Aviation Authority (CAA) published in 2013 an Airport Operating Expenditures Benchmarking report.[47] This report provides for 2012 a review and assessment of the airport Opex benchmarking indicators available to the CAA. This evidence has been used by the CAA to develop benchmarks of Heathrow, Gatwick and Stansted against comparators based on publicly available data.

The CAA has reviewed several benchmarking studies, each of which has used different data and methods to assess the relative level of efficiency of each airport. The ATRS benchmarking already has been discussed. The Leigh Fisher benchmarking study provides benchmark comparisons of airports across key matrices including opex per passenger. The study is based on financial reporting for 2009, and includes 50 airports (including Heathrow and Gatwick, but not Stansted).

A study of Booz & Company, commissioned by British Airport Authorities – BAA, benchmarks the opex performance of Heathrow against a variety of European comparators, based on 2011 financial account data.[48] Interesting is that in addition to the partial productivity metrics, the study also use an econometric model to account for differences associated with uncontrollable factors. This Is based on a 'residual' approach whereby costs are separated into 'Inherent', 'Structural', 'Systematic' and 'Realized' costs, which can then be defined as controllable and non-controllable. The residual productivity analysis, having taken account of these factors, is then used to estimate the relative efficiency gap of Heathrow against the specific comparators.

Building on the evidence and methodologies of the above mentioned studies, the CAA has undertaken additional analysis on the relative performance of Heathrow, Gatwick and Stansted against comparators. Data has been collected for a sample of 16 airports from 2000 to 2011, subdivided in two sub-samples. Two partial productivity metrics have been used to examine the relative performance of the airports. First the ordinary opex per passenger (not been adjusted to take account of differences in airport activities). Secondly the adjusted opex per passenger, intended to adjust for airport activities to provide a more consistent estimate of each airport's core operating costs for comparative purposes.

7.8. Concluding remarks

It is clear that both seaports and airports have a lot in common. Both are very competitive sectors, but it remains the case that geographical factors and constraints on development of new facilities may lead to particular ports and airports having monopoly power. Both are very complex environments, with a multitude of actors influencing each other's behaviour and results. That means that there always is the possibility that some actors start to increase their market power and that we get a risk that market power gets abused. To a considerable extent, this risk may be reduced by competitive tendering for concessions to provide these services, provided that this process is handled effectively and that real competition emerges.

[47] CAA, (2013), *CAA Airport Operating Expenditure Benchmarking Report 2012*, Civil Aviation Authority, CAP 1060, London.

[48] The following airports are included: Amsterdam (AMS) (Group); Athens (ATH): Birmingham (BHX); Paris Charles de Gaulle (CDG) (Group); Copenhagen (CPH); Dublin (DUB) (Group); Edinburgh (EDI); Frankfurt (FRA); London Gatwick (LGW); Manchester (MAN); Munich (MUC) and Zurich (ZRH).

In short, then, ports and airports raise difficult questions regarding whether regulation is needed and if so where and of what form. Benchmarking, both of ports and airports as a whole and of individual functions within them, may play a useful role in the first place simply as a means of monitoring whether there appear to be problems of inefficiency and of abuse of monopoly power. Where problems are identified, then there may be the need for regulation, based on benchmarking and yardstick competition.

Chapter references

ATRS, (2017), *The ATRS Global Airport Benchmarking Report*, Air Transport Research Society, Embry-Riddle Aeronautical University, Daytona Beach, Florida, USA. This benchmarking project was initiated in 2000 at the University of British Columbia. It is now hosted at the College of Business at Embry-Riddle Aeronautical University.

Barros, C.P., (2005), "Decomposing growth in Portuguese seaports: a frontier cost approach", *Maritime Economics & Logistics*, 7 (4), pp. 297-315.

Bichou, K., (2006), "Review of port performance approaches and a supply chain framework to port performance benchmarking", *Research in Transportation Economics*, 17 (6), 567-598; and Gonzales, M.M., Trujillo, L., (2009), "Efficiency measurement in the port industry: a survey of the empirical evidence", *Journal of Transport economics and Policy*, 43 (2), 157-192.

Bruzzone, G., (2017), *The role of benchmarking in competition policy and how to use yardstick competition in a procompetitive regulatory strategy*, Meeting of the Advisory Board, Turin.

CAA, (2013)*, CAA Airport Operating Expenditure Benchmarking Report 2012*, Civil Aviation Authority, CAP 1060, London.

Coppens, F., Lagneaux, F., Meersman, H., Sellekaerts, N., Van de Voorde, E., Van Gastel, G., Van Elslander, T., Verhetsel, A., (2007), *Economic impact of port activity: a disaggregate analysis. The case of Antwerp*, Working paper document n° 110, National Bank of Belgium, Brussels, 79.

Cullinane, K., Song, D.-W., Wang, T.-F., (2004), "An application of DEA windows analysis to container port production efficiency", *Review of Network Economics*, 3 (2), 184-206. Pinto et al, 2017, p. 515.

Goss, R., (1990), "Economic policies and seaports: Are port authorities necessary?", *Maritime Policy and Management*, 17 (4), 257-271.

Heaver, T., Meersman, H., Moglia, F., Van de Voorde, E., (2000), "Do Mergers and Alliances Influence European Shipping and Port Competition?", *Maritime Policy and Management*, 27 (4), 363-373.

Heaver, T., Meersman, H., Van de Voorde, E., (2001), "Co-operation and Competition in International Container Transport: Strategies for Ports", *Maritime Policy and Management*, 28 (3), 293-305.

Kupfer, F., Meersman, H., Pauwels, T., Struyf, E., Van de Voorde, E., Vanelslander, T., (2013), "Economic regulation of airports: The case of Brussels Airport Company", *Case Studies on Transport Policy*, 1 (1), 27-34.

Meersman, H., Van de Voorde, E., Vanelslander, T., (2009), "The Economic Fabric of Ports", in Meersman, H., Van de Voorde, E., Vanelslander, T., (eds.), *Future Challenges for the Port and Shipping Sector*, Informa, London, 89-107; and to Meersman, H., Van de Voorde, E., Vanelslander, T., (2010), "Port Competition Revisited"*, Review of Business and Economics*, 55 (2), 210-232.

Meersman, H., Van de Voorde, E., Vanelslander, T. (2009), "The Economic Fabric of Ports", inMeersman, H., Van de Voorde, E., Vanelslander, T., (eds.), *Future Challenges for the Port and Shipping Sector*, Informa, London, 89-107.

Meersman, H., Van de Voorde, E., Vanelslander, T., (2010), "Port Competition Revisited", *Review of Business and Economics*, 55 (2), 221.

Pinto, M.M.O., Goldberg, D.J.K., Cardoso, J.S.L., (2017), "Benchmarking operational efficiency of port terminals using the OEE indicator", *Maritime Economics & Logistics*, 19 (3), 504-517. Pinto et al, 2017, p. 305.

Tovar, B., Rodriguez-Déniz, H., (2015), "Classifying ports and efficiency benchmarking: a review and a frontier-based clustering approach", *Transport Reviews*, 35 (3), 378-400. This paper contains an excellent and extensive review of the literature on classification methods for port efficiency. Moreover an approach has been presented that combines stochastic frontier analysis, clustering and self-organized maps. The methodology is applied to a database of Spanish port authorities.

UNCTAD, (2016), *Port Management Series, Volume 4: Port Performance: Linking performance indicators to strategic objectives*, Unitd Nations Conference on Trade and Development, 45 p.

Van de Voorde, E., Winkelmans, W., (2002), "A general introduction to port competition and management", in Huybrechts, M., Meersman, H., Van de Voorde, E., Van Hooydonk, E., Verbeke, A., Winkelmans, W., (eds.), *Port Competitiveness. An economic and legal analysis of the factors determining the competitiveness of seaports*, De Boeck Ltd., Antwerp, pp. 1-16.

Zaharan, S.Z., Bin Alam, J., Al-Zahrani, A.H., Smirlis, Y., Papadimitriou, S., Tsioumas, V., (2017), "Analysis of port authority efficiency using data envelopment analysis", *Maritime Economics & Logistics*, 19 (3), 518-537.

8. THE RELEVANCE OF BENCHMARKING AS A POLICY INSTRUMENT IN THE ITALIAN CONTEXT

Ginevra Bruzzone, Giovanni Fraquelli e Marco Ponti

8.1. Looking at Italy and at the tasks of the Regulatory Authority

As pointed out in the Introduction, benchmarking is one of the instruments available to policy makers to improve efficiency and promote an effective use of public resources. In the absence of effective competition, either in the market or for the market, incentive-wise regulatory schemes based on competition by comparison may be used to avoid excessive prices/low quality for users and/or unjustified public expenditures, depending on how the funding of the sector is organised. Looking at land-based transport modes, for instance, Table 1 shows the significant impact of rail transport and local public transport on public budgets in Italy.

Table 1. Land-based transport modes - Net Public Expenditures - year 2014 [millions €]

	Expenditures	Fiscal and commercial revenues	Net public expenditures
Railways and related modes	11.913	3.233	8.680
Local public transport and internal ship services	7.371		7.371
Ports	1.939	1.363	576
Road mode	16.139	55.540	-39.401
Total	37.362	60.136	-22.774

Source: Marco Ponti, Francesco Ramella (2018), *Trasporti: conoscere per decidere*, Egea.

Benchmarking supports public policy in several ways: it can be used to improve the information basis of public decision makers, to set incentive targets so as to increase productive efficiency with respect to cost-based models, to define the optimal dimension of service areas, to assess the pros and cons of vertical and horizontal integration.

Thus, benchmarking is relevant for many of the institutional tasks assigned by Italian legislation to the Transport Regulatory Authority (ART). According to Art. 37 of Decree Law no. 201/2011, as modified by Legislative Decree no. 50/2017, these tasks include:

a. ensuring fair and non-discriminatory *access conditions to rail, port, airport and highway infrastructures*, by methods encouraging competition, productive efficiency and cost containment for users, undertakings and consumers;

b. defining, when justified by competition conditions in local and national markets for transport services, the *criteria for the establishment of tariffs*, charges, tolls, ensuring the economic equilibrium of regulated undertakings, productive efficiency and cost containment for users, undertakings, consumers;

c. establishing *minimum quality conditions* for national and local transport services subject to public service obligations (PSO), on the basis of demand and supply conditions in the relevant geographical area;

d. for public transport services provided on an exclusive basis, defining *schemes for tender calls and for service contracts*; establishing objective criteria for exceptions to the principle whereby the contract should be divided into *small lots* (taking into account actual and potential demand, economies of scale, integration of services and other criteria established by the relevant legislation); for rail transport, verifying that the tender calls do not contain discriminatory conditions or impede access of potential competitors to the market, with particular reference to availability of the rolling stock;

e. defining the *schemes of service contracts also for in house contracts or direct awards* in local public transport; both for tenders and for in house or direct awards, establishing the types of efficiency and effectiveness goals that must be pursued by the undertaking, as well as the objectives of financial equilibrium. For all service contracts, imposing accounting separation between activities provided under public service obligations and other activities;

f. for *highways*, establishing *toll schemes based on price cap regulation*, with the determination of the X productivity index every five years; defining *schemes of contracts* to be included in tender calls for the management or construction of highways and *schemes of tender calls*; defining the *optimal management areas* for toll highways, so as to promote a plural management and encourage competition by comparison;

g. with reference to airports, defining models for the adoption of *airport charges* and fulfilling the supervisory tasks established by the implementing provisions of Directive (EC) 2009/12;

h. with respect to access to the rail infrastructure, defining inter alia the criteria for the *determination of access charges*.

Notably, in enforcing such provisions, ART has the power to establish the criteria for regulatory accounting and may require, if necessary to ensure competition, accounting or functional separation. It may also propose the competent administration to terminate or withdraw concessions, conventions and other contracts, when the conditions established by the law are met.

In the rail sector, further legislative measures have broadened the mission of the Authority. Pursuant to Article 37 of Decree Law no. 1/2012, ART is entrusted with the task of defining the scope of public service obligations (PSO) in the different areas and the way in which PSO should be financed, after hearing the Ministry for Infrastructure and Transport and the relevant regional and local administrations. It should also assess the pros and cons of different degrees of separation between the manager of the infrastructure and train operating companies, in the light of the experience of other EU Member States and of the need to protect regular users of regional rail transport services.

As to data collection, the Authority is empowered to request from the relevant entities any information necessary and proportionate for the fulfilment of its institutional tasks. Thus, within this legislative framework, ART is empowered to collect the relevant data needed to promote efficiency in the transport sector, including by means of benchmarking, and to use them either for the adoption of regulatory measures or in the exercise of its advocacy powers.

The status of independent administrative authority ensures that ART supports public policy by exercising its regulatory powers on the basis of the law and technical expertise and by using its advocacy powers acting at arm’s length with respect to all undertakings, either State-owned or privately owned.

As discussed in the previous chapters, an efficiency enhancing strategy should be based on market studies analysing whether competition in the market, competition for the market or competition by comparison is the most effective solution, taking the interests of users and the impact on public budgets into account. In this respect, both the regulatory authority and the competition authority can provide important support to public decision-makers by using their competition advocacy powers.

