Performance Review Body
Advice on the Union-wide target ranges for RP4

Annex II - Academic study on cost-efficiency

September 2023

Performance Review Body of the Single European Sky | Rue de la Fusée 96, Office 50.659, 1130 Brussels
Office Telephone: +32 (0)2 234 7824 | cathy.mannion@prb.eusinglesky.eu | prb-office@prb.eusinglesky.eu | wikis.ec.europa.eu/display/eusinglesky

Academic Group

Nicole Adler and Gianmaria Martini

Nicole Adler, Hebrew University of Jerusalem
Gianmaria Martini, University degli Studi di Bergamo

31 July 2023

Executive Summary

a. The cost efficiency of Air Navigation Service Providers (ANSPs) is an important element in the development of an efficient Single European Sky. Each ANSP serves an individual airspace and in so doing is a natural monopoly. Since there is little direct competition in the market, efficiency is not encouraged by sound competitive pressure.

b. Benchmarking can provide a useful substitute for such settings. Benchmarking allows to identify best practices, and if ANSPs are asked over time to adjust to best-practice levels, their cost efficiency will converge as if they are working in a competitive setting. Hence, instead of competing in the market, it is possible to create pseudo competition via benchmarking-based regulation, whereby the ANSPs compete via a model.

c. Two such benchmarking models are implemented and the results are subsequently combined. One is based on data envelopment analysis (DEA) and another on stochastic frontier analysis (SFA). They may be combined in different ways (minimum or maximum estimated scores or an average between the measures) and across different time periods in order to determine cost-efficiency targets.

d. The report analyzes en-route activities only rather than gate-to-gate provision. En-route provision has remained a monopolistic service provided by a single ANSP in each Member State (with the only exception being MUAC).

e. We present the Union-wide estimated efficiency scores after accounting for negative externalities (i.e., delays) and the operational environment (i.e., variability (seasonality), and complexity).

f. Union-wide, the DEA model presents estimated efficiency levels of approximately 79%, while the SFA model estimates efficiency levels of 89%. The weighted average therefore suggests potential efficiency levels of 84%.

g. We find that the ANSPs could save just under one billion euros annually by adjusting to best practices (based on the 2019 PPP-adjusted costs). However, there are substantial differences in potential cost saving levels across the individual ANSPs. It is therefore natural to work not only with a general cost reduction requirement that captures technological progress, but also to work with additional individual requirements encouraging the less efficient ANSPs to catch-up to best practices.

h. It is noteworthy that there seems to be a significant reduction in efficiency between RP1 (covering 2012 to 2014 inclusive) and RP2 (covering 2015 to 2019). Concerns are also raised over the reporting of capital expenditures, suggesting the possibility of some data manipulation or gaming, which presents a challenge for the regulatory authorities. In setting the x% savings target for the RP4 period, it is important to guarantee that the 1% annual cost efficiency improvements realized over the eight years analyzed is not negated in the process.

Contents

List of Tables

List of Figures

1. Introduction

1.1. The task of the academic group (AG) is to provide advice to the Performance Review Body (PRB) on target setting for cost efficiency for Reference Period 4 (RP4, years 2025-2029).

1.2. The object of this document is to provide a report and a model to the PRB with meaningful and scientifically robust Union-wide targets on cost efficiency on a benchmarking of ANSPs from the efficient cost frontier based on proven models. These targets take into account the estimates of inefficiency provided by the AG.

1.3. The document reports the data analysis by the AG, the steps implemented to construct the variables for the empirical analyses, the models estimated to study the Union-wide efficiency of ANSPs, the descriptive statistics regarding the variables included in the empirical analyses, and the set of results assessing ANSP efficiency during the period of observation. Finally, we suggest a possible range of improvements relevant for RP4.

1.4. The Academic Group has been tasked with delivering a report and modeling methodologies to the PRB, offering scientifically robust benchmarking of ANSPs' cost efficiency.

2. Cost efficiency

2.1. The principle of cost efficiency is broadly employed in business management and regulatory agencies. Intuitively, the concept revolves around the capability of producing a specified volume of product or service at the lowest possible cost. This practice ensures that a company optimizes its resources and avoids wastage. In effect, any increase in costs implies expenses exceeding what is deemed necessary. For this reason, in the realm of business management, the objective of cost minimization is continuously monitored and assessed, utilizing a variety of tools.

2.2. The costs of a business are tied to the acquisition of factors necessary for the production of a good or service. Consequently, management strategies aim to monitor the costs by checking the utilization of production factors, also referred to as inputs, according to economic theory.

2.3. Inputs utilized by a company encompass a wide range of elements, including personnel, facilities, buildings, computers, energy, raw materials, etc. For purposes of simplification, economic theory generally groups them into two broad categories: labor and capital. Inputs that are exhausted after usage fall under the labor classification, for example an hour of work once performed is irrecoverable. The consumption of 10 kWh of electricity cannot be reused. Conversely, the capital category includes all inputs that are not immediately depleted upon usage. Examples of capital inputs include computers, buildings, plants, radars, etc. However, the capital goods will gradually become obsolete over time, necessitating maintenance or upgrades, and eventually replacement, for example with newer generation models over time.

2.4. In business management, monitoring is typically conducted using simple, easily calculable indices that are quickly updated. These are the key performance indicators (KPIs), which are usually calculated and monitored according to the two broad input categories described above.

2.5. In the context of the labor production factor, KPIs are typically calculated based on the product per employee, or the product per hour worked, etc. A similar process is undertaken, albeit less frequently, for the capital factor of production. A common KPI in this case is the volume of product produced per hour of use of the facility, or by the value of the fixed assets indicated in the financial statements. In terms of cost efficiency, a typical KPI is the labor cost per full-time equivalent employee.

2.6. While these KPIs are extensively utilized by managers, they do present an issue: they represent partial measures of efficiency. They concentrate on a single input and neglect the contribution of the other input to production. For example, one firm may have lower unit labor costs than another, thereby appearing more efficient. However, this may be attributable to a larger endowment of capital, which increases the volume of output, rather than higher worker productivity. If one also considers the capital endowment, it could reveal that the second firm is more efficient than the first.

2.7. For this reason, the correct measure for calculating cost efficiency is based on methods that take into account all production factors. In order to estimate whether an ANSP is carrying out the activity cost-effectively, a computational method is therefore needed that takes into account both categories of inputs, i.e., labor and capital, the volume of output that is produced, and whether this is achieved at minimum cost.

2.8. This method incorporates two dimensions: the costs of the company are determined by the expenditure on production factors, namely the cost of labor and the cost of capital. The first dimension therefore concerns whether the firm's labor and capital endowment are the minimum necessary to achieve a specific volume of output. In economic theory, this dimension is called technical efficiency, and is represented in Figure 1. The second dimension involves the optimal mix of labor and capital, according to their individual prices. This is referred to as cost efficiency, and is represented in Figure 2.

Technical efficiency

2.9. The initial step in determining technical efficiency is to establish the level of output. This is represented by the isoquant, which illustrates the combinations of labor (L) and capital (K) that could potentially be sufficient to produce the target level of production, such as the number of flight hours controlled in a year. The second step involves identifying the ANSPs position with respect to the isoquant. In Figure 1, two ANSPs are depicted, one with a black dot and the other with a red dot. The black dot lies on the isoquant, suggesting that the ANSP operates relatively efficient. Conversely, the red dot is located above the isoquant, indicating that the ANSP is producing the same number of controlled flight hours as the black dot ANSP but utilizes more labor and capital. This is indicative of technically inefficiency. The level of technical inefficiency may be estimated by the vertical segment projecting the red dot onto the isoquant.

Cost efficiency

2.10. The second step consists of estimating cost efficiency, as shown in Figure 1. In this figure, we depict three ANSPs as black, red and blue dots. The target level of production is identified by the isoquant (as in Figure 1), and an isocost function is also depicted. The isocost is depicted as a straight line, which represents all possible input combinations that yield the same cost. The gradient of the isocost function is determined by the cost of labor and the cost of capital, where the former is given by the price of labor times the amount of labor used by the ANSP, and the latter by the price of capital times the amount of capital available. Hence, cost efficiency takes into account the costs of the inputs.

2.11. In Figure 2, three isocost lines are illustrated, one in bold and two in dashed lines. The bold line, extending from the origin of the graph towards the top right, indicates the lowest cost, while the two dashed lines signify higher costs. Hence, the ANSP represented by the blue dot incurs the highest cost because the isocost line passing through it is the highest. The ANSP denoted by the black dot is cost-efficient for two reasons. First, it is technically efficient i.e., it lies ON the isoquant. Second, it operates at minimum cost because, given the price levels of the inputs (denoted by the slope of the isocost line), it employs the optimal combination of inputs to produce the target output, i.e., it is located on the isocost line tangent to the isoquant. The blue-dot ANSP is technically efficient, as it lies on the isoquant, meaning it is not using more inputs than necessary to meet the output target. However, it is not utilizing the best combination of inputs relative to their prices. The isocost line passing through the blue dot intersects the isoquant, indicating that this dashed-line isocost is higher than the bold one. Thus, the blue-dot ANSP is technically efficient but cost-inefficient, i.e., it is not operating at minimum costs. The measure of its cost inefficiency is indicated by the vertical segment in Figure 2. The red-dot ANSP is both technically and cost inefficient because it lies above the isoquant, and the isocost line passing through it is highest.

2.12. The definition of cost efficiency provided by economic theory implies that the ANSP selects the input mix (i.e., the combination of labor and capital) that yields the minimum expenditure for the required level of operations, given the current level of input prices.

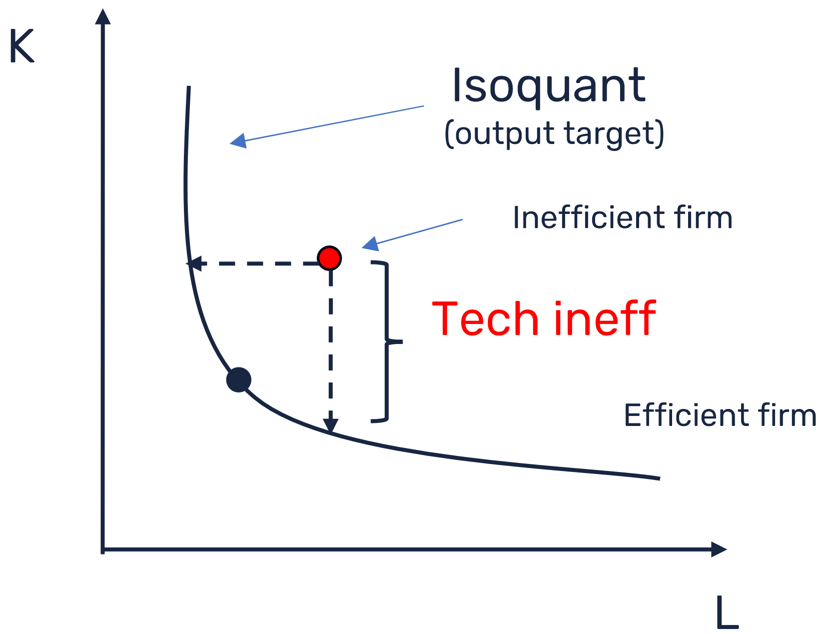

Figure 1 - Technical efficiency

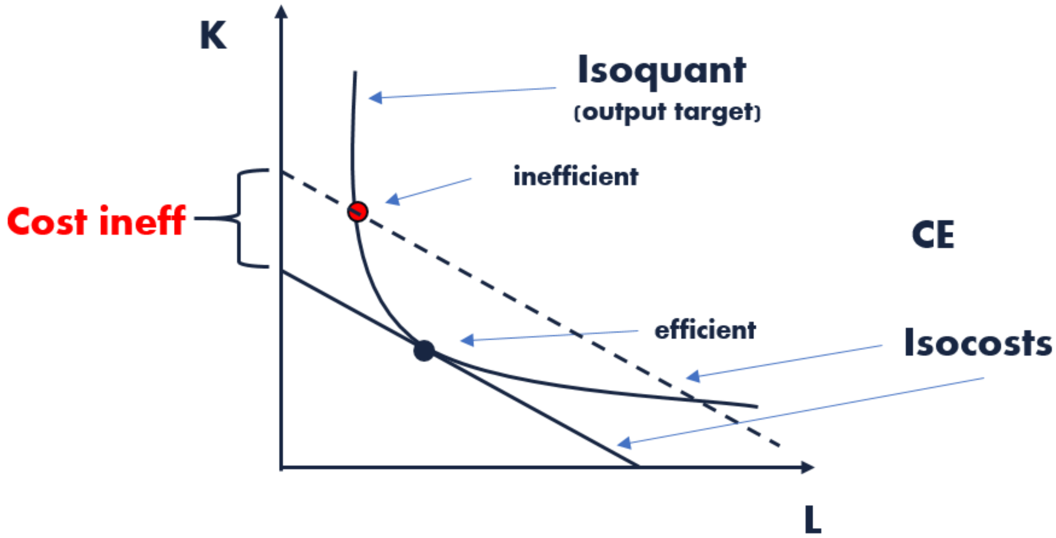

Figure 2 - Cost efficiency

From data to cost functions

2.13. Estimates of ANSP cost efficiency are obtained from the observed data, including input quantities, output, input prices and other factors that may influence the ability of the ANSPs to operate at the minimum level of costs, as depicted in Figure 3.

2.14. Operations may be influenced by exogenous factors that are beyond the control of management. Such factors include the possible presence of random shocks, such as striking air traffic controllers in another country or a volcanic eruption. Furthermore, non-random factors may impact the ability of management to minimize costs, such as seasonality which necessitates adequate staff and capital levels to handle peak traffic during specific periods of the year. Finally, the quality of management also influences costs, as it impacts the level of effort required to attain the minimum cost level. It is important to note that only this last component signifies true cost inefficiency.

2.15. The methodology outlined in Figure 3 is applied to the observed variables describing the ANSPs operations. Since these data points pertain to costs, they require standardization. This is necessary as costs are measured in different currencies and span various time periods that may be affected by inflation. Moreover, the purchasing power of different ANSPs must be considered due to variations in input prices across different countries.