The importance of benchmarking as a policy instrument depends on whether, in the light of competition conditions, public policy has to introduce efficiency-enhancing incentives; more generally, benchmarking is useful for the design of public contracts.

In this chapter we look at the Italian picture and at the implications of the analysis carried out in the previous chapters for the different transport modes. For each mode, we will briefly describe the current organization of the sector, identify areas where it is useful to explore the possibility of efficiency improvements (e.g. natural monopolies, exclusive rights, design of tenders, contract design for in house or direct awards) and discuss the relevance of benchmarking within the public policy toolbox.

8.2. Toll highways and roads

8.2.1. Description of the sector

The road and highway network in Italy is around 257 000 kilometres, much less than in France (more than 1 million km), Poland and the UK (more than 400 000 km) but larger than in other European countries.[49]

The management of national roads is entrusted to Anas, a State-owned company currently within the FSI group. For other roads the competence is of regional or local administrations, at the provincial and municipal level. Pursuant to a recent decree,[50] approximately 6250 km of the latter network are being brought within the competence of ANAS ("rientro strade"). The construction and maintenance of roads are usually contracted out, following the rules on public procurement.

As to highways, in 2017 the total length of the network was around 6900 km, almost the double than the one in the United Kingdom (3768) and slightly over than half the ones in France (11599) and Germany (12993) (Eurostat). Around two thirds of the network have 2+2 separated lanes, while there are 1870 km with 3+3 lanes and 131 km at 4+4 lanes.[51]

All the network is managed by means of concessions. The toll component is dominant, with about 6000 km; the remaining 938 kilometres are managed by Anas.

Three quarters of the network were built between the '60 and the '70 of the past century, following law no. 463/1955 ("Legge Romita") and law no.729/1961 ("Piano Zaccagnini"). The largest part of the network was built by Società Autostrade, which at that time was a State-owned enterprise, controlled by the IRI holding.

The traffic on the toll highway network has grown almost fivefold from 1970 to 2010, going from 15 to 83 billion vehicles-km; the trend was reversed from 2010 to 2013, with a reduction of almost 10 per cent of the traffic; since 2013 the traffic started growing again, reaching 82,7 billion vehicles-km in 2017 The heavy vehicle component of the traffic is 23 per cent of the total figure.

The toll highway network is divided among 25 concessionaires (Table 2 and Annex), although the number is much lower (approximately reduced to half) if one considers control shareholdings. Autostrade per l'Italia Spa (ASPI) is the largest operator, with slightly less than 2900 km (3020 km taking into controlled companies). The second largest operator is SIAS Spa (Gavio group) that accounts for around one fifth of the entire network (1423 km), mainly in the north-western regions of Italy.

[49] Eurostat, which however suggests using these data with caution, due to problems of comparability.
[50] Decree of the President of the Council of Ministers of 20 February 2018.
[51] AISCAT, data at 31 December 2017.

In terms of ownership, slightly less than 66 per cent of the network is operated by entirely private companies, about 13 per cent is totally public, and the remaining 23 per cent is operated by public-private companies, mainly under control of the public partners.

The sector is highly concentrated also in some other European Member States: in France the 3 largest concessionaires have a market share of more than 90%, in Spain the largest concessionaire has a market share of more than 60%. Interestingly, concentration as measured by the Herfindahl-Hirschman Index is higher in terms of traffic than in terms of toll revenues or network length.

Table 2. Concessionaires (2017)

Concessionaires	Km	% of total	Traffic [mln of v-km]	% of total	Toll revenue [thousands/€]	% of total
Autostrade per l'Italia S.p.A.	2.854,9	47,6%	47.915,5	55.2%	4.054.125	49,9%
Autostrada del Brennero S.p.A.	314,0	5,23%	4.962,2	5,7%	430.780	5,3%
Consorzio per le Autrostrade Siciliane	298,4	4,97%	1.661,9	1.9%	91.846	1,13%
S.A.T.A.P. S.p.A.	291,9	4,86%	4.308,5	5%	554.899	6,8%
Strada dei Parchi S.p.A.	281,4	4,69%	2.135	2,5%	214.860	2,64%
Autostrada Brescia-Verona-Vicenza-Padova S.p.A.	235,6	3,91%	5.568,3	6,4%	453.077	5,58%
Autovie Venete S.p.A.	210,2	3,5%	2.678	3,1%	257.089	3,16%
Milano Serravalle-Milano Tangenziali S.p.A.	179,1	2,98%	3.140,5	3,6%	280.959	3,46%
Autostrada Torino-Ivrea-Valle d'Aosta S.p.A.	155,8	2,59%	1.846,8	2,1%	170.099	2,09%
Società Autostrada Ligure Toscana S.p.A.	154,9	2,58%	1.925,1	2,2%	244.751	3,01%
Società Autostrada Ligure Toscana – Tronco Autocisa	101,0	1,68%	843,2	1%	128.310	1,58%
Autostrada dei Fiori – Tronco Torino-Savona S.p.A.	130,9	2,18%	945,3	1,1%	91.508	1,12%
Autostrada dei Fiori – Tronco Savona-Ventimiglia	113,3	1,88%	1.263,1	1,5%	202.041	2,4%
Autostrade Centro Padane S.p.A.	88,6	1,47%	983,1	1,1%	82.008	1%
Soc. Italiana per il Traforo Aut.le del Frejus S.p.A.	75,7	1,26%	318,1	0,4%	62.960	0,77%
Concessioni Autostradali Venete - C.A.V. S.p.A.	74,1	1,23%	1.800,3	2,1%	184.255	2,2%
Società Autostrade Valdostane S.p.A.	67,4	1,12%	357	0,4%	86.835	1,06%
Società di progetto Brebemi	62,1	1,03%	378	0,4%	80.740	0,99%
Autostrada Asti-Cuneo S.p.A.	55,7	0,92%	153,4	0,2%	24.524	0,3%
Autostrade Meridionali S.p.A.	51,6	0,86%	1.701,9	2%	103.037	1,26%
Società Autostrada Tirrenica S.p.A.	54,6	0,91%	308,1	0,3%	46.317	0,57%
Tangenziale Esterna di Milano S.p.A.	33,0	0,55%	277,1	0,3%	67.600	0,8%
Raccordo Autostrada Valle d'Aosta S.p.A.	32,4	0,54%	117,1	0,1%	24.050	0,2%
Autostrada Pedemontana Lombarda S.p.A.	41,5	0,69%	196,4	0,2%	32.424	0,3%
Tangenziale di Napoli S.p.A.	20,2	0,33%	927,8	1%	80.717	0,9%
Società italiana Traforo Gran S.Bernardo S.p.A.	12,8	0,22%	7,5	0%	8.574	0,1%
Società italiana Traforo del Monte Bianco S.p.A.	5,8	0,09%	11,78	0%	61.252	0,7%

Source: AISCAT 2018.

8.2.2. Areas which should be monitored within an efficiency enhancing strategy

a. Local and national non-toll roads

Local and national non-toll roads serve three quarters of the Italian traffic of freight and passengers. Formally no regulation is deemed possible, since, if not congested, they are "pure" public goods, and neither natural nor legal monopolies. Nevertheless, since the main national roads often present physical and functional features identical or very close to the ones of toll roads, monitoring their maintenance and investment costs may both substantially enlarge the information base for the entire sector, and as a by-product provide useful information on their efficiency to the different administrative bodies in charge of them.

b. Toll highways

In general, concessions have been awarded both to public and private operators without any competitive and transparent process, on the basis of undisclosed financial plans and for a long duration recently extended for some important concessionaires.

Although reference to a price cap methodology is required since 1996, the approaches followed in practice are heterogeneous and often do not include reference to an efficiency-enhancing X factor attaining to the expected productivity. Several methods entail the ex-ante reimbursement of investment plan, with no penalties if the investments are not carried out. Moreover, there are heavy compensations to be paid in case of termination of the contracts.

A description of six different methods for setting tolls in existing contracts can be found in the illustrative Report of the ART Decision no. 3/2018.[52]

The most frequent approach is "footing the bill", i.e. a "cost-plus" approach with no efficiency-enhancing factors, that assumes all the expenditures fully authorised and controlled by the conceding body. Under this approach concessionaires are incentivised to maximise investments, maintenance and the related unit costs, even in terms of site duration, in the absence of specific penalties for delays. This last feature clashes with the main formal justification for setting a toll on a road, which is the speed allowed to the users of a toll highway compared to a slower non-toll one. Works like adding a lane etc., generally reduce the original design speed of a substantial percentage for a number of years.

A different approach is followed in the case of the largest concession. ASPI has been privatized with a form of regulation that can be defined a "lump sum sale" to the concessionaire, i.e., even if still the investments have to be authorized by the conceding body, the level of profits is set free, in exchange of an up-front payment to the State of about 7 billion €. As well known (Laffont etc.) this approach incentivizes efficiency (the concessionaire has direct advantages in minimising costs), but information asymmetries are left untouched.[53]

As anticipated, only for a small number of concessions the price cap formula includes an X factor aimed to enhance productivity and quality incentivising factors.

[52] http://www.autorita-trasporti.it/wp-content/uploads/2018/09/Relazione-illustrativa-delibera-n.-88_2018.pdf, in particular pp. 13-23. See also F. Ramella (2017), cit., and P. Sestito (2015), *Hearing before the 8th Commission of the Chamber of Deputies within a Sectoral Inquiry on highway concessions*, https://www.bancaditalia.it/pubblicazioni/interventi-vari/int-var-2015/sestito-audizione-110615.pdf.

[53] The price cap in this case depends on the real inflation rate, the remuneration of investments approved in the additional Convention of 1997 and the remuneration of new investments.

A distinct overall regulatory issue concerns the quota of works that are allowed to be assigned "in house" in relation to what has to be tendered out. A sharp debate is still going on about this issue, including the impact in terms of increased transaction costs. The main objective of monitoring this aspect is collecting evidence on efficient benchmarks, since the financial plans of concessionaires apparently are quite differentiated.

Already in 1997-1998, the importance of periodic tendering of concessions and of yardstick competition had been advocated with a view to enhancing efficiency of the sector.[54] As to empirical research, as illustrated in the sectoral chapter on highways, Massiani and Ragazzi (2008) found evidence of inefficiencies and suggested that yardstick competition would be useful to improve incentives, whereas Benfratello, Iozzi and Valbonesi (2009) found strong economies of density and economies of scale at least up to networks of 300 km. More recently, the Italian General Accounting Office (Corte dei Conti), with decision no. 1/2018 G, has opened an investigation on the existing concessions, with a focus on alleged inefficiencies and lack of transparency.

In a broader perspective, a revision of public policy with respect of roads and highways should also reconsider the separation of the sector in toll and free motorways, especially since the larger part of the traffic is served by under- financed and under-planned local roads in the metropolitan areas.

8.2.3. Regulatory issues and the relevance of benchmarking

Currently in Italy the issue of how to ensure a proper public policy for toll highways is being hotly debated, and it is widely acknowledged that in this sector both competition for the market and benchmarking can play a role in ensuring efficiency and effectiveness in the construction, maintenance and management activities.

The regulatory powers originally assigned by the law to ART mainly attained to the design of new awards of concessions. The studies and regulatory measures adopted so far, however, may provide broader guidance for an efficiency enhancing public policy in the sector, in terms of more effective yardstick competition and reduction of information asymmetries.

First of all, as anticipated in the introduction of this chapter, decree law no. 201/2011 (Art. 37(2)(g)) requires ART to ensure fair and non-discriminatory access conditions to highway infrastructures according to methods that encourage competition, productive efficiency and cost-containment to the benefit of users, both undertakings and consumers. In particular, ART is required to define the optimal management areas for toll motorway sections so as to promote plural management thereof and foster competition by comparison.

Pursuant to this provision, ART has conducted proceedings for the definition of optimal management areas, collecting historical data from 23 highway concessionaires and estimating an efficient cost frontier. The outcome is Decision no. 70/2016, stating that significant economies of scale exist up to a length of 180 km, and that beyond 315 km there are no further significant economies of scale. Therefore, the optimal management area for individual concessions should be in the range between 180 and 315 km. Thus, in the award of new concessions or modification of essential elements thereof, the authorities should take this structural efficiency level into account.

[54] See for instance AGCM, AS 135, *Opinion on the extension of highway concessions*, 22 May 1998.

Benchmarks resulting from the estimation of efficient cost frontiers may also play a role in the definition of efficiency enhancing incentives in service contracts with concessionaires, since on the basis of these estimates ART can indicate the X productivity parameter for the price cap formula.

The methodology set out by the Authority in Decision no. 70/2016 and the related estimates have already been used to this aim.

In a first case, in Decision no. 119/2017, ART has defined, upon request of the adjudicating body, schemes for the award of new concessions to be included in specific tender offers as well as toll schemes, based on a price cap methodology and a X productivity factor to be revised every 5 years. The tariff scheme, which innovates on the approach followed by CIPE decision no. 39/2007, identifies the admissible costs and investments, and contains both a component concerning the operation of the infrastructure (including ordinary maintenance costs) and a component to cover construction costs, (for the recovery of new investment costs, including costs possibly incurred in case of takeover by a new concessionaire).