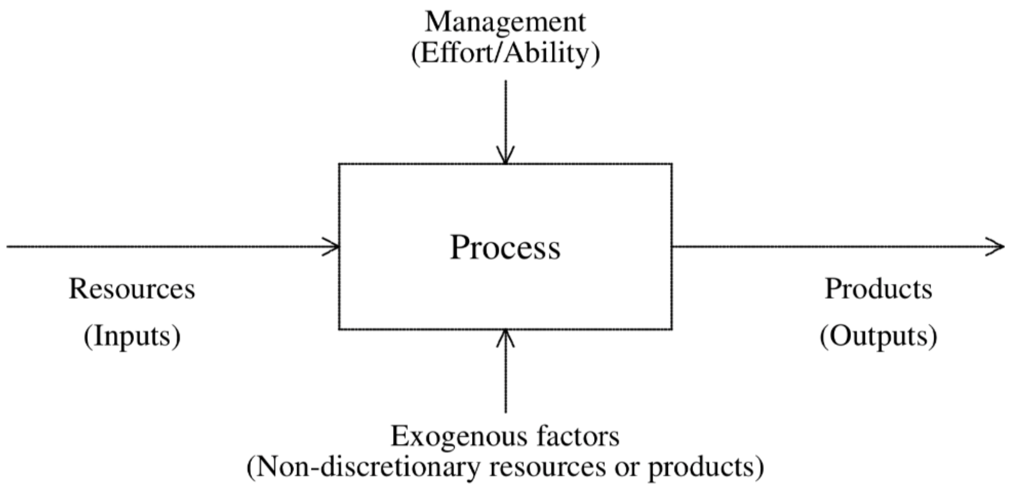

Figure 3 - Cost efficiency model

Regulatory benchmarking

2.16. The evaluation of cost-efficiency is a common objective in regulated sectors. In these sectors, public agencies regulate the market due to the potentially significant market power of a company. Regulated sectors are characterized by the presence of natural monopolies, i.e., single companies that control the entire market. ANSPs, in fact, are local natural monopolies. They are the sole organizations that control air traffic in a specific country or territory. In other sectors, such as electricity transmission, gas, water, telecommunications and transport networks, activities are typically concentrated in the hands of a small number of companies due to economies of scale. In these sectors, costs are characterized by a high proportion of fixed costs, leading to decreasing unit costs. By centralizing all activity within a single firm, the provision of various products or services are achieved at the lowest unit costs.

2.17. To prevent monopolies from reducing production and/or inflating prices due to the lack of competition, regulatory agencies oversee the costs of the company by defining supply prices, also known as tariffs. These regulatory bodies ensure that the tariffs cover production costs while also providing a reasonable return on invested capital. For the purposes of tariff setting, regulatory agencies need estimates of the levels of cost-efficiency of such natural monopolies. This is necessary for achieving two objectives. First, to establish tariffs that ensure a reasonable level of quality. Second, to incentivize cost-inefficient monopolies to exert effort towards achieving efficiency.

2.18. There are multiple regulatory methods, including cost-plus regulation, price cap regulation, yardstick competition and concessions through auctioning.

2.19. Cost-plus regulation is based on the idea that the regulated monopoly should only reach the break-even point where the regulated price (the tariff) is equal to the average costs (Alexander and Irwin, 1996). In this case the regulatory agency must know the economic costs of the monopolist, including the opportunity cost of capital invested. The regulation model is such that the tariff is computed by aggregating two elements: (1) the average monetary costs, observed from operations, and (2) a fair rate of return on the capital invested. The cost-plus method does not provide strong incentives toward cost efficiency. The management knows that any cost level will be covered by the tariff and has no strong incentive to reduce costs by cutting inefficiencies. Furthermore, increasing investments will also be covered by the tariff through the granted return on the capital. Hence, the cost-plus regulation model leads to over-investment.

2.20. Price cap regulation is a different approach in which the regulatory agency sets the price levels that the monopolist can charge over the next four to five years (Alexander and Irwin, 1996). The pattern of prices decreases over the subsequent years, providing the monopolist with an incentive to reduce its costs. Moreover, if the monopolist reduces costs beyond that defined by the regulatory body, it will gain profits. For example, if the price in a given year is 3% lower than the previous year, and the costs are decreasing by 5%, the 2% difference translates into an increased surplus for the monopolist. This perspective fosters efficiency and innovation incentives. However, there is a downside because the monopolist has an incentive to cut costs, which may lead to under-investment.

2.21. Yardstick competition is a price regulation scheme in which the regulated price established for a given firm is derived from the cost structure of similar firms operating in different niches (Shleifer, 1985). The approach requires extensive data and can be challenging to implement because identifying comparable benchmarks may prove difficult.

2.22. The final model of regulation is based on auctions. The regulatory agency grants the right to manage a sector for a specified period of time to the company that wins the auction based on the best bid. The auction can be designed in different ways, such as the English style (where the winner is the one who places the highest bid in an ascending order auction), the Dutch style (which is similar to the English auction, but with bids in descending order), the first-price sealed-bid (where the winner is the one who places the highest bid in a scenario where each firm places a single bid confidentially), and the second-price sealed-bid auction (Vickrey, 1961), in which the winner is the one with the highest bid but pays the second highest price offered. Concessions through auctions are typically implemented for long-term periods, such as 20 or 30 years. It is noteworthy that auctions are now being used at various airports across Europe for the selection of a terminal ANSP provider.

2.23. The regulatory model promoting cost efficiency in the Union-wide ANSP sector presently utilizes a price cap approach. This approach dictates an annual percentage of inefficiency that must be addressed. The incentive aspect of the ANSPs' regulatory framework sets a target for an annual percentage reduction in costs over the five year review period.

3. Empirical methods

3.1. Benchmarking methods, and in particular Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA), have become well-established and informative tools for economic regulation. DEA and SFA are now routinely used by European regulators to set reasonable revenue and price caps for energy transmission and distribution system operators for example. The application of benchmarking in regulation, however, requires specific steps in terms of data validation, model specification and outlier detection that are not systematically documented in open publications.

3.2. We note that the Performance Review Unit (PRU) of Eurocontrol has been collecting data systematically on ANSP services since 2002. Furthermore, since year 2012, Member States submit cost data to the European Commission as defined by the Single European Sky (SES) framework. Substantial work on data verification is undertaken, leading to the likelihood that the information for the timeframe analyzed (2012 to 2019 inclusive) should be reasonably reliable.

3.3. In this chapter, we explain the modern foundations for frontier-based regulation, and we discuss its use in the present project aimed at regulating en-route ANSP charges.

Benchmarking

3.4. In the business world, benchmarking is traditionally thought of as a managerial tool that helps improve performance by identifying and quantifying the impact of applying best documented practice. Managers compare the performance of their respective organizations, products and processes externally with competitors and best-in-class companies and internally with other operations within their own organizations that perform similar activities.

3.5. The idea of best practice is important. In benchmarking the idea is not to compare existing organizations to some theoretical ideal or green-field solution. Rather, the idea is to use best realized practice as the benchmark. This naturally implies that the benchmarking targets are achievable, relative to the comparators and evolving from the action of the firms. Consequently benchmarking in both models applied here are reasonably conservative since they estimate only relative efficiency.

Key Performance Indicators

3.6. Traditionally benchmarking focuses on key performance indicators (KPIs). KPIs are ratio numbers that are assumed to reflect the purpose of the ANSP in some essential way. KPIs are widely used by operators, shareholders, regulatory agencies, researchers and others with an interest in performance evaluation. Well-known KPIs are related to the analysis of financial accounts. They include indicators like Return on Investments (=net income/total assets), gross margin, etc.

3.7. Unfortunately, the use of KPIs has its limits. First, when we compare a small ANSP to a large ANSP on a ratio (say support staff cost per flight hour controlled), we implicitly assume that we can scale input and output proportionally. That is, we assume constant returns to scale. A second limitation of the KPI approach is that it typically involves only partial evaluations. One KPI seldom reflects the purpose of the ANSP. We may have multiple inputs and outputs and therefore form several output-input ratios each of which provides an incomplete representation of the ANSP. KPIs in this case do not account for substitution between inputs and between outputs.

3.8. A third limitation is that KPIs seldom capture the allocation properly. One ANSP may be better in all conceivable sub-processes and still be inferior by relying more on the relatively less efficient processes.

Model based

3.9. For these reasons, advanced benchmarking is model based. We try to account for multiple effects that may interact in complicated ways. To handle this, we use a systematic approach to the ANSP. An ANSP is seen as a transformation of multiple resources into multiple products and services. The transformation is affected by non-controllable factors as well as by non-observable skills deployed and efforts made within the organization. The idea is to measure the inputs, outputs and non-controllable factors and hereby to evaluate the managerial characteristics, like skills and effort. Note that in benchmarking, we usually think in economic production terms, and we refer to different performance dimensions as inputs and outputs. Non-controllable factors are also often thought of as special non-controllable inputs and outputs depending on whether they facilitate or complicate the production process.

Frontier methods

3.10. In the scientific literature, different state-of-the-art estimation techniques have been presented. The best-practice methods go under the name of frontier analysis methods, as they combine the best-practice observations to form a continuous frontier towards which any observation can be gauged. A taxonomy of these methods is illustrated in Table 1 below.

	Deterministic	Stochastic
Parametric	Corrected Ordinary Least Squares (COLS) Aigner and Chu (1968), Lovell (1993), Greene (1990, 2008)	Stochastic Frontier Analysis (SFA) Aigner et al. (1977), Battese and Coelli (1992), Coelli et al. (1998a)
Non-Parametric	Data Envelopment Analysis (DEA) Charnes et al.(1978), Deprins et al. (1984)	Stochastic Data Envelopment Analysis (SDEA) Land et al. (1993), Olesen and Petersen (1995), Fethi et al. (2001)

Table 1 - State of the art frontier methods

3.11. The different estimation methods used for benchmarking are basically suggestions for how to compare individual observations, as illustrated by the dots (ANSPs) in Figure 4 below, given the relationships between input costs and outputs.

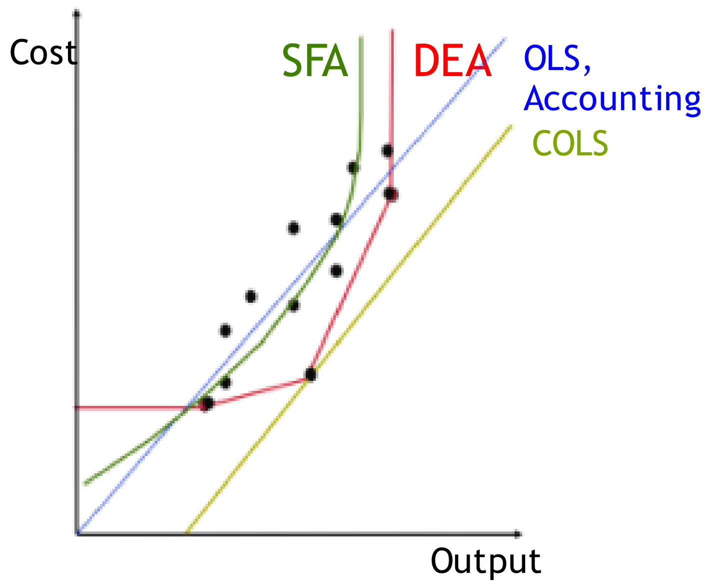

Figure 4 – Multiple estimation methods

3.12. The most frequently applied methods are Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) methods (see Bogetoft and Otto (2011) for a full review). Both approaches have their advantages and disadvantages. In this project, we therefore apply both.

Efficiency measures

3.13. The most frequently applied methods are Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) methods (see Bogetoft and Otto (2011) for a full review). Both approaches have their advantages and disadvantages. In this project, we therefore apply both.

$$Efficiency = \frac{Minimal\ cost}{Actual\ costs}$$

3.14. A cost efficiency measure of, for example, 90% suggests that the ANSP could have produced the same services spending only 90% of its real costs. In other words, there is a savings potential of 10% of the benchmarked cost.

3.15. The relationship to potential savings is illustrated in Figure 5.

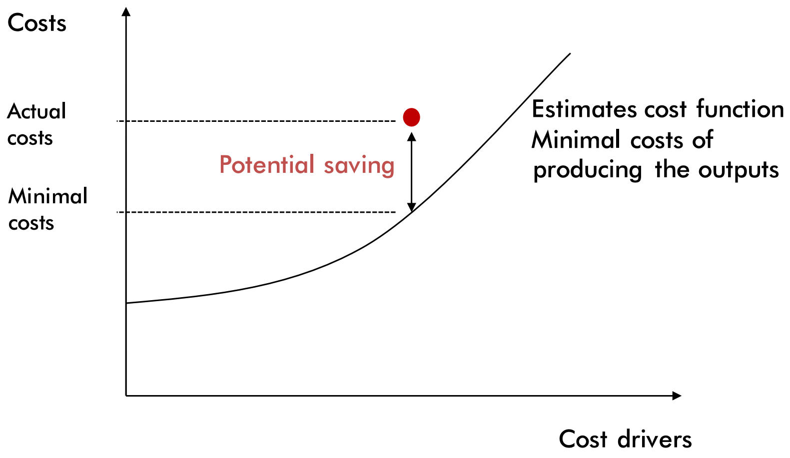

Figure 5 - Efficiency measurement

The benchmarking process

3.16. The development of a regulatory benchmarking model based on international comparisons is a considerable task due to the diversity of the ANSPs involved and the procedural constraints. In this section, we shall highlight some of the typical steps of a regulatory benchmarking analysis and we shall discuss what creates a good benchmarking model. Some of the important steps in a careful benchmarking exercise include the following.

3.17. Choice of variable standardizations: opting for appropriate accounting standards, cost allocation guidelines, inclusion/exclusion criteria, asset definitions, and operating standards is crucial to obtain a consistent data set from ANSPs with varied internal practices.

Choosing a good model

3.18. *Choice of variable aggregation*: Selection of aggregation parameters, such as interest and inflation rates, is necessary for determining standardized capital costs. Additionally, identifying relevant combined cost drivers, possibly through engineering information, helps streamline and reduce the complexity of pertinent data.

3.19. *Initial data cleaning*: Data collection is an iterative process where definitions are likely to be adjusted and refined and where data collected are constantly monitored by comparing simple KPIs across ANSPs and using more advanced econometric outlier detection methods.

3.20. *Average-cost model specification*: To complement expert and engineering model results, econometric model specification methods can be used to investigate which cost drivers / ANSP services best explain average cost. This can be useful to estimate the variability of the data, to validate the fit on the model specification to data and to determine how many cost drivers are necessary.

3.21. *Frontier model estimations*: To determine the relevant best practice model using DEA and SFA models, they must be estimated, evaluated and tested on full-scale data sets. The starting point is the cost drivers derived from the model specification stage, but the role and significance of these cost drivers is further examined in the frontier models, and alternative specifications derived from using alternative substitutes for the cost drivers should be investigated, taking into account the outlier detection mechanisms. In frontier models, special outlier criteria are typically used. The aim is to protect the evaluated ANSPs against a small number of special ANSPs, potentially deploying an incomparable technology or serving an incomparable context, that have an excessive influence on best practices. Two frontier criteria are often used in regulatory benchmarking. One is based on the idea of super-efficiency and says that a single ANSP that is doing very much better than all other ANSPs is most likely an outlier. The other is based on the idea of the average impact on the efficiency of the other ANSPs. An ANSP that has a sizable impact on the efficiency of a large share of the other ANSPs might also be considered an outlier.

3.22. The choice of a benchmarking model in a regulatory context is a multiple criteria problem. There are several objectives, which may conflict with one another.

3.23. *Conceptual*: It is important that the model makes conceptual sense both from a theoretical and a practical point of view. The interpretation must be easy and the properties of the model must be natural. This contributes to the acceptance of the model in the industry and provides a safeguard against spurious models developed through data mining and without much understanding of the industry. More precisely, this has to do with the choice of outputs that are natural cost drivers and with functional forms that, for example, have reasonable returns to scale and curvature properties.