Subsequently, the same approach has been followed to outline tariff schemes, in the absence of a tender, for the direct awards of a toll highway concession (Decision no. 73/2018).

In its 2018 Annual Report, ART argues that empirical evidence and benchmarking may provide important guidance for a proper design of future public policies in the highway sector in Italy. In particular, it stresses that so far the tariff schemes based on the new approach have been provided upon request of the authorities awarding new concessions, but might have a broader use (for instance, for the annual update of motorway tolls).[55] Recently, Decree Law 28 September 2018, no. 109 ('decreto Genova') has made a step in this direction. The task of establishing toll schemes based on price cap regulation, with the determination of the X productivity index every five years, has been extended to the update or revision of existing contracts. In case of update or revision of contracts, the adjudicating authority is required to monitor the application of existing toll systems with reference to the actual implementation of investments already covered by the tariff scheme.

Summing up, benchmarking methodologies are expected to play an increasing role for highways in Italy, both for the design of tenders and the design of contracts, also in case of in house awards. The role of ART in carrying out the tasks established by the law may be useful to promote more uniform efficiency-enhancing regulatory methods. Monitoring the maintenance and investment costs of non-toll highways and roads would also be useful to broaden the information basis for the entire sector.

8.3. Rail

8.3.1. Description of the sector

In Italy, as in some other European countries such as France, Germany and Austria, a State-owned vertically integrated undertaking, structured as a holding company (FSI), plays a major role both in the management of the rail infrastructure (through its subsidiary RFI) and in rail transport services.[56]

As illustrated in Chapter 5 of this Report, benchmarking as a regulatory tool can be relevant both for rail infrastructure, due to its natural monopoly features, and for those public transport services which are provided on the basis of exclusive rights or, in any case, are subsidized to compensate for public service obligations. In Italy, since freight transport, international passenger transport and high-speed services have

[55] ART Annual Report 2018, pp. 42-45.
[56] Among the sources of data, see MIT, IRG-Rail Annual Reports on Market Monitoring, ART Annual Reports, the balance sheets of Trenitalia and RFI, National Observatory for Local Public Transport Policies.

been liberalized and are not subject to public service obligations, the analysis of services will focus on regional rail transport services, provided on the basis of public franchises, and on subsidized medium to long distance passenger transport services.

Looking at the current organization of the sector, the national infrastructure managed by RFI coexists with local networks representing less than 20 per cent of the total rail network[57] and which in some cases are not fully interoperable. The conditions for the management of the national rail network, including investment plans, are set in a contract between the Ministry for Infrastructure and Transport and RFI.

Medium to long distance services under public service obligations are provided by Trenitalia on the basis of a service contract for the period 2017-2026. In 2017 such subsidized services accounted for 4239 million passengers-km, whereas non-subsidised medium to long distance services provided by Trenitalia accounted for 16643 million passengers-km.[58]

As anticipated, regional transport services are provided on the basis of public franchise contracts. In most cases, these contracts are awarded directly by the competent administrations; only in a very small number of Regions the contract has been awarded on the basis of a tender or competitive dialogue.[59]

Trenitalia, directly or by means of controlled companies, plays a major role in the provision of regional passenger rail transport services (with 24412 passengers-km in 2017, with an increase of 3% over 2016), although in some areas contracts are held also by other companies, often publicly-owned. The contracts for local rail transport are approximately fifty; according to the estimates of the Italian Competition Authority, 28 per cent of the 2.8 billion compensation granted by Regions for local rail transport services are granted to companies not controlled by FSI.[60]

Pursuant to Regulation (EU) 2016/2338, which has modified Regulation (EC) 1370/2007, direct awards for rail passenger transport are admitted only until 2023; subsequently, public administrations should normally resort to public tenders for contract awards. In Italy, since the end of 2016 several new contracts have been awarded on the basis of the transition rules, often with a duration of ten/fifteen years.

8.3.2. Areas which should be monitored within an efficiency enhancing strategy

In the light of this market structure, the areas which should be monitored within an efficiency enhancing strategy include operating costs and the cost of investment for infrastructures, as well as operating costs for subsidized transport services.[61] In both areas, the quality dimension should also be looked at.

Focussing on a competition/regulatory perspective, the main challenges for public policy are the following:

a. addressing the issue of rail transport within an overall programming strategy including all alternatives;

b. reassessing on a regular basis whether the traditional boundaries for public service obligations and the way in which they are financed are still justified. This issue is relevant both for regional rail transport services and for subsidized medium to long distance transport and, although it has not much to do with benchmarking as a regulatory tool, it is worthwhile recalling that there is not a

[57] ART Annual Report 2018, p. 121.
[58] Trenitalia, 2017.
[59] ART Annual Report 2018, p. 51.
[60] Autorità garante della concorrenza e del mercato (2016), Indagine conoscitiva sul trasporto pubblico locale, § 116.
[61] The further issue of the allocative efficiency of investment plans in rail infrastructure compared to alternatives, which is also crucial for public policy, goes beyond the scope of this Report.

once and for all answer, since it strongly depends on the existence of alternatives and on their costs;

c. defining the optimal dimension which should be considered in the award of contracts and in benchmarking exercises;

d. deciding how to award contracts (directly or via public tenders);

e. establishing the content of contracts, both for tenders and for direct awards, so as to ensure efficiency enhancing incentives.

8.3.3. Empirical evidence and data for benchmarking

Infrastructures

For rail infrastructure, an estimate of the minimum efficient dimension for infrastructure management, similar to what the regulatory authority has already done for highways, although not expressly included in the tasks of ART, can be linked to the goal of ensuring, according to methodologies incentivizing competition, productive efficiency and limiting the cost for users, fair and non-discriminatory access conditions to infrastructure (art. 37(2)(a) d.l. 201/2011).

Most of the data which can be used for benchmarking exercises are provided directly by the incumbent, and refer to the different contexts in which it operates. Other information on operating and investment costs can be collected by the other local infrastructure managers, but as already indicated, they represent only a minor share of the network.

Therefore, for the management of infrastructures the most reasonable strategy is promoting international benchmarking, by means of a common methodology and the collection of high quality comparable data, taking the quality dimension into account, at the UIC level or IRG-Rail level. This kind of international initiatives would also help overcoming the current asymmetries of information between public authorities and incumbents in the different Member States. In such a comparison, it should be taken into account that, especially for State-owned companies, the strong power of labour unions may entail that the combination of assets and staff at a particular point of time may not be optimal. National data could be used; or regional data within countries could be combined together to increase the sample size and richness of the comparison.

Rail passenger transport services

For rail passenger transport services, the evidence relating to Italy shows that economies of scale are more pronounced than for bus transport (with a minimum efficient scale at 40 million train-km per year), justifying the award of contracts at least at the regional or supraregional level.[62] Economies of density are also significant.

With respect to economies of scope in the intermodal provision of services, the Italian Competition Authority has repeatedly warned against the risk of undue restrictions of competition if bus services (with lower barriers to entry) are bundled with rail services in the award of contracts.

The data on the cost of service are mainly provided by Trenitalia, which for the purposes of public service compensation at the regional level submits an average of the costs incurred in the provision of services in the different areas.[63]

[62] See for instance A. Boitani (2016), *I costi standard del trasporto ferroviario e la concorrenza per il mercato*, Sipotra.
[63] AGCM (2016), Sectoral Inquiry on Local Public Transport.

Internal benchmarking is a good starting point. However, in order to ensure its effectiveness as an efficiency-enhancing tool, the different segments of the company should enjoy a significant degree of autonomy and be incentivised to outperform each other.

8.3.4. Policy issues and the tasks of the regulatory authority

As to infrastructures, ART may use international benchmarking, looking in particular at the activities performed within IRG-Rail, to strengthen efficiency-enhancing incentives in the future revision of the regime already established for the management of the network by RFI.[64]

As to rail transport services, the tasks of the Italian regulatory authority include:

- defining model-tender calls for public franchises and model service contracts (art. 37.2.f decree law 201/2011);
- defining, after hearing the Ministry for Infrastructure and Transport, the regions and local authorities concerned, the areas subject to public service obligations (PSO) on the different tracks and the way in which PSOs should be financed (Art. 37.1 decree law no. 1/2012).

With Decision 49/2015, ART adopted regulatory measures on model tender calls and related contracts for the award of local transport services, including rail. This important Decision focuses on the essential and instrumental goods and information which should be made available to participants, on requirements for participation to tenders and adjudication criteria, on how the adjudicating entity should prepare a simulated economic and financial plan as a basis for the assessment of the offers it receives and on the criteria which should be followed in the design of the economic and financial plan by participants to the tender. Moreover, Decision no. 49/2015 sets requirements for the content of contracts between the administration and the provider of the service, which should apply also to contracts awarded directly and to revisions of contracts undertaken after its entry into force.

Looking in particular at benchmarking, Decision no. 49/2015 is relevant in several respects:

- administrations are invited to prepare the simulated economic and financial plan taking as a benchmark an average efficient undertaking (although for labour costs they should refer to the existing contracts, considered exogenous variables; notably, for public transport companies national contracts are more generous of those in the private sector);
- for tenders, the basis for public service compensation should be established in the tender call not only according to the principles of Art. 4(1) of Reg. 1370/2007, but also of Art. 17 of Legislative Decree no. 422/1997, i.e. standard costs;[65]
- for direct awards, although the Decision makes no express reference to an efficient benchmark,[66] it states that the contract should include incentives aimed at a progressive improvement of performance, accompanied by sanctions if the objectives are not reached. Exceptions are allowed only for "*de minimis*" contracts and emergency situations (Art. 5, paragraphs 4 and 5, Reg.

[64] Pursuant to Art. 37(2)ART. See *inter alia* Decisions 96/2015, 152/2017.

[65] For a discussion of the standard cost methodology, see paragraph 8.4 on bus transport.

[66] For in house or direct awards, compensations of public service obligations must respect the principles of Art.4.1.b of Regulation 1370/2007, Annex 1 thereof as well as the criteria set forth in Annex 1, Table 4 of Decision no. 49/2015. Reference should be made to incremental costs deriving from the PSO plus a portion of the common costs, according to regulatory criteria for their allocation. Regulation 1370/2007 does not make reference to an efficient benchmark but only to the observed costs of the company, requiring only separate assessment of PSO for rail and bus transport services in case of multimodal contracts.

1370/2007). The objectives may include not only an increase in the number of paying passengers and quality improvements, but also the reduction of costs or an increase in productivity with reference to comparable efficient benchmarks. Thus, in the case of direct awards, although Regulation no. 1370/2007 simply requires no overcompensation with respect to the observed costs of the train operating company, public authorities may introduce in the service contract efficiency enhancing incentives with reference to appropriate benchmarks.

In particular, ART recommends the competent public authorities to adopt a price cap mechanism accompanied by a symmetric subsidy cap mechanism (including a X factor relating to productivity), which will be especially relevant in case of direct awards.

As to the choice of in house/direct awards instead of a tender, in Decision no. 49/2015 ART argues that the awarding administrations should provide justifications for their choice, in terms on universal and social goals, efficiency, cost effectiveness and quality of service, pursuant to Article 34 of Decree Law no. 179/2012 on local public services.

Decision no. 49/2015 will be revised by December 2018 so as to take into account not only lessons learned from the first three years of application, but also legislative developments. The main developments include:

- the new Code for Public Contracts – legislative decree no. 50/2016;
- (EU) Regulation 2016/2338 which has modified (EC) Regulation no. 1370/2017 in EU, setting in particular year 2023 as the general deadline for direct awards in regional rail transport services;
- Legislative Decree no. 50/2017 on local public transport, which:
    a. strengthens the efficiency enhancing criteria in the allocation of the resources of the National Transport Fund, also with reference to standard costs (Art. 27);
    b. assigns RFI a preminent role in upgrading local infrastructures (Art. 47);
    c. assigns Regional governments the task of defining catchment areas (*bacini di mobilità*) for regional and local transport services, on the basis of the analysis of demand. The governing body of each area is required to divide contracts into lots, which would support benchmarking. Exceptions will be allowed, on the basis of criteria which will be determined by ART, for reasons of economies of scale and efficiency. For rail services, the Decree acknowledges that exceptions can be provided also for lots covering several Regions, with the consent of the interested regional governments (Art. 48)
    d. expressly assigns ART the task of establishing the content of the service contract also for in house awards. Both for such contracts and for tender calls ART is required to establish the effectiveness and efficiency goals, and the objectives of financial equilibrium for providers. Accounting separation should always be ensured between activities carried out under PSO and other activities. ART is also required to set rules on the access to essential and instrumental goods for the new entrant winning a tender and on the transfer of the workforce from the incumbent to the newcomer, on the basis of criteria of social protection established in the law.

A separate investigation was launched by ART in 2016 (Decision no. 83/2016) with the aim to elaborate methods for defining the scope of public service obligations and the most efficient ways of financing them for local transport services. The approach undertaken at the EU level in the control of State aid can provide some guidance: although the Member States have an appreciable discretion when establishing the scope of services of general economic interest, before imposing public service obligations public administrations should ascertain whether market forces are sufficient to meet the public interest objectives (accessibility of mobility services at affordable prices) with no need of public service obligations, taking all the available alternatives into account. Moreover, in order to avoid overcompensation, from a national perspective it makes sense to go beyond what is required by Regulation no. 1370/2007 (no overcompensation with respect to the observed costs of the incumbent), by means of efficiency benchmarks.