3.24. *Statistical*: It is, of course, also important to discipline the search of a good model with classical statistical tests. We typically seek models that have significant parameters of the right signs and that do not leave large unexplained variation. At the same time, there must be a balance between the complexity of the model used and the sample size. In statistical approaches, this is the question of degrees of freedom. In a DEA context, there are less guidance although some rules of rules-of-thumb has been proposed. One is to require a sample of size of at the very least 3*(number of inputs + number of outputs) and (number of inputs)*(number of outputs). With 30 observations, we should therefore have no more than 9 output parameters. Experience suggests however that this number of output parameters is exaggerated and may lead to models that cannot separate between the efficient and the inefficient firms. Another informal heuristic is to say that DEA models, since they are non-parametric, are extremely flexible and that we therefore need at least enough observations to estimate a translog cost function (Coelli, 2004). With two cost drivers, a translog has 1+2+3 = 6 unknown parameters and with 3 cost drivers it has 1+3+6 = 10 unknown parameters.

3.25. *Regulatory and pragmatic*: The regulatory and pragmatic criteria calls for conceptually sound, generally acceptable models as discussed above. Also, the model will ideally be stable in the sense that it does not generate too much fluctuation in the parameters or efficiency evaluations from one year to the next.

The regulatory perspective also comes into the application of the model. In other words, let us not forget the trivial but very important requirement to comply with the specific conditions laid out in the regulatory directives of the individual jurisdictions.

3.26. The multiple criteria nature of model choice is a challenge. When we have multiple criteria, they may conflict, and this means that there is no optimal model that dominates all other models. We have to make trade-offs between different concerns to find a compromise model, to use the language of multiple criteria decision making, and such trade-offs can be challenged by the regulated parties.

Output based cost functions

3.27. The focus of this project is on the estimation of best practice cost functions and the use of these to estimate potential savings across multiple ANSPs.

3.28. We can distinguish two types of cost functions. Output based costs function explain cost directly as a function of the services provided and the contexts in which they are provided:

$$Cost\ =\ f(Outputs, Context)$$

3.29. Price based cost functions explain costs by the outputs provided, the prices of input factors, and the context:

$$Cost\ =\ f(Outputs, Input\ prices, Context)$$

3.30. Both approaches have their advantages and disadvantages in a practical, regulatory context. The output based approach requires less data since it does not require data on factor prices. Factors prices are often not observed directly but constructed from allocated costs and measures of the physical inputs. An advantage of this approach is therefore that it is also less dependent on the cost allocation of different ANSPs and the use of these costs together with the number of full time equivalents to construct the prices. On the other hand, the output case approach does not allow us to take into account that the relative factor prices may be different across ANSPs and that this may explain some of the cost differences. Note that it is the relative price difference, not the general price levels (which we corrected by inflation and PPP as described below) that matters. If, for example, the cost of capital and the cost of labor are very different across the ANSPs, we would expect them to use different factor combinations, with one relying more on labor and the other more on capital inputs. The consequence of ignoring such differences in price relations might be that some ANSPs are held responsible for aspects of the environment that they cannot entirely control, namely the relative prices of factor input. For these reasons, the output based cost function may potentially lead to harsher evaluations.

3.31. We have chosen to estimate the output based cost function using DEA and the price based cost function using SFA. In this way, we obtain intervals of efficiency scores for each ANSP which may capture some of the methodological uncertainty of any benchmarking study.

Combining DEA and SFA results

3.32. The results from DEA and SFA could be merged in various different ways, with examples of every type of aggregation found in regulatory practices throughout Europe.

3.33. Interval estimates could be created from the efficiency score estimated by each of the two methods (DEA, SFA). It would create a hopefully small band from which the regulator could choose an appropriate level or bound on the individual ANSP price.

3.34. The minimum efficiency score, min(DEA,SFA), would be the toughest estimate of potential cost reduction identified by at least one of the models results.

3.35. The maximum efficiency score between the results of the two models, max(DEA,SFA), could be referred to as the 'benefit of the doubt' regulatory approach. This would lead to the lowest possible cost reductions.

3.36. Calculating the average score of the results of the two models, median(DEA,SFA), would balance the advantages and disadvantages of each model equally. This would lead to results similar to that of the interval estimates and is the approach that we have chosen as described in Section 6.

4. Data and descriptive statistics

4.1. The data regarding the performance of the ANSPs have been provided and validated by the Performance Review Body (PRB). The AG received the data covering the period 2012-2021. The data includes 28 Member States plus MUAC. The costs considered here include the actual costs reported in the charging zones except for National Supervisory Authorities (NSA) and Eurocontrol costs. Therefore, the cost efficiency models only focus on the ANSPs cost base.

4.2. The countries included are (the ANSP is indicated in parenthesis): Austria (Austro Control), Belgium-Luxembourg (Belgocontrol), Bulgaria (BULATSA), Croatia (Croatia Control), Cyprus (DCAC Cyprus), Czech Republic (ANS CR), Denmark (NAVIAIR), Estonia (EANS), Finland (Finavia), France (DSNA), Germany (DFS), Greece (HCAA), Hungary (HungaroControl), International (MUAC),Ireland (IAA), Italy (ENAV), Latvia (LGS), Lithuania (Oro Navigacija), Malta (MATS), Netherlands (LVNL), Norway (Avinor Continental), Poland (PANSA), Portugal (NAV Portugal Continental), Romania (ROMATSA), Slovakia (LPS), Slovenia (Slovenia Control), Spain (ENAIRE), Sweden (LFV), Switzerland (Skyguide).

4.3. The years 2020 and 2021 have been excluded from the empirical analyses due to the severe impact of the COVID-19 pandemic on the air transportation sector. For example, in March-April 2020, 100% of available seats were grounded in Europe (Andreana et al., 2021), showcasing the dramatic effect of COVID-19 on the industry. Additionally, current global capacity is still lower than that of 2019.[1] Data from the Official Airline Guide (OAG) indicate that worldwide capacity reached 111.4 million seats in May 2023, which is still 3.6% lower than the same month in 20192. Relating the ANSP costs to their operations and traffic volumes in the years 2020 and 2021 would likely lead to skewed results. Therefore, the benchmarking analysis uses a dataset that includes only the period from 2012 to 2019 to evaluate ANSP performance.

4.4. Most of the data are in monetary values and have been converted to euros. The costs of an ANSP are primarily composed of two main elements: operating costs (OPEX) and capital costs (CAPEX). OPEX encompasses the variable staff costs and other operating expenses. CAPEX, on the other hand, comprises the cost of capital and depreciation. The cost of capital represents the opportunity cost associated with investing money in air traffic control, which is economically linked to the return on the capital invested. Depreciation accounts for the necessary funds to maintain the quality of assets at a consistent level. Capital is estimated by the ANSP net book value.

4.5. Much of the data draws from accounting records and therefore requires standardization to be comparable across ANSPs. First, data in currencies other than the euro must be converted to euros (the values have been calculated by the PRB supporting team using the average exchange rate for the year 2017

[1] See reports on the website oag.com.

as a reference). Second, since purchasing power varies among the countries included in the analysis, all monetary values must be adjusted using a Purchasing Power Parity (PPP) index to account for these differences.

4.6. Purchasing Power Parity (PPP) denotes the number of currency units required to buy a specific quantity of goods and services in different countries. PPPs can be used as currency conversion rates to convert expenditures expressed in national currencies into an artificial common currency, thereby eliminating the effect of price level differences across countries. The index has been estimated by Eurostat, using the average purchasing power of the EU across 29 member states in 2020 as a reference.

4.7. All monetary values are transformed into constant terms using a Producer Price Index (PPI) for each country (source: Eurostat). In general, producer price indices measure the average change in prices paid by domestic producers for goods and services sold in domestic and/or export markets between different time periods. The Producer Price Index is used to represent the cost of purchasing materials and supplies from local producers. The PPI is set at 100 for the base year, which is 2012.

4.8. In summary, if a variable X is defined in monetary terms and in a currency different from the euro, it is first converted to euro, and then standardized according to the following formula: $\frac{X/PPP}{PPI} \times 100$.