In 2017, a further proceeding was opened to adopt measures aimed at promoting the efficiency of regional rail transport services (Decision no. 69/2017). Interestingly, data collection aimed to estimate efficient costs played a crucial role in this procedure. Ideally, in order to adopt efficiency enhancing regulatory measures public decision makers should have access to an adequate dataset, with safeguards aimed at ensuring that the data are not only reliable but also sufficient for the specific purposes of regulatory activities. In order to overcome the persistent information asymmetry between TOCs and regulators, the procedure started with an appreciable effort of data collection from operators to carry out estimates of efficient costs. Within this procedure, ART launched a public consultation on a model of service contracts, to be used in case of direct awards, aimed at increasing efficiency and effectiveness in regional rail transport services. The model is based on the estimation of efficient cost frontiers, following an approach similar to the one already used for highways (Decision 66/2018). The final decision was adopted at the end of November 2018 (Decision 120/2018).

In a broader perspective, the collection of proper data on costs in support of effective public policies should be organised on a regular basis. On the other hand, especially in view of an increasing use of public tenders, collecting information for a proper benchmarking is not sufficient to make the most of competitive forces. It remains crucial to tackle complementary issues such as the availability of rolling stocks, access to infrastructures and social clauses, in order to reduce, when necessary and proportionate, existing barriers to entry by means of appropriate pro-competitive regulatory measures.

As to the issue of horizontal and vertical integration in the sector, providing evidence on the pros and cons of alternative models of separation of infrastructure management from the provision of transport services is, from the very beginning, among the tasks of the regulatory authority, which therefore is supposed to continue collecting the relevant empirical evidence, also in the light of international experiences. So far, as shown in Chapter 5 of this Report, empirical evidence does not seem to strongly support a full unbundling scenario compared to less intrusive alternatives aimed at avoiding cross-subsidization.[67] Empirical studies on the impact on efficiency of different growth strategies by incumbents in the rail sector can also provide useful information to public decision makers, which would be useful, in particular, for the assessment of the external growth strategy of the publicly-owned FSI group both horizontally (control of smaller rail concessions) and in other transport sectors (services and infrastructure).[68] An important but not yet researched question is whether the appropriateness of vertical separation / or more integrated structures could vary within a country, leading to a mixed approach (since previous studies have used national data).

Summing up, in the rail sector benchmarking is relevant both for infrastructures and for subsidized services. As to infrastructure, ART may promote international benchmarking for investment and operating costs, taking the quality dimension into account. For regional passenger transport services and for medium to long distance passenger transport services subject to public service obligations, benchmarking, which is especially relevant in case of direct award of contracts, should be included in a broader efficiency-enhancing strategy.

[67] CER (2012); *The effect of vertical separation in the railway sector.*

[68] For a discussion, also in an international perspective, see M. Sebastiani (2017), “La questione dimensionale nel settore ferroviario e dintorni”, in S.I.Po.Tra, *Razionalizzazione dei mercati e aggregazioni fra imprese di trasporto*, Quaderno n. 2/2017, (http://www.sipotra.it/wp-content/uploads/2017/04/Quaderno-AGGREGAZIONI-FINALE.pdf).

8.4. Bus transport

8.4.1. The current organization of the sector

At present, in Italy the main features of local public transport by bus are the following:

- a large number of firms in the urban sector, almost one company in each town and city; overall, approximately 900 operators were active in 2016 (1230 in 2000)[69]; approximately one quarter provides only urban services, more than one half only extra urban services, the remaining part both urban and extra urban services;
- a strong prevalence of publicly-owned firms, especially in the urban sector; private companies, which overall provide less than 20 per cent of local transport services, are present only in small towns or for non-urban services;
- only 10,5% of undertakings have more than 100 employees, only 20 undertakings more than 1000; fragmentation is especially high in Southern Italy;
- few foreign companies, mainly in the North;
- more than 20% of operators carry out losses, high subsidies notwithstanding (Atac and Cotral accounting for 70% of annual deficits).[70]

8.4.2. Main problems in terms of efficiency

Improving efficiency and ensuring a more effective use of public resources are key challenges for local public transport in Italy.

On the one hand, compensation for public service obligations has a significant impact on public budgets. According to the most recent data, the share of revenues from traffic on the public cost of the service (measured as the sum of revenues from traffic and public support) is approximately one third[71].

Comparisons among Regions are difficult because of differences in contracts, which in some cases are based on gross costs and therefore do not provide information on revenues from traffic. The available evidence shows that in some Regions, including Veneto, Valle d'Aosta and Sicilia, revenues from traffic represent more than 40 per cent of total public cost.72

Public compensation for public service obligations (operating costs) amounts to 2,4 euro/km, whereas in France it is 2,2 euro/km, in Spain 1,7 euro and in Germany only 0,9 euro. The impact on public expenditure amounts to 6,5 billion euro a year, the largest share coming from the National Transport Fund (Fondo nazionale trasporti) established in 2013. Decisions on public expenditure are taken at the regional and local level, whereas funding comes mostly from State resources ("finanza derivata").

On the other hand, unit production costs are high compared to the costs in other European countries. In particular, unit production cost of the service in the larger Italian urban areas is higher than 5,5 €/bus-km,

---

[69] Conto nazionale infrastrutture e trasporti 2016-2017.

[70] Detailed information on the sector can be found in the annual reports of Osservatorio per il trasporto pubblico locale (TPL) and on the website of the Osservatorio TPL, in *Conto nazionale infrastrutture e trasporti-CNIT*, in the publications by ASSTRA and ANAV, in the sector inquiry on local public transport published by the Italian Competition Authority (AGCM) in 2017, in the Annual Reports of the Italian Transport Authority, in the reports by Corte dei conti on publicly-owned enterprises. See also V. Karantounias, D. Pinelli, *Local State-Owned Enterprises in Italy: inefficiencies and ways* forward, European Economy, Economic Brief 10, April 2016.

[71] The data on revenue from traffic and total public costs can be found in the annual Report of the Osservatorio TPL.

[72] Osservatorio TPL, October 2017.

whereas in Sweden it is around 3,4 €/bus-km and in the British metropolitan areas (London excluded) is in the order of 2,5 €/bus-km. All cost components are higher in Italy, but the difference is especially high for the labor cost per unit of production, that in Italy is almost double the size of that of Sweden and the UK.

Services supply on average is abundant and exceeds demand more than in other countries (0.6 passengers/km versus 2,8 places/km; whereas in France the corresponding values are 0.8 passengers/km versus 1,9 places/km and in Spain 1,2 passengers/km v. 2,7 places/km). The load factor is 22%, compared to 28% in the UK, 42% in France and 45% in Germany.

Tariffs, which are established at the local level, are comparatively low[73], and the perceived quality of the service lower than for private transport.[74] However, the situation is not uniform across Italy. Significant variance of situations at the local level (revenues and average costs/km, productivity, share of public compensation), which apparently do not depend only on geographic and socio-economic differences, indicate that there is room for improvement. In this context, both competitive pressure (either for the market or in the market) and benchmarking may play a positive role in steering the incentives of companies so as to improve the efficiency of the sector.

8.4.3. The evolution of public policies: from the reform of 1997 to legislative decree no. 50/2017

In Italy attempts to modernize local public transport and enhance its efficiency started in the late Nineties. With law no. 422/1997, the Italian government, in view of a situation where bus operators on average enjoyed subsidies covering more than 70 per cent of their operating costs, adopted a radical reform of the sector. The programming of services and the management of subsidies were shifted from the national to the regional level. Firms were required to sign formal agreements with local governments (service contracts) defining the rules that the providers of the service must comply with and addressing important issues such as reimbursement and risk-sharing schemes. The reform required that at least 35% of costs should be covered by revenues from traffic. Moreover, it established an obligation on public administrations to rely on competitive tendering for the allotment of service concessions; incentive mechanisms in the allocation of subsidies (e.g. through a subsidy cap) were also introduced, which were expected to enhance efficiency.

Subsequently, this legislative framework was subject to repeated changes and suffered from the overall instability of rules which characterized the local public services sector in Italy in the last two decades. Since 2009, national legislation follows the approach of Regulation (EC) no. 1370/2007, whereby public administrations may choose between tenders and direct awards to in house providers. In this context, public service obligations, their compensation and exclusive rights should be transparent and clearly spelled out in service contracts, in compliance with the principles of non- discrimination and proportionality.

A huge resistance to pro-competitive reforms of local public services led to a referendum in 2011 in which citizens voted against a framework which favoured call for tenders for the award of contract services. Although the focus of the referendum was not on the transport sector, the result slowed down substantially the liberalization process also in local public transport.

In parallel, Decree Law no. 138/2011 provided that the Regions should establish optimal management areas (*ambiti territoriali ottimali* - ATO) and gave the corresponding governing entities competences on the

[73] ART Annual Report 2018, Figures 96 and 97.
[74] ISFORT -Audimob, (2015).

organization of services and award of contracts.[75] Moreover, Decree Law no. 201/2011 entrusted the newly established Transport Authority with the task of establishing minimum quality conditions for transport services subject to public service obligations (Art. 37(2) (d))).

As to the award of contracts, current legislation (decree law no. 179/212, Art. 34(20)) only requires the ATO governing bodies to publish a report explaining the scope of public service obligations, the corresponding public compensations, the reasons for the choice to organise a tender or proceed via direct awards, with express reference to the compatibility with EU rules and with reference to universality, social goals, efficiency, cost-effectiveness, quality of service.[76] Such reports, to be made available on the website of the Observatory for local public services at the Ministry for Economic Development, should be accompanied by an economic and financial plan, with specific provisions for in house awards.

In Decision no. 49/2015 on local public transport, discussed in paragraph 8.3 with reference to regional rail transport, ART published a standard tender notice, with a focus on the conditions of access to infrastructures and the minimum set of information which should be provided to participants; the criteria for the economic and financial planning; the minimum content of service contracts; the composition of selection boards; the obligation to carry out a consultation with stakeholders before proceeding with the tender. As to in house awards, a further constraint is represented by the limits set by Regulation no. 1370/2007 on the possibility to provide services outside the relationship with controlling entities.

In 2016, the reform of publicly-owned companies (legislative decree no. 175/2016) clarified that also in-house companies may fail, although so far, huge deficits notwithstanding, almost no company has been left to go bankrupt, delivering a strong message through the sector that efficiency is not the main issue. The reform provides that all publicly controlled companies should adopt specific programmes to prevent the bankruptcy scenario. Moreover, public shareholders are required to prepare rationalization plans aimed to reduce costs and increase the efficiency of publicly owned companies, which may include dismissals of shareholdings and aggregations.

Since 2017 detailed information on public service compensations should be published on the National register on State Aid, established at the Ministry for economic development, also for "*de minimis*" aid and for aid which should not be notified to the European Commission being consistent with Regulation no. 1370/2007.[77]

Legislative Decree no. 50/2017 entailed a further quite significant reform of the local public transport legislation aiming, inter alia, at strengthening the incentives to an efficient provision of services. In particular, the public compensation is linked both to the share of costs covered by traffic revenues, to standard costs pursuant to Art. 1(84) of Law no 147/2013, and to whether the award of the service is based on a public tender. The role of the National Observatory for Local Transport Services as the national database for all data concerning the provision of the services is strengthened. The new legislative framework stresses that programming of services should be based on a careful analysis of demand and of the available alternatives for mobility. The same holds the definition of catchment areas (*bacini di mobilità*) at the regional and local level and their governing bodies; contracts should be divided into lots; exceptions are allowed only in the presence of economies of scale, other efficiency reasons and territorial specificities, on the basis of criteria spelled out by the Transport Authority pursuant to Art. 37(2)(f) taking into account

[75] Art. 3bis, §1.
[76] Art. 3-bis of decree law no. 138/2011, modified in 2014.
[77] Only when compensation of public service obligations is not State aid because it satisfies the four cumulative criteria set by the Court of Justice in the *Altmark* judgment (including either public tender or benchmarking), the relevant information on compensation of public service obligations should not be included in the Register.

current and potential demand, economies of scale and service integration. Moreover, the Transport Authority is given the task to prepare also draft contract services for in house companies and publicly controlled companies. Both for tender notices and for service contracts in case of in house or direct awards, the Authority should specify the objectives of efficiency, effectiveness and financial and economic balance for providers; all contract services should ensure accounting separation between activities under public service obligations and other activities.

8.4.4. The current approach to standard costs

Within the overall current legislative framework, standard costs matter in different respects, including in particular in the allocation of state resources and as a basis for public service compensation in tender notices.[78]

A methodology for the definition of standard costs pursuant to law no. 147/2013 has been introduced by a recent ministerial decree (no. 157/2018 of 28 March 2018).

The definition of "standard costs" is a progress compared to the former rationale, based on "historical costs", although there may be different views on the correct way to calculate the "standard". The approach adopted by the government is explained in Avenali at al. (2016)[79] and summarized in the Box.

BOX 8.1. STANDARD COSTS FOR LOCAL PUBLIC TRANSPORT IN ITALY

The dataset concerns 45 firms producing more than 500 million bus-kilometres. Data have been collected by means of questionnaires sent to managers of Italian companies (20 private and 25 public-owned companies) providing LPT services. The 45 firms contributed to create a dataset composed of 54 separate units referred to services bundles ("a service bundle is the set of one or more service contracts for which the firm is able to measure (only) jointly its direct and indirect costs,".. Avenali, at al., p. 167). Basically, the cost obtained for every unit is not derived from accounting data as stated in the financial reports, but is an evaluation provided by the firm, as reported in the questionnaire. The total cost of every unit is obtained as the sum of the cost of the different activities and materials that contribute to generate the output: the cost of driving, depots and movement personnel, the cost of fuel, maintenance, depreciations, administrative costs and the cost of capital. The cost of capital has been recalculated by the estimation of the WACC of the LPT sector on the net invested capital approximated by the net value of the vehicles owned by each firm (net book value). The average cost per bus-Kilometre is 3.6 euro (min 2.1, max 7.9).

The authors explain the variability of the unit costs of the 54 units by a linear regression model. The main explanatory variables concern commercial speed, million of bus-kilometres, the degree of renewal of the fleet, defined as the ratio between a monetary value (current value of depreciation, calculated by Current Cost Accounting method and adjusted by assuming 15 years depreciation life) and bus- kilometres. In order to take into account the nonlinear relationships concerning the effects of the commercial speed, three nonlinear functional forms have been tested.

Summarizing the results, the commercial speed is indicated as the most important driver of the unit cost and the relationship assumes an L-shape.[80] The correlation between the unit cost per bus-Kilometre and the investments in bus fleet and the cost incurred for the provision of the service is positive and also very important.

[78] Decree law no. 95/2012.

[79] A. Avenali, A. Boitani, G. Catalano, T. D'Alfonso, G. Matteucci, (2016), "Assessing standard costs in local public bus transport: Evidence from Italy", in *Transport Policy*, 52, 164-174.

[80] This suggests that the use of reserved lanes should be promoted; in general, the overall impact on private traffic should also be considered.

Some doubts remain about the measure of costs and the methodological approach.

Firstly, the questionnaire used to collect the data was meant to create a complete, certified and updated database for the monitoring of the industry, but the values are not certified by a direct relationship between the questionnaire and the income statement of every firm.

Each company has specific managerial peculiarities which induce it to attribute different values to the different activities that characterize the service and the way of combining the inputs has a strong impact on the dimension of the unit cost *per bus-Kilometres*. Relying on the questionnaire, the costs of the activities that characterize the service are added analytically and in this way the authors obtain the total cost of the 54 units. Given this approach, there is no comparison with the actual accounting data. In addition, and more importantly, the managerial ability of optimizing in a different way the various activities is neglected. The result is a cost dimension that could be far from the real situation of the firm. In order to improve the quality of the findings, it is necessary to consider the accounting data of the firms, obviously by paying attention to the inflation rate and to the opportunity cost of financial capital. It would therefore be important to compare the unit costs built from the report data and the unit costs obtained from the questionnaires.

Secondly, the use of the variable “renewal of the fleet” as it is defined could be problematic, since it raises an endogeneity problem. Basically, the variable is defined as depreciations + rents per bus-kilometre, that is a component of the dependent variable (cost per bus-kilometre). It does not seem correct, even intuitively, to explain a variable using a portion of the same variable. In addition, from the econometric point of view, the variable “renewal of the fleet” is likely to be an endogenous regressor. Suppose that the physical capital is managed inefficiently (e.g. some firms in the sample have overinvested and have too many – underutilized - buses). The effect of such inefficiency (omitted, since it is not directly observable) on the dependent variable should be collected by the residual; however, it has an impact also on the mentioned regressor, which is correlated with the residual itself. This endogeneity problem is serious in an OLS model. The variable “renewal of the fleet” could be substituted by a more suitable option (e.g. average age of the fleet, if available).

As to policy targets, Avenali *et al.* (2016) explain that they “…do not use a frontier approach in order to define the minimal efficient cost for the provision of LPT services, because of the general will of the Minister of Infrastructures and Transport (MIT) and of the Italian Regions for a gradual financial reorganization of the LPT sector..- (p.165)”. In table 3 (p.168) the authors report the descriptive statistics of the unit cost per bus-kilometre, where the mean value is 3.6 €/Km, ranging between a minimum level of 2.1 and a maximum of 7.8. Given the strong variability of the unit cost, it is hard to accept the average unit cost as a “standard cost”, and we feel that a better measure is necessary.

The unit cost per bus kilometre is a partial productivity measure and the use is quite problematic especially in multiple output industries, such as the bus industry, where the output may be also described by the passenger-Km or by the size of the network.

Thus, the current “standard costs” approach, although it represents an improvement over past approaches, may tend to freeze the past inefficiency of the sector. The alternative economic approach to verify whether firms operate in an efficient manner is to estimate a cost function using econometric methods, or a non-parametric methods – such as data envelopment analysis (DEA). “An approach that simultaneously takes account of multiple denominators / drivers of costs is therefore needed - this being a key advantage of the econometric approach as multiple cost drivers can be included in the model at the same time.” (Smith, Chapter 3 of this Report)”. Another fundamental advantage of the econometric approach is the possibility to estimate the shape of the cost frontier inside of the same function, while the reliance on unit cost (as exemplified by Smith, Chapter 3) implicitly assumes constant return to scale (although the scale issue is considered in the subsequent regression analysis).

Finally, as highlighted in Agrell and Bogetoft (2017), relying on rigorous efficiency techniques does not imply the application of a rigid efficiency-recovery system: the inefficiency detected through a model provides an indication of the potential performance improvement for each firm, but the timing of the catching-up process remains, ultimately, a regulatory choice.

8.4.5. Benchmarking, the award of contracts and the role of ART

In Italy, for local public transport competition has never been allowed "in the market". Competition "for the market" (competitive tendering of temporary concessions) has been promoted by law in 1997 but postponed several times. After 20 years, only about one hundred urban concessions have been tendered out, with dismal results: more than 90% have been won with minimal discounts by the incumbent companies. Very soft budget constraints, close links between incumbents and municipalities, labour unions sponsored by the municipalities, that in turn were allowed to be both competitors and judges on the tendering process, all contributed to this result.

In this context, benchmarking comes to play a crucial role in order to ensure some external incentive on providers to enhance efficiency and avoid overcompensation of public service obligations. The absence of significant economies of scale militates in favor of the division in lots of contracts, creating better conditions for competition for the market and benchmarking. Studies on efficient cost frontiers may provide hints for the proper definition of lots of minimum efficient dimension.

On the other hand, as in the rail sector, benchmarking should be included in a more general public policy strategy aimed at ensuring that the incentives of public administrations and operators are aligned in pursuing the objectives of efficiency, quality of service, effective use of public resources.[81] A crucial condition for an efficiency-enhancing strategy is getting rid of conflicts of interest in programming activities and in the award of contracts when local public authorities are also shareholders of incumbent

[81] For instance, the boundaries of public service obligations should be defined within an overall programming strategy, taking into account their costs (buses running empty are a very frequent case) and the available alternatives. In its sector inquiry on local public transport, the Italian Competition Authority suggested that profitable activities which can be provided by competition in the market, under authorization schemes, should be excluded from PSO. When demand is very low, alternatives can be found (e.g. subsidizing demand instead of organizing regular supply of the service). Legislative decree no. 50/2017 takes some steps in this direction, much depends on its implementation. On the other hand, experience in other countries indicates that attempts to introduce competition in this sector should take into account the problems which can arise because if there is not adequate coordination among different operators.
The issue of high unit costs and of low tariffs should also be addressed. Labour cost is especially high within the sector, near 50% higher than in the manufacturing sector. Further special laws ensuring that, in large companies, drivers will remain employed in the same company after a limited number of years in the driving role, are a factor both of inefficiency and of reduced economies of scale. The link between company size and production costs is weak above a small threshold, and even negative (Boitani et al.) The explanation, at least in part, may reside in the strong political clout of the labour unions, that are much stronger in large public companies than in small ones. Tariffs are among the lowest in Europe (at purchasing power parity), and especially so for monthly and yearly passes. This is supposed to be both for income redistribution and for environmental goals. But legitimate doubts exist that the real reason is also related to the search for political support by local public administrators.
Moreover, programming and planning in the LPT sector in Italy would strongly benefit from innovative evaluation tools, based on modern quantitative analysis, economic, social, environmental etc. Only Milan has a transport plan based on an extensive traffic model and a Cost-Benefit Analysis comparing different possible strategies, even if still lacking the financial side. Also in this respect, the new provisions of Legislative Decree no. 50/2017, if properly implemented, may provide the basis for better informed and more rational policy making.

undertakings. Consistently with the directives on public procurement and concessions, the selection board in case of tenders is composed of independent members, with no direct or indirect interests in the sector. Transparency constraints on the relevant choices and detailed guidelines on how to proceed (including standard tender notices and model service contracts prepared by the regulatory authority) may also contribute to this aim.

Public administrations should take benchmarks into account when they have to justify their choice with respect to how to award contracts, especially in order to show that in house award is not inefficient compared to a tender. When the award of contracts is based on competition for the market, standard costs as defined by Ministerial Decree no. 147/2018 represent an improvement, compared to historic costs, as a basis for the tender, but as already illustrated, they tend to reflect the status quo more than a challenging efficiency target. Thus, further improvements in the definition of the benchmark would contribute to stimulate the search for efficiency in cost-based tenders.

On the other hand, the significant impact that very strict social clauses may have on efficiency-pursuing strategies should not be underestimated. In particular, close to last general elections a new constraint has been imposed on the possible “new entrants” in the sector: not only they have to keep the existing workforce, and this not only following the national labour contract for the sector, but also all the local contracts, generally much more expensive, have to be maintained unvaried. This approach may severely shrink any possible benefit of both the “standard costs” rules and any competitive tendering process.[82]

8.5. Seaports and airports

8.5.1. Introduction

As illustrated in the sectoral chapter, there are common elements in the analysis of how competition works and of the role of regulation in seaports and airports. Three types of competition can be considered: competition between undertakings within a single port/airport, competition between undertakings from different ports/airports, competition between ports/airports. Although competition between and within ports and airports is usually sufficient to ensure market discipline, there may remain areas of market failure which justify some form of regulatory intervention.

Both sectors have been subject to a significant market evolution and deserve careful monitoring by competition authorities in order to avoid anticompetitive agreements, abuses of dominant position and ensure that mergers do not result in a significant impediment to effective competition.

Several studies benchmarking performance are available. For airports, the perspective is often that of private investors, whereas for ports the focus is mainly on international competitiveness.

The challenge, for regulatory policy, is being capable of intervening only in so far as market forces are not sufficient to perform the task, in particular with reference to access to infrastructures. The broadly worded tasks of ART pursuant to Article 37 of Decree Law no. 201/2011 appear consistent with such a flexible approach: ART is required to ensure fair and non discriminatory access conditions to port and airport infrastructures, by methods encouraging competition, productive efficiency and cost containment for

[82] As recently recalled by the Italian Council of State, social clauses are compatible with the provisions of the Italian Constitution on freedom of enterprise, competition and good governance of the public administration, as well as with EU rules on fundamental freedoms, only if they do not entail an automatic obligation on the newcomer to maintain the existing workforce, and can be balanced with the needs, in terms of workforce, related to the implementation of the new contract and with the autonomous organizational and entrepreneurial choices of the new provider. Council of State, Section III, 6 June 2018, no. 3471.

users, undertakings and consumers, and has also the power of defining, when justified by competition conditions in local and national markets for transport services, the criteria for the establishment of tariffs, fees, tolls, considering the need to ensure the economic equilibrium of regulated undertakings, productive efficiency and cost containment for users, undertakings, consumers.

Both for ports and airports, the Italian legislation must be seen within the broader EU legislative framework establishing common rules on airports and ports.

8.5.2. Seaports: mean features of the port sector in Italy

Italy has 282 ports, with an overall traffic volume of 462 million tons for freight and 70,3 millions passengers.[83] Only 12 ports reach a traffic of more than 1 million passengers, 15 ports a freight traffic exceeding 10 million tons, 10 ports a traffic of containers exceeding 200 000 TEU.[84]

The relevance of Italian ports at the EU level decreased, from 13,7 per cent in 2005 to 12 per cent in 2015. The issue of international competitiveness of the Italian port system as part of a broader production and logistic chain is crucial within a broader strategy for growth and employment.[85] A national plan with a strategy for ports and logistics, including investment in infrastructures, was adopted in 2015 and updated in 2016.[86] The Strategy pursues ten objectives, which include the promotion of competition, transparency and upgrading of services, improving accessibility and the integration of logistics, closer coordination of initiatives at the national level.