Physical production factors

4.9. We include measures of the physical amounts of multiple production factors utilized by the ANSPs. We consider the working hours of air traffic controllers (ATCO) for this purpose (i.e., area control center air traffic control (ACC ATCO)-hours on duty, taken from the Eurocontrol ATM Cost-Effectiveness (ACE) benchmarking reports).

4.10. Output is measured by the total instrument flight rules (IFR) flight hours controlled by each ANSP on an annual basis, information provided in the Eurocontrol ACE benchmarking reports.

Pricing factors

4.11. The labor cost is determined by dividing staff expenses by ATCO hours, offering an approximation of the hourly labor rate.

4.12. The price of capital is derived from the ratio of CAPEX to the annual sector opening hours of the corresponding ANSP. Economic theory posits that the price of capital represents the cost a firm incurs for utilizing capital, closely aligning with the rent paid for asset usage, including buildings. Therefore, the economic price of capital intrinsically links to the operational hours of an asset, a perspective adopted in this report.

4.13. We also tested alternative estimations using different definitions for the price of capital, based on an evaluation of each ANSP's total assets provided by the PRB support. For example, CAPEX divided by the regulated asset base (RAB), or alternatively, CAPEX divided by the net book value of the fixed assets (NBV). However, as depicted in

4.14. Figure 6, the RAB trend throughout the observation period showcases pronounced fluctuations, particularly in France, Germany, Greece, Italy, Spain, and Switzerland. These variations were particularly evident towards the end of the timeline, coinciding with the onset of Regulatory Period 2 (RP2, 2015-19).

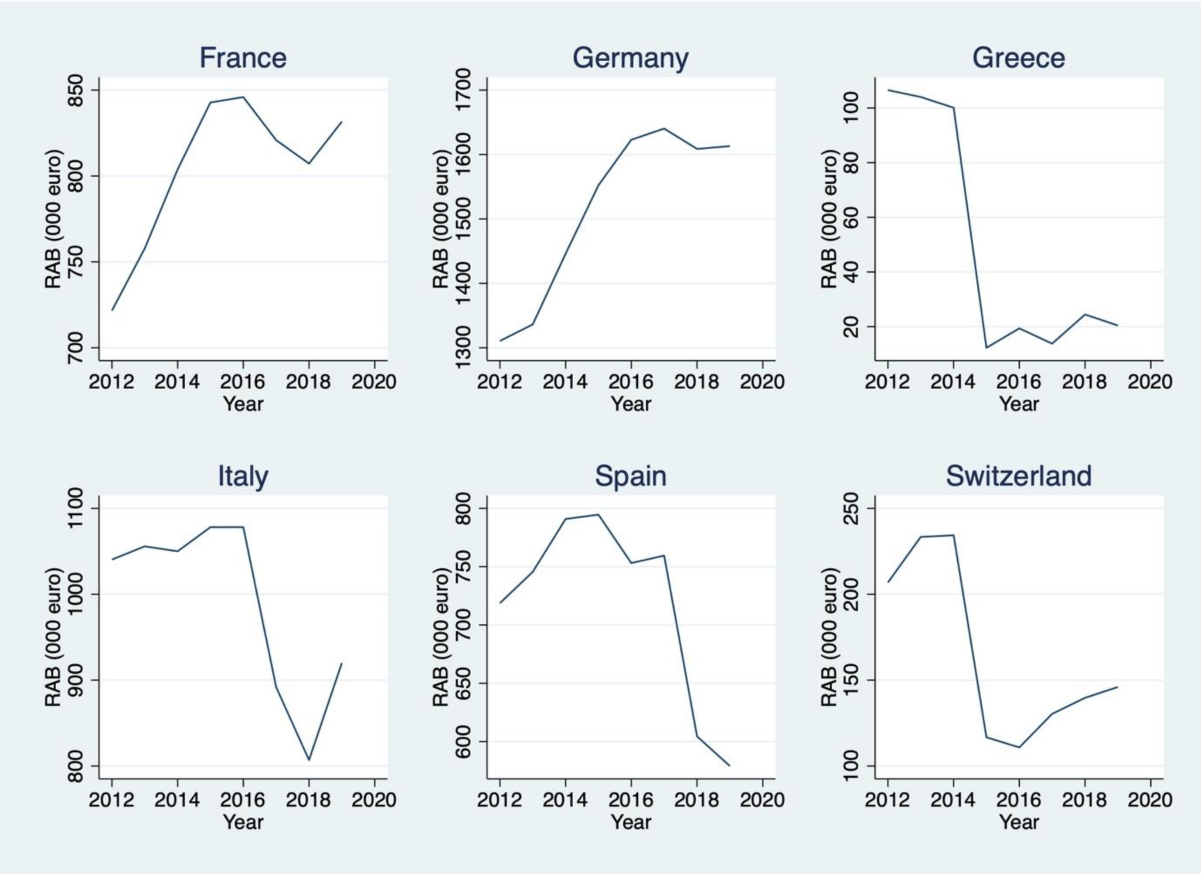

Figure 6 – Variability in RAB of six ANSPs, 2012-19

4.15. Similarly, as demonstrated in

4.16. Figure 7, we analyzed the data using the NBV. In this scenario, even more ANSPs show dramatic fluctuations. The effect of these trends becomes significant when we consider CAPEX, which experiences a marked decrease towards the end of the observed period. These features render the price of capital, if calculated as the ratio between CAPEX and either RAB or NBV, statistically insignificant when we apply stochastic frontier analysis to estimate cost efficiency. Consequently, we chose to employ the definition of the price of capital given by the ratio of CAPEX over the sector opening hours, which performs satisfactorily in the estimation procedure.