The Italian legislation (Law no. 84/1994) has been broadly revised between 2016 and 2017 (Legislative Decree no. 169/2016, modified by Legislative Decree no. 232/2017). At the EU level Regulation no. 2017/352, which is directly applicable in all Member States, has established common rules on the provision of port services and the financial transparency of ports which, although applicable only to the ports of the trans European transport network (TEN-T, established by Regulation (EU) no. 1315/2013) and to some services (for instance, pilotage is not included), set out a common approach to competition advocacy and promoting efficiency in European ports.

The new governance of the sector in Italy, since the reform of 2016-2017, is based on 16 Port System Authorities, in charge of the 57 main commercial ports. The reform entailed a reduction in the number of port authorities and envisages enhanced coordination between the Ministry for Infrastructure and Transport and the Port Authorities, among the Authorities, between the Authorities and other public administrations involved in the port and logistics system.

As to competition policy, since its establishment in 1990 the Italian Competition Authority has often used its enforcement and advocacy powers to protect and promote competition in ports.[87]

[83] MIT, 2015-2016.

[84] For more detailed data, see the figures in the ART Annual Report 2018, pp. 128-141 and MIT Annual Reports on the Activities of Port Authorities.

[85] See L. Antonellini, P. Costa, F. Munari, P. Spirito (2017), "Riforma e problematiche della portualità", in Sipotra, *Le politiche dei trasporti in Italia. Rapporto 2017*, 271-308.

[86] *Piano strategico della portualità e della logistica*, approved by d.P.C.M. 26 August 2015; *Connettere l'Italia: fabbisogni e progetti di infrastrutture, Annex III to the Documento di economia e finanza 2017 (Allegato infrastrutture 2017).*

[87] The relevant decisions and opinions are available at the AGCM website (www.agcm.it).

8.5.3. Areas which should be monitored within an efficiency enhancing strategy

The focus of this Report is on economic regulation and therefore, on areas of possible market failure. However, it is worth recalling that there are complementary issues which an efficiency enhancing public policy should address. The first one, considered both in Regulation (EU) no. 2017/352 and in the National Plan for Ports and Logistics, is administrative simplification in ports. The other one is the assessment and financing of investment in port infrastructures, looking both at whether PPP is feasible and at the choice of priorities in the use of public funds.

Focusing on possible areas of interest in a microeconomic perspective, a first candidate[88] is conditions of access to port areas and infrastructures, which remain in the public property, especially when competition between ports is not sufficient to incentivise public authorities to ensure access at fair, reasonable and non discriminatory conditions. With specific reference to ports (although similar remarks could be made for other transport modes) in 2017 the Italian Anticorruption Authority, in its annual plan for the prevention of corruption, stressed that the access policy issue is also relevant in the perspective of transparency of the activities of public entities and the prevention of corruption.

A related issue is how to ensure that the choices of the competent authorities with respect to authorizations and franchises for the provision of services in the port area make the most of the possibilities for competition in the market (between different operators within the port) or, when it is efficient to award exclusive or special rights, competition for the market.

As to whether also the economic conditions applied by providers of port services should be monitored and constrained, in order to ensure productive efficiency and fair economic conditions to users, in principle a role for the competition authority or for ex ante regulation should be limited to cases of lack of effective competition.

A specific problem, in Italy, is that these issues have been addressed at the local level with different approaches, whereas a shared framework would facilitate benchmarking and the sharing of best practices.

Finally, still in a microeconomic perspective, the use of public funds either to compensate public service obligations in the provision of services of general economic interest within ports or to support investment in infrastructure should not go beyond what is necessary and proportionate to pursue a well defined objective of public interest, when market forces are not sufficient, and ensuring that distortions of competition are minimised.

At the EU level, the economic framework behind (EU) Regulation no. 352/2017 apparently acknowledges these areas of policy concern. The starting point is that the full integration of ports in seamless transport and logistic chains is needed to contribute to growth and the internal market. This requires modern port services, that contribute to the efficient use of ports, and a climate favourable to investments for the development of ports in line with current and future logistics requirements. The attractiveness of maritime transport depends on the availability, efficiency and reliability of port services, and on whether questions regarding the transparency of public funding and port charges, administrative simplification and restrictions on the provision of services at ports are properly addressed. Although, as anticipated, the prescriptions of the Regulation apply only to maritime ports of the TEN-T and to a subset of services, Member States may choose to follow the same approach also for other ports and services (e.g. cargo-handling, passenger services, pilotage services). In any case, the principles set out by the case law of the Court of Justice apply.

[88] The public policy issues concerning the assessment and financing of investment in port infrastructures goes beyond the scope of this Report.

In particular, pursuant to (EU) Regulation no. 352/2017, qualitative requirements and any limitation on the number of providers of port services should be strictly justified and the procedure for granting the right to provide services must be transparent and non-discriminatory. If the Member States or the contracting parties demonstrate that the relevant activity is directly exposed to competition in markets to which access is not restricted pursuant to Arts. 34 and 35 of Directive 2014/25/EU, the port sector or subsector is not subject to the rules framing market access limitations under the Regulation. For the compensation of PSO and for public support to investment in infrastructures, the Regulation refers to the rules on the control of State aid. Specific indications on accounting separation in ports are provided.

Moreover, as to charges for access to port infrastructure, the managing body of the port is required to set up a transparent and non-discriminatory charging system, which contributes to the maintenance and development of infrastructures and the provision of services, in accordance of the port's own commercial strategy and investment plans and where relevant with the requirements of the general ports policy of the Member State concerned. The Regulation acknowledges that the charging policy may involve some differentiation, provided that the criteria for variation are non-discriminatory and consistent with competition rules.[89]

As to charges for port services, the Regulation acknowledges, with no prejudice to EU competition rules, that Member States may adopt regulatory measures in order to avoid overcharging where the situation of the market is such that effective competition cannot be achieved. In general, for services provided under public service obligation and pilotage services which are not exposed to effective competition and entail a higher risk of price abuse in cases where monopoly power exists. Member States are required to ensure that the charges are set in a transparent, objective and non-discriminatory way and are proportionate to the cost of the service provided.[90]

Port users and other stakeholders should be consulted on the main issues related to the development of the port, including its charging policy. As to the enforcement of these rules and principles, Member States are required to set up an effective procedure to handle complaints and empower specific authorities to adopt binding decisions for dispute resolution.

Moreover, Member States should identify the authorities in charge of collecting information on the elements that serve as a basis for determining the structure and level of port service charges and port infrastructure charges.

8.5.4. The role of benchmarking and regulatory initiatives

Whereas benchmarking the performance of port authorities and comparing the competitive appeal of different ports has a broader relevance in a public policy perspective, from a narrower competition and regulatory viewpoint, collecting information on cost functions and productivity for benchmarking is useful for those services which, at least in some contexts, are not exposed to effective competition and, thus, may require public measures aimed at enhancing efficiency. In particular, benchmarking can play a role in the design of incentivising mechanisms.

In Italy, an open issue is the appropriate design of the governance of the port sector, in particular with respect to the role of an independent regulatory authority such as ART versus other authorities and public bodies (Ministry, maritime authorities, Port System Authorities, Conference of Port System Authorities). Article 37 of Decree Law no. 201/2011 expressly includes port infrastructures among the infrastructures to

[89] Regulation 352/2017, Art. 13.
[90] Regulation 352/2017, Art. 12.

which ART should ensure fair and non-discriminatory access according to methodologies promoting competition, productive efficiency and cost-containment for users. On the other hand, Law no. 84/1994 assigns specific competences to the Ministry for Infrastructures and Transport, Maritime Authorities and Port System authorities, also with respect to access to port infrastructure and services and to charging policies.[91]

In this context, taking Article 37 of Decree Law no. 201/2011 as the legal basis, the Authority has opened a public consultation and adopted draft first regulatory measures on fair and non-discriminatory access to port infrastructures.[92] Concessions relating to the construction and management of infrastructures, which fall within the scope of application of the Code of Public Contracts,[93] are expressly excluded.

The draft measures focus on four main issues:

a. procompetitive criteria for the award of concessions for the use of port infrastructures (port areas and docks), including principles for planning, duration, selection procedures, charging policy. As to charging policy, ART recalls the principles of fairness and non-discrimination and the need to take into account each port development objectives and regulatory accounting principles and requires that the charge includes both a fixed and a variable component. The variable component should include incentivising mechanisms aimed at improving *inter alia* productive efficiency and levels of service, also by means of yearly updates;
b. criteria and procedures for authorizing port operations and services, based on Reg. 2017/352 but applicable also to commercial ports not included in the TEN-T;
c. supervision of port operation and service charges, whereby ART stresses the need of a special focus on those operations and services which require the use of an essential infrastructure and indicates some general criteria for assessing admissible costs and investments;
d. regulatory accounting as a tool to ensure a proper allocation of costs between the different activities.

Leaving aside the discussion of the overlap of competences in the governance of ports and of the actual powers of ART in this area, and focussing instead on the substance of prescriptions, since in Italy the fragmentation of approaches at the local level has traditionally represented a weakness, it is beyond doubt that a common approach to accounting and to admissible costs is useful to promote transparency, comparability and spread of best practices between ports and between port service providers.

In addition, collecting information on minimum efficient scale for the different services would also support port authorities' pro-competitive franchise policies, in particular when establishing limits on the number of service providers.

On the other hand, the boundaries of the activities for which the application of competition rules is not sufficient to prevent the application of unfair conditions and it is appropriate to adopt a regulatory approach should be clearly defined, at least in terms of general principles, in order to avoid either over-regulation or under-regulation, as well as unjustifiedly different practices in different ports.

[91] See, in particular, Art. 16 and 18 of Law no. 84/1994.
[92] Decision no. 57/2018.
[93] Legislative Decree no. 50/2016.

8.5.5. Airports: main features of the airport sector in Italy

The airport sector in Italy, like in other European countries, undertook an in-depth transformation since the 1990s with the development of low cost carriers and widespread privatization. Moreover, in recent years traffic is growing at a rapid pace, both for passengers (175 millions in 2017, i.e. +6,2 per cent from 2016) and freight (cargo) (1,1 mln tons, i. e. + 9,2%).

Although in Italy there are 42 commercial airports (including 13 strategic airports and other 22 airports of national interest), more than 44 per cent of the traffic, both for passengers and freight, is concentrated in the Rome and Milano areas. Only four airports have more than 10 mln passengers, 7 airports between 5 and 10 mln, 12 airports between 1 and 5 mln and 19 airports less than 1 million.

On average, 70 per cent of the turnover of airport companies comes from aviation activities, and 30 per cent from non-aviation activities.

A national plan for investment in airport infrastructures has been adopted in 2015 and a national overall strategy for transport infrastructures, including airports, was outlined in the Annexes to the *Documento di programmazione economica e finanziaria* 2017 and 2018. As in other EU Member States, when investment in airport infrastructures involves the use of public resources, national initiatives have to comply with the EU rules on State aid.[94]

For airport charges, the national framework is strongly influenced by the EU common rules set by Directive 2009/12. The Directive, which complements the EU rules on the provision of air navigation services and ground handling services,[95] stresses that airport managing bodies, offering facilities and services the cost of which they generally recover through airport charges, should endeavour to operate on a cost efficient basis. Establishing a common EU framework regulating the essential features of airport charges and the way they are set is deemed necessary to ensure that basic requirements in the relationship between airport managing bodies and airport users are met. The Directive applies to airports above a minimum size (5 mln passengers) and leaves to Member States the choice whether to allow revenues from airport's commercial activities to be taken into account in establishing airport charges (i.e. double till).

In order to ensure that the system is transparent and non-discriminatory, the Directive provides that airport users should obtain from airport managing bodies information on how and on what basis airport charges are calculated. Applying different conditions is allowed but the system should entail no discrimination. Consultation between airport managing bodies and airport users has to be carried out before taking decisions on airport charges. In case of disputes, each party should be able to have recourse to an independent supervisory authority, empowered to adopt binding decisions.

The Directive is with no prejudice to the right of each Member State to apply additional regulatory measures when appropriate, e.g. economic oversight measures such as the approval of charging systems and/or the level of charges, including incentive-based charging methods or price cap regulation. Member States may decide not to provide a dispute resolution role for the independent authority when there is a mandatory procedure at the national level whereby airport charges or their maximum level are determined or approved by the independent supervisory authority, or there is a mandatory provision whereby the authority assesses on a regular basis or upon request whether airports are subject to effective competition

[94] See the rules on airports and ports in the General Block Exemption Regulation (EU) no. 651/2014 and the Commission Guidelines on State aid to airports and airlines (2014/C 99/03).
[95] Commission Regulation (EC) no. 1794/2006 and Council Directive 96/67/EC.

and, whenever wanted on the basis of such an examination, the Member States decides that the airport charges or their maximum level shall be determined or approved by the independent supervisory authority.

The implementing provisions of Directive 2009/12, contained in Decree Law no. 1/2012 (Arts. 71-82), read in combination with Art. 37 of Decree Law no. 201/2011, entrust ART with tasks of economic regulation and supervision in the sector. In particular, ART is entrusted with the task of approving models for airport charges which should be cost oriented, promote efficiency and be capable of incentivizing investments so as to foster innovation, security and the quality of services.