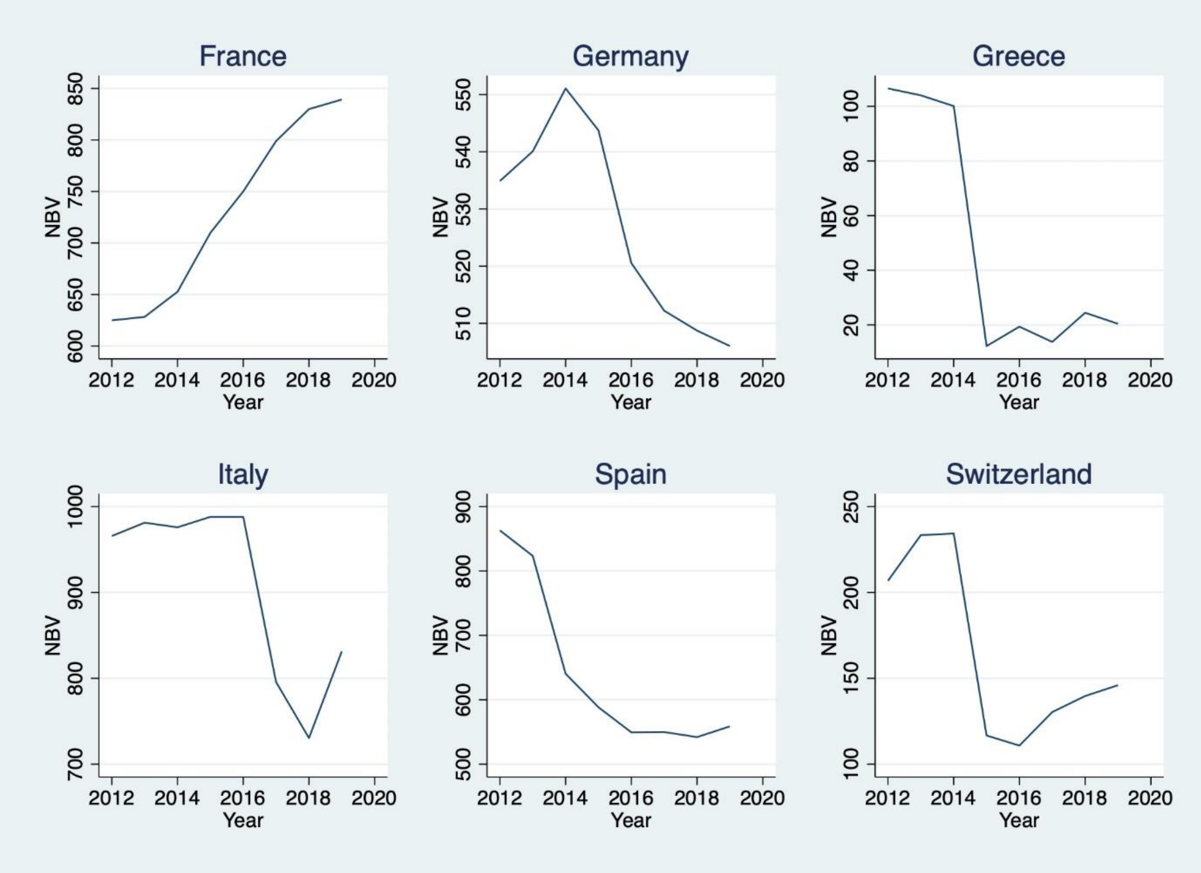

Figure 7 – Variability in NBV of six ANSPs, 2012-19

Negative Externalities

4.17. ANSP operations may inadvertently result in negative externalities, often referred to as 'bad outputs', which lead to undesired outcomes that detract from the overall experience of travelers due to unforeseen extended travel times. Therefore, the AG also included delays in the benchmarking analysis. In an ideal scenario, increased delays should correlate with lower total costs because the ANSP does not utilize all necessary inputs to ensure punctuality. However, the actual impact of delays on total costs is a subject left to the empirical analysis.

4.18. Since delays are considered a negative output, they are treated in such a way that an ANSP is penalized for higher levels of delay. The reasoning behind this is that an ANSP's performance is considered particularly good in terms of delays when this negative output is minimized. Therefore, the time lost due to delays is inverted in the cost efficiency estimation. This implies that the greater the delays attributed to an ANSP, the lower the output. Furthermore, since we have several instances with zero delays, this negative output variable is computed as follows: $\frac{1}{1+delays}$. This approach ensures that the ratio is always defined and that the ANSP's performance in this dimension ranges between 0 and 1.

4.19. To accurately assess the workload of each ANSP, we must take into account the complexity of the flight paths managed by them. To assist in this, Eurocontrol generates an index reflecting the complexity of each ANSP's flight paths on an annual basis.

4.20. The fluctuation of traffic load over a year can also influence the relative cost base of an ANSP. This variability is calculated by dividing the traffic levels in the peak month by the average monthly traffic. Since it is not feasible to employ ATCOs seasonally, high variability could result in increased annual costs compared to an ANSP with a similar output distributed evenly throughout the year. Variability is computed as an index by Eurocontrol.

4.21. Complexity and variability are characteristics of the air traffic controlled by the ANSPs and can be incorporated in the benchmarking analysis in different ways. Most commonly, they are included as explanatory variables in the inefficiency model or they are used to construct additional volume-based output measures that can be considered as outputs. In the first case, they are added to the estimated regression as explanatory variables. In the second case, we have constructed the following additional variables for the data envelopment analysis model:

$$Complexityhours = Complexity \times Total\_IFR\_hours$$

$$Variabilityhours = Variability \times Total\_IFR\_hours$$

Data transformation

4.22. Table 2 presents all the variables included in the cost efficiency estimation, their unit of measure and description. The data draws on different sources: Reporting tables (costs, capital), EUROCONTROL (traffic, staff) and Eurostat (inflation, purchasing power parity).

	Variable	Unit of Measure	Description/Computation
Output	Flight Hours	hours	Total IFR flight-hours controlled by ACC (aggregated at ANSP level)
Costs	Staff costs Other operating costs Depreciation Cost of capital	000€	Normalized by average euro exchange rate 2017
Staff	ACC ATCO-hours on duty	hours	Staff
Explanatory Variables	Sector opening hours	hours	Sum of sector hours
	ATFM delay	minutes	Minutes of en route ATFM delay
	Variability score	index	Traffic levels in the peak month divided by average monthly traffic
	Complexity score	index	Potential number of interactions between aircraft per flight-hour controlled, considering traffic density and structural index
Indices	PPI	%	Producer Price Index (annual growth rate)
	PPP		Purchasing Power Parity (EU27_2020=1)

Table 2 - Variables definition, unit of measure and description

Descriptive statistics

4.23. Exploratory data analysis are presented in Figure 8 and Figure 9, having been normalized at the base year 2012. Figure 8 depicts a consistent increase in flight hours-controlled by upwards of 20% compared to 2012, and similar patterns are observed for IFR flights per ANSP.

4.24. Sector opening hours remain relatively constant from 2012 to 2015, increase between 2015 and 2017 and then decrease slightly, leading to an overall 7% percentage increase between 2012 and 2019.

4.25. Complexity in managing flight traffic control gradually increases over the timeframe by about 15%, whilst variability remains relatively constant over the same timeframe.

4.26. Delays increase over three-fold by 2018 and decrease slightly in 2019. In order to visualize the bad output, it is necessary to use a two-scale plot. The scale of the delays index is shown on the right of the graph in Figure 8. At the end of the period the increase is approximately 300%.

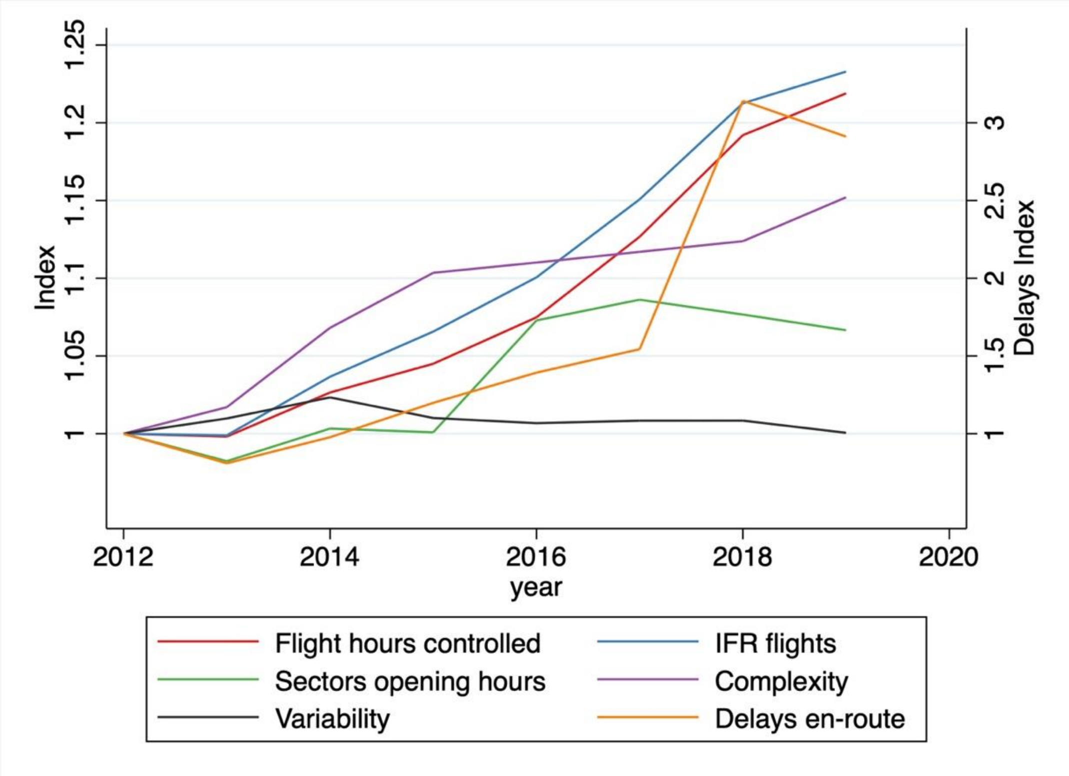

Figure 8 - Output indices, 2012-19

4.27. Figure 9 illustrates a +9% increase in total costs. Labor is quantified using the full-time equivalent (FTE) metric. Staff costs have increased by about 12%, despite a decline of about 2% in the total number of employees, which indicates a significant rise in wages.