The managing body of the airport should establish its charging policy on the basis of these models, after consultation with the users of the airport, as provided in the Directive. Art. 80 of Decree Law no. 1/2012 requires in particular that ART monitors that, for charges for the use of infrastructures and services provided on an exclusivity basis, the managing body of the airport complies with the principles of cost-orientation, transparency and fairness, taking into account the average at the EU level of airport charges in comparable airports.

Notably, however, some of the largest airports (Milano, Rome, Venice), subject to *contratti di programma* according to the provisions of Article 17, §34 bis, of Decree Law no. 78/2009, are outside the scope of ART regulatory powers established by the implementing provisions of Directive 2009/12, and remain supervised by the Ente Nazionale Aviazione Civile (ENAC), i.e. the Italian Civil Aviation Authority.[96]

8.5.6. Areas which should be monitored within an efficiency enhancing strategy and empirical studies

For airports, as for ports, an efficiency enhancing strategy has to consider, on the one hand, issues relating to public support to investment in infrastructures, which go beyond the scope of this Report, and, on the other, issues relating to access to infrastructures and the provision of services. As to the latter, public decision makers face two distinct challenges: promoting, when feasible, competition in the market or for the market for activities provided within airports; ensuring, when effective competition is not sufficient, that access conditions to airport infrastructures and airport service charges are transparent and based on efficiency-incentivizing mechanisms. In this context, economic benchmarking can be useful to incentivise efficiency.

Airport efficiency is a debated issue in the empirical literature (Abrate and Erbetta, 2010)[97], where a number of empirical studies provide an excellent support for developing proper benchmarking approaches.

Even if specific analysis of the Italian case is limited, it is possible to find some relevant examples of application of both consolidated and more advanced techniques. For instance, Malighetti *et al.* (2007)[98] rely on DEA and Malmquist indexes, finding higher efficiency for larger structures as well as a positive effect of privatization and hub premium. Among the studies based on DEA, it is relevant to mention also the works by Curi *et al.* (2008)[99] and Gitto and Mancuso (2012a)[100], where the latter work identifies potential

[96] ENAC website: www.enac.gov.it.
[97] G. Abrate and F. Erbetta, (2010*), "*Efficiency and patterns of service mix in airport companies: An input distance function approach", *Transportation Research Part E: Logistics and Transportation Review*, 46 (5), 693-708.
[98] P. Malighetti, G. Martini, S. Paleari, R. Redondi, (2007), "An empirical investigation on the efficiency, capacity and ownership of Italian airports", *Rivista di Politica Economica*, 97, 57–188.
[99] C. Curi, S. Gitto and P. Mancuso, (2008), "An application of Data Envelopment Analysis (DEA) to measure the efficiency of the Italian airports after the privatisation", *L'Industria*, 4, 689–712.
[100] S. Gitto and P. Mancuso, (2012a), "Two faces of airport business: A non-parametric analysis of the Italian airport industry", *Journal of Air Transport Management*", 20, 39-42.

efficiency gains related to several factors (airside activities, presence of private capital, liberalization of handling services and type of concession). Barros and Dieke (2008)[101] use DEA with bootstrapping and second-stage analysis, and the same method is employed by Curi *et al.* (2010)[102], that find a positive impact of public ownership while a negative effect is related to the presence of two hubs. Also Curi *et al.* (2011)[103] rely on bootstrapped DEA, and confirm the negative effect of opening a second hub in Italy involving the performance of the first hub; additionally, the authors highlight the role of the type of concession, while other characteristics (dimension, price regulation) play different roles with respect to operational or financial efficiency. Gitto and Mancuso (2012b)[104] use bootstrapping to correct Malmquist indexes of TFP, finding a significant technological regress (from 2000 to 2006), with few airports improving their productivity thanks to efficiency recovery; a gap exists between North-Central and Southern airports, and the type of concession affects the performance, while ownership does not have a significant role. Merkert and Mangia (2014)[105] rely on a Norwegian-Italian sample and focus on the impact of competition on DEA-based efficiency measures, while Barros and Peypoch (2008)[106] compare Italian and Portuguese airports productivity using Luenberger productivity indicators based on directional distance functions accounting for both input contraction and output expansion, finding an almost generalized increase in productivity. Also Martini *et al.* (2013)[107] rely on the directional distance function framework to include noise and air pollution as undesirable outputs; their results show better performance for public airports, while also the fleet mix impacts on efficiency. With respect to studies based on parametric methods, Abrate and Erbetta (2010) estimate a parametric input distance function in order to assess the presence of cost complementarities between aeronautical and auxiliary activities. The results support the outsourcing of handling activities, while there is not clear evidence with respect to the strategy of diversification towards commercial activities. Scotti *et al.* (2012), instead, employ stochastic frontier analysis to identify the impact of competition on performance, showing that this effect appears to be negative. Also in this case, public airports show the highest efficiency.

---

[101] C.P. Barros and P.U.C. Dieke, (2008), "Measuring the economic efficiency of airports: a Simar–Wilson methodology analysis", *Transportation Research,* Part E 44, 1039–1051.

[102] C. Curi, S. Gitto and P. Mancuso, (2010), "The Italian airport industry in transition: a performance analysis", *Journal of Air Transport Management*, 16 (4), 218-221.

[103] C. Curi, S. Gitto and P. Mancuso, (2011), "New evidence on the efficiency of Italian airports: a bootstrapped DEA analysis", *Socio-Economic Planning Sciences*, 45 (2), 84-93.

[104] S. Gitto and P. Mancuso, (2012b), "Bootstrapping the Malmquist indexes for Italian airports", *International Journal of Production Economics*, 135 (1), 403-411.

[105] R. Merkert and L. Mangia, (2014), "Efficiency of Italian and Norwegian airports: A matter of management or of the level of competition in remote regions?", *Transportation Research Part A: Policy and Practice*, 62, 30-38.

[106] C.P. Barros and N. Peypoch, (2008), "A comparative analysis of productivity change in Italian and Portuguese airports", *International Journal of Transport Economics*, XXXV (2), 205–216.

[107] G. Martini, A. Manello and D. Scotti (2013), "The influence of fleet mix, ownership and LCCs on airports' technical/environmental efficiency", Transportation Research, Part E: Logistics and Transportation Review, 50, 37-52.

8.5.7. Regulatory initiatives

As anticipated, ART has been entrusted with the application of the implementing rules of EU Directive 2009/12, with the exception of the 5 largest airports supervised by ENAC. In order to facilitate negotiations between airport managers and users of airport facilities, already in 2014 the Authority has developed three different models for the adoption of airport charges, depending on the size of the airport, i.e. above 5 million passengers, between 3 and 5 million and below 3 million (Decision no. 64/2014).

The choice of whether to adopt single-till, hybrid till or dual-till approaches, i.e. to take into account revenues from commercial activities in establishing airport charges, is left to airport managers. Airport managers adopt charging schemes after negotiating with airport users but the Authority is empowered to adopt corrective measures with respect to the proposed scheme when needed to ensure productive efficiency and contain costs for users. A description of the three models, whose complexity decreases for smaller airports, can be found in Cambini-Perrotti (2015).[108] The negotiation process has turned out to be rapid, lasting only a few months. [109]

In the mid term review of the models, the Authority observed that in the airports supervised by ART, the current charging schemes have spurred investment plans for more than 1 billion euro in four years, with 80 per cent financed by private capital. Moreover, it announced the intention to enhance the role of benchmarking and competition by comparison on the basis of estimates of efficient frontiers in order to strengthen incentivizing factors; a revision of the models along these lines has already been launched.

With respect to the provision of non-aviation activities, the Authority commits to further analyse efficiency and competition issues in order to ascertain whether any efficiency enhancing regulatory measure would be justified.

In the meantime, the 2009/12 Directive is being revised by the European Commission. One of the main goals of the revision is increasing the effectiveness of the enforcement procedures, since the approach based on consultation between airport managers and airport users has turned out to be quite ineffective, paying more attention to the process than to what should be the objective of the consultation. A more focused regulation for airports with the higher market power is being advocated, as well as greater recognition of the link between airport charges and service quality.[110] The outcome of the ongoing debate at the EU level on the proper approach to airport charging policies will obviously have an impact also on the choices at the national level.

8.5.8. Benchmarking for seaports and airports: main conclusions

Summing up, both for Italian seaports and airports market studies are crucial to understand the market evolution and how it affects the need for regulatory intervention. In particular, policy makers should refrain from imposing regulatory measures for the provision of services unless they are necessary and proportionate in the public interest.

For seaports, the main challenge for regulation is to promote efficiency-enhancing conditions of access to port areas and infrastructures which contribute to the maintenance and development of infrastructures

[108] C. Cambini and L. Perrotti, (2015), "The New Transport Regulation Authority in Italy: Structure, Competencies and First Regulatory Decisions", *Network*, 54, March. For a more general discussion of the use of price cap models in the regulation of airports in Italy, see also F. Ramella, (2017), *La regolazione di autostrade e aeroporti. Efficienza e investimenti*, TRASPOL Report 3/2017, Milano, Italy.
[109] On the impact of the models, see the ART 2018 Annual Report, p. 25.
[110] European Commission, *Support Study to the Ex Post Evaluation of Directive 2009/12 on airport charges. Final Report*, Steer Davies Glear, December 2017.

and to the efficient provision of services. Estimates of minimum efficient scale for the different services may support pro-competitive franchise policies in the port areas. The traditional fragmentation of regulatory approaches in the port sector has represented a major weakness. Thus, ART may play a useful role in promoting a common approach to accounting, admissible costs, etc. so as to improve transparency, comparability and spread of best practices.

For airports, the main challenge in the ongoing revision of models for airport charges within ART's jurisdiction is to strengthen efficiency enhancing incentives compared to the first version of such models by means of estimates of efficient frontiers.

Chapter references

Abrate, G., Erbetta, F., (2010*), "*Efficiency and patterns of service mix in airport companies: An input distance function approach", *Transportation Research Part E: Logistics and Transportation Review*, 46 (5), 693-708.

AGCM, (1998), AS 135, *Opinion on the extension of highway concessions*, 22 May 1998.

AGCM, (2016), *Sectoral Inquiry on Local Public Transport*.

Antonellini, L., Costa, P., Munari, F., Spirito, P., (2017), "Riforma e problematiche della portualità"*,* in S.I.Po.Tra, *Le politiche dei trasporti in Italia. Rapporto 2017*, 271-308

ART, Annual Report 2018.

Avenali, A., Boitani, A., Catalano, G., D'Alfonso, T., Matteucci, G., (2016), "Assessing standard costs in local public bus transport: Evidence from Italy", *Transport Policy*, 52, 164-174.

Barros, C.P., Dieke, P.U.C., (2008), "Measuring the economic efficiency of airports: a Simar–Wilson methodology analysis", *Transportation Research Part E*, 44, 1039–1051.

Barros, C.P., Peypoch, N., (2008), "A comparative analysis of productivity change in Italian and Portuguese airports", *International Journal of Transport Economics*, XXXV (2), 205–216.

Boitani, A., (2016), *I costi standard del trasporto ferroviario e la concorrenza per il mercato*, Sipotra.

Cambini, C., Perrotti, L., (2015), "The New Transport Regulation Authority in Italy: Structure, Competencies and First Regulatory Decisions", *Network*, 54, March. For a more general discussion of the use of price cap models in the regulation of airports in Italy, see also Ramella, F., (2017), *La regolazione di autostrade e aeroporti. Efficienza e investimenti*, TRASPOL Report 3/2017, Milano, Italy.

CER (2012), *The effect of vertical separation in the railway sector.*

Curi, C., Gitto, S., Mancuso, P., (2008), "An application of Data Envelopment Analysis (DEA) to measure the efficiency of the Italian airports after the privatisation", *L'Industria*, 4, 689–712.

Curi, C., Gitto, S., Mancuso, P. (2011), "New evidence on the efficiency of Italian airports: a bootstrapped DEA analysis", *Socio-Economic Planning Sciences*, 45 (2), 84-93.

European Commission, (2017), *Support Study to the Ex Post Evaluation of Directive 2009/12 on airport charges. Final Report*, Steer Davies Glear, December.

Gitto, S., Mancuso, P., (2012a), "Two faces of airport business: A non-parametric analysis of the Italian airport industry", *Journal of Air Transport Management*, 20, 39-42.

Gitto, S., Mancuso, P., (2012b), "Bootstrapping the Malmquist indexes for Italian airports", *International Journal of Production Economics*, 135 (1).

Karantounias. V., Pinelli, D., (2016), *Local State-Owned Enterprises in Italy: inefficiencies and ways* forward, European Economy, Economic Brief 10, April 2016.

Malighetti, P., Martini, G., Paleari, S., Redondi, R., (2007), "An empirical investigation on the efficiency, capacity and ownership of Italian airports", *Rivista di Politica Economica*, 97, 57–188.

Martini, G., Manello, A., Scotti, D., (2013), "The influence of fleet mix, ownership and LCCs on airports' technical/environmental efficiency", *Transportation Research Part E: Logistics and Transportation Review*, 50, 37-52.

Merkert, R., Mangia, L., (2014), “Efficiency of Italian and Norwegian airports: A matter of management or of the level of competition in remote regions?”, *Transportation Research Part A: Policy and Practice*, 62, 30-38.

Ponti, M., Ramella F. (2018), *Trasporti: conoscere per decidere*, Egea

Ramella, F., (2017), *La regolazione di autostrade e aeroporti. Efficienza e investimenti*, TRASPOL Report 3/2017, Milano, Italy.

Sebastiani, M., (2017), “La questione dimensionale nel settore ferroviario e dintorni”, in S.I.Po.Tra, *Razionalizzazione dei mercati e aggregazioni fra imprese di trasporto*, Quaderno n. 2/2017, (http://www.sipotra.it/wp-content/uploads/2017/04/Quaderno-AGGREGAZIONI-FINALE.pdf).

Sestito, P., (2015), *Hearing before the 8th Commission of the Chamber of Deputies within a Sectoral Inquiry on highway concessions*, https://www.bancaditalia.it/pubblicazioni/interventi-vari/int-var-2015/sestito-audizione-110615.pdf.

9. CONCLUSIONS

Chris Nash

Most economists have a preference for competition in or for the market where it is practicable. However, transport infrastructure displays strong natural monopoly characteristics, so competition in the market is often not feasible. Competition for the market does exist, through tenders for the provision of transport services, concessioning of whole sections of infrastructure, or contracting out construction and maintenance, although it is not widespread and does not always work well. In practice, provision of transport infrastructure and some services remains dominated by public sector monopolies.

This means that there is a strong need for the use of benchmarking to support yardstick competition by providing estimates of efficient costs to use in setting targets.

However, benchmarking is also of value where there is competition for the market, in that it may be used to determine the characteristics of efficient contracts, such as size, length, gross or net cost and conditions attached. Moreover, benchmarking may also be useful where there is competition in the market if there is any doubt as to its effectiveness, as it may suggest where there are areas of inefficiency.

In the past, benchmarking often relied on crude comparisons of unit cost, with an attempt to ensure that the decision making units being compared were broadly comparable. There are now benchmarking techniques available which can deal with issues such as multiple outputs and inputs and heterogeneity of operating conditions which are ignored by simple measures. The biggest problem is the quality and quantity of data. Benchmarking studies using these techniques exist for all the main modes of transport and the technique is already being used in some cases by regulators.

In Italy, there is scope for the use of benchmarking in all modes covered by ART. We believe that there is much more scope for the use of competitive tendering in awarding contracts than is currently the case in Italy, and benchmarking may help determine such factors as the appropriate size and length of concessions. But where competitive tendering is not introduced, or until it is, then benchmarking as a way of informing setting of targets for efficiency (yardstick competition) is a crucial tool. The use of 'standard costs' on negotiating contracts in local public transport is an important step in the right direction, but we believe further work is needed in order to ensure that these are derived using the most appropriate techniques.

ANNEX: CONCESSIONS OF TOLL HIGHWAYS IN ITALY (2017)

Code	Highway	In operation [km]	Under construction[km]	Planned [km]	Total [km]	Concessionaire
A5	Aosta-Traforo Monte Bianco	32,4	–	–	32,4	Raccordo Autostrada Valle d'Aosta S.p.A.
A5	Quincinetto-Aosta	59,5	–	–	59,5	Società Autostrade Valdostane p.A.
A5	Raccordo A5-S.S. 27 del G.S. Bernardo	7,9	–	–	7,9	Società Autostrade Valdostane p.A.
A5	Torino-Ivrea-Quincinetto	51,2	–	–	51,2	Autostrada Torino-Ivrea-Valle d'Aosta S.p.A.
A4-A5	Ivrea-Santhià	23,6	–	–	23,6	Autostrada Torino-Ivrea-Valle d'Aosta S.p.A.
A32	Torino-Bardonecchia	75,7	–	–	75,7	Soc. Italiana per il Traforo Aut.le del Frejus p.A.
A6	Torino-Savona	130,9	–	–	130,9	Autostrada Torino-Savona S.p.A.
A55	Sistema Tangenziale di Torino	81	–	–	81	Autostrada Torino-Ivrea-Valle d'Aosta S.p.A.
A21	Torino-Piacenza	164,9	–	–	164,9	S.A.T.A.P. S.p.A.
A33	Asti-Cuneo	55,7	–	34,5	90,2	Autostrada Asti-Cuneo p.A.
A26	Voltri-Gravellona Toce e diramaz. per Bettole, Santhià	244,9	–	–	244,9	Autostrade per l'Italia S.p.A.

Code	Highway	In operation [km]	Under construction[km]	Planned [km]	Total [km]	Concessionaire
A8	Milano-Varese	45,3	–	–	45,3	Autostrade per l'Italia S.p.A.
A8-A26	Diramazione Gallarate-Gattico	24	–	–	24	Autostrade per l'Italia S.p.A.
A9	Lainate-Como-Chiasso	32,4	–	–	32,4	Autostrade per l'Italia S.p.A.
A36	Dalmine-Como-Varese-Valico, dir. A8 e Tang.le Est di Milano	30,2	–	64,2	94,4	Autostrada Pedemontana Lombarda S.p.A.
A58	Tangenziale Est Esterna di Milano (TEEM)	33	–	–	33	Tangenziale Esterna S.p.A.
A59	Tangenziale di Como	4,3	–	–	4,3	Autostrada Pedemontana Lombarda S.p.a.
A60	Tangenziale di Varese	7	–	–	7	Autostrada Pedemontana Lombarda S.p.a.
A50	Tang.le Ovest di Milano	33	–	–	33	MI Serravalle-MI Tangenziali S.p.A.
A51	Tang.le Est di Milano	29,4	–	–	29,4	MI Serravalle-MI Tangenziali S.p.A.
A52	Tang.le Nord di Milano	12,9	6	–	18,9	MI Serravalle-MI Tangenziali S.p.A.
A54	Tang.le di Pavia	8,4	–	–	8,4	MI Serravalle-MI Tangenziali S.p.A.
A53	Raccordo Bereguardo-Pavia (b)	9,1	–	–	9,1	MI Serravalle-MI Tangenziali S.p.A.
A7	Milano-Serravalle	86,3	–	–	86,3	MI Serravalle-MI Tangenziali S.p.A.
A7	Serravalle-Genova	50	–	–	50	Autostrade per l'Italia S.p.A.

Code	Highway	In operation [km]	Under construction[km]	Planned [km]	Total [km]	Concessionaire
A1	Milano-Napoli (c)	803,5	–	–	803,5	Autostrade per l'Italia S.p.A.
A4	Torino-Milano	127	–	–	127	S.A.T.A.P. S.p.A.
A35	Brescia-Milano	62,1	–	–	62,1	Società di Progetto Brebemi S.p.A.
A4	Milano-Bergamo-Brescia	93,5	–	–	93,5	Autostrade per l'Italia S.p.A.
A4	Brescia-Verona-Vicenza-Padova	146,1	–	–	146,1	Autostrada Brescia-Verona-Vicenza-Padova S.p.A.
A4	Padova est-bivio A4\A57	14,3	–	–	14,3	Concessioni Autostradali Venete - C.A.V. S.p.A.
A4	Bivio A4\A57 - Quarto d'Altino	32,3	–	–	32,3	Concessioni Autostradali Venete - C.A.V. S.p.A.
A57	Tangenziale di Mestre Raccordo tangenziale Mestre Aeroporto Venezia (Tessera)	18,1 9,4	–	–	18,1 9,4	Concessioni Autostradali Venete - C.A.V. S.p.A.
A27	(VE) Mestre-Belluno	82,2	–	–	82,2	Autostrade per l'Italia S.p.A.
A4	(VE) Mestre-Trieste	115,4	–	–	115,4	Autovie Venete S.p.A.
A57	Tangenziale di Mestre	10,5	–	–	10,5	Autovie Venete S.p.A.
A23	Palmanova-Udine	18,5	–	–	18,5	Autovie Venete S.p.A.
A28	Portogruaro-Conegliano	48,8	–	–	48,8	Autovie Venete S.p.A.
A34	Villesse-Gorizia	17	–	–	17	Autovie Venete S.p.A.
A23	Udine-Carnia-Tarvisio	101,2	–	–	101,2	Autostrade per l'Italia S.p.A.
A21	Piacenza-Brescia	88,6	–	–	88,6	Autostrade Centro Padane S.p.A.

Code	Highway	In operation [km]	Under construction[km]	Planned [km]	T otal [km]	Concessionaire
A22	Brennero-Verona-Modena	314	–	–	314	Autostrada del Brennero S.p.A.
A31	Valdastico	89,5	–	39,1	128,6	Autostrada Brescia-Verona-Vicenza-Padova S.p.A.
A15	Parma-La Spezia e completamento e coll.to con la A22	101	14	69	184	Autocamionale della Cisa S.p.A.
A13	Bologna-Padova	127,3	–	–	127,3	Autostrade per l'Italia S.p.A.
A14	Bologna-Taranto	781,4	–	–	781,4	Autostrade per l'Italia S.p.A.
A10	Ventimiglia-Savona	113,3	–	–	113,3	Autostrada dei Fiori S.p.A.
A10	Savona-Genova	45,5	–	–	45,5	Autostrade per l'Italia S.p.A.
A12	Genova-Sestri Levante	48,7	–	–	48,7	Autostrade per l'Italia S.p.A.
A12	Sestri Levante-Livorno e diramaz. per Lucca e La Spezia	154,9	–	–	154,9	Società Autostrada Ligure Toscana p.A.
A12	Livorno-Civitavecchia	54,6	–	187,4	242	Società Autostrada Tirrenica p.A.
A12	Civitavecchia-Roma	65,4	–	–	65,4	Autostrade per l'Italia S.p.A.
A11	Firenze-Pisa Nord	81,7	–	–	81,7	Autostrade per l'Italia S.p.A.
A24	Roma-L'Aquila-Teramo	166,5	–	–	166,5	Strada dei Parchi S.p.A.
A25	Torano-Avezzano-Pescara	114,9	–	–	114,9	Strada dei Parchi S.p.A.
A56	Tangenziale Est-Ovest di Napoli	20,2	–	–	20,2	Tangenziale di Napoli S.p.A.
A16	Napoli-Canosa	172,3	–	–	172,3	Autostrade per l'Italia S.p.A.

Code	Highway	In operation [km]	Under construction[km]	Planned [km]	Total [km]	Concessionaire
A30	Caserta-Nola-Salerno	55,3	–	–	55,3	Autostrade per l'Italia S.p.A.
A3	Napoli-Pompei-Salerno	51,6	–	–	51,6	Autostrade Meridionali S.p.A.
A20	Messina-Palermo	181,8	–	–	181,8	Consorzio per le Autostrade Siciliane
A18	Messina-Catania	76,8	–	–	76,8	Consorzio per le Autostrade Siciliane
A18	Siracusa-Gela	39,8	19,7	73,3	132,8	Consorzio per le Autostrade Siciliane
Toll highways TOTAL		**5.978**	**39,7**	**467,5**	**6.485**	

Code	Highway	In operation [km]	Under construction[km]	Planned [km]	Total [km]	Concessionaire
T1	Traforo del Monte Bianco (tratta di competenza italiana)	11,6	–	–	11,6	Soc. Italiana per il Traforo del M. Bianco p.A.
T2	Traforo del Gran S. Bernardo (tratta di competenza italiana + compresa autostrada di accesso)	5,8	–	–	5,8	Soc. Italiana per il Traforo del G.S. Bernardo p.A.
T4	Traforo del Fréjus (tratta di competenza italiana)	12,8	–	–	12,8	Soc. Italiana per il Traforo Austostradale del Fréjus p.A.
Toll tunnels TOTAL		**30,2**	-	-	**30,2**	

Code	Highway	In operation [km]	Under construction[km]	Planned [km]	Total [km]	Concessionaire
A90	Grande raccordo anulare di Roma (GRA)	68,2	–	–	68,2	A.N.A.S.
A91	Roma - Aeroporto di Fiumicino	17,4	–	–	17,4	A.N.A.S.
A3	Salerno - Reggio Calabria	443	–	–	443	A.N.A.S.
A29	Palermo - Mazara del Vallo e diramaz. Punta Raisi	118,8	–	–	118,8	A.N.A.S.
A29 DIR	Dir - Alcamo - Trapani e diramaz. per Birgi	50	–	–	50	A.N.A.S.
A19	Palermo – Catania	192,8	–	–	192,8	A.N.A.S.
A19 DIR	Diramazione per via Giafar	5,2	–	–	5,2	A.N.A.S.
A29 RACC. BIS	Raccordo per Via Belgio	5,6	–	–	5,6	A.N.A.S.
A18 DIR	Catania nord - Catania Centro	3,7	–	–	3,7	A.N.A.S.
	Sistiana – Rabuiese	10,1	–	–	10,1	A.N.A.S.
	Catania – Siracusa	25,1	–	–	25,1	A.N.A.S.
A.N.A.S. TOTAL		**939,9**	–	–	**939,9**	
Overall TOTAL		**6.943,2**	–	–	**7.455,1**	

Source: AISCAT 2018

Autorità
di Regolazione
dei Trasporti

Via Nizza 230 – 10126 Torino
www.autorita-trasporti.it