4.28. The operating costs (OPEX) index has increased by 11%. CAPEX at the end of the period (2019) is at the same level as the beginning year (2012), due to a 7% decrease between 2018 and 2019.

4.29. Hence, the observed trend at the descriptive level is that output increased more than costs, and that the lower increase in costs is due to constant capital costs whereas labor costs have risen.

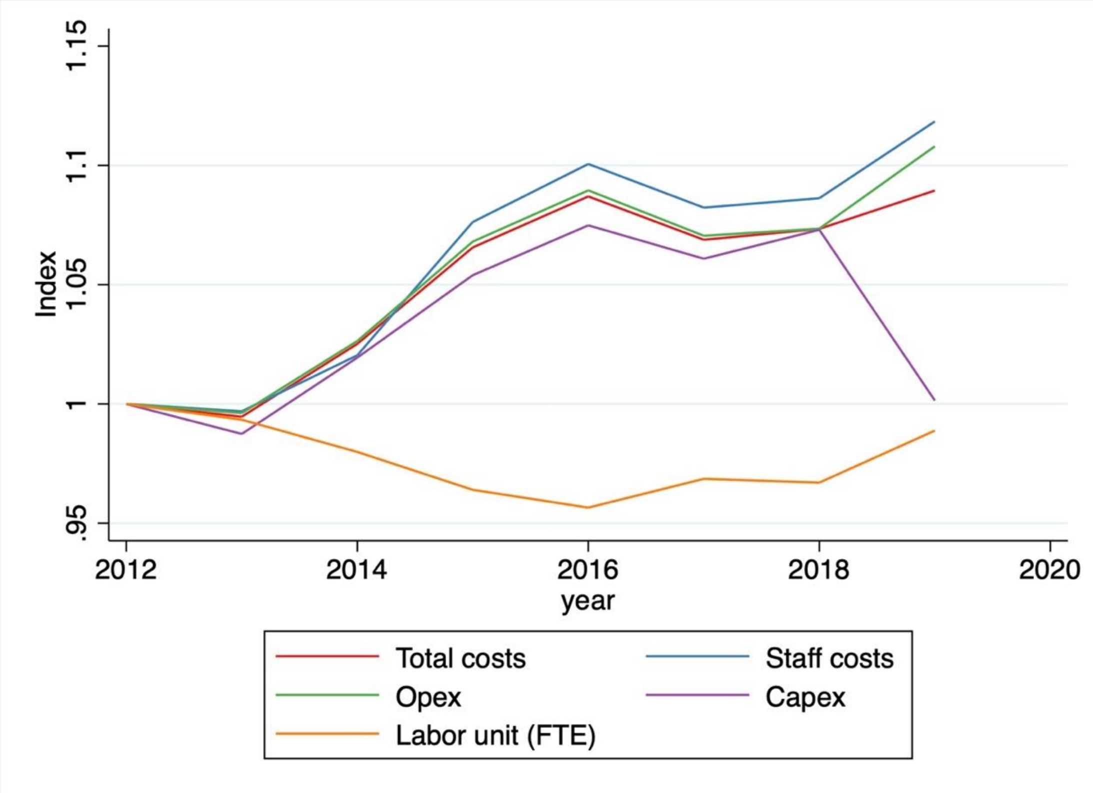

Figure 9 - Cost indices, 2012-19

4.30. Table 3 presents a summary of the data over the timeframe analyzed. Depreciation costs and the economic cost of capital are relatively invariant over time. In year 2012, about 21,700 full-time-equivalent employees worked for the ANSPs, which decreased until 2015 and then increased in the last years, leading to 21,500 by year 2019.

Years	2012	2013	2014	2015	2016	2017	2018	2019
All currency in 000€ PPPd				**Costs**				
Staff costs	3,527,512	3,486,527	3,509,171	3,598,576	3,607,221	3,673,799	3,774,476	3,932,105
Other operating costs	886,229	880,305	900,210	875,124	867,922	885,350	912,554	961,861
Depreciation	634,453	626,779	637,873	617,410	622,640	636,870	661,689	652,581
Cost of capital	285,645	268,529	287,882	308,576	304,477	304,240	323,724	277,848
Opex	4,413,741	4,366,832	4,409,381	4,473,700	4,475,143	4,559,149	4,687,030	4,893,966
Capex	920,098	895,309	925,755	925,985	927,117	941,110	985,413	930,429
Total costs	5,333,839	5,262,140	5,335,136	5,399,685	5,402,260	5,500,259	5,672,442	5,824,394
				Inputs				
ATCO hours on duty	9,529,471	9,412,612	9,487,217	9,331,701	9,454,511	9,530,020	9,514,383	9,728,853
Labor unit (FTE)	21,719	21,575	21,282	20,937	20,775	21,037	21,003	21,476

Table 3 - Annual trends in ANSPs costs and labor inputs

5. Models applied

Data Envelopment Analysis (DEA)

5.1. The non-parametric DEA approach uses linear programming to evaluate the performance of the firms or organizations. In the DEA literature is common to refer to the evaluated as Decision Making Units (DMUs). A DMU can be an observation of inputs and outputs for a firm at a given time (cross section) or across time periods (panel data).

5.2. DEA does not use maximum likelihood estimation, which is common in more statistical approaches, to determine the underlying model. Instead, DEA is based on the idea of minimal extrapolation.

5.3. In DEA, the estimate of the technology T, which is the empirical reference technology, is constructed as the smallest set of input-output combinations that contains data from the different DMUs, (xk,yk), k = 1,...,K and satisfies certain technological assumptions specific to the given approach.

5.4. By constructing the smallest set that contains the actual observations, the method extrapolates the least. As long as the true technology T satisfies the regularity properties, the approximation T* that we develop will be a subset of the true technology. We refer to this as an inner approximation of the technology. By choosing the smallest set, we are making a cautious or conservative estimate of the technology set and therefore, also a cautious or conservative estimate of the loss due to inefficiency. We can say also that the approximation is based on best practices rather than on speculation as to what may be technologically feasible. A popular understanding of the property is also that we estimate the technology so as to present the evaluated units in the best possible light.

5.5. We note that DEA is based on the implicit assumption that there is no noise in the data. If the data are somewhat random, due to exogenous shocks, bad reporting practices or ambiguity in accounting practices, the result may not be an inner approximation of the true possibilities.

Assumptions of DEA models

5.6. The basic DEA models mainly differ in the assumptions that they make about the technology T. The most important assumptions include free disposability (we can produce less with more), convexity (a weighted average of feasible production plans is feasible), scaling (production may be scaled) and additivity (the sum of two feasible production plans is feasible).

5.7. Given the size of the data set, and our aim to discriminate among efficient and inefficient firms, it is useful to assume convexity. Convexity is an assumption that complies with standard cost and production theory and that is also invoked in most parametric approaches.

5.8. With respect to returns-to-scale, we choose between the following:

a) Constant Returns to Scale (CRS) means that we do not believe there to be significant disadvantage of being small or large.

b) Non-Increasing Returns to Scale (NIRS), sometimes referred to as Decreasing Returns to Scale (DRS), means that there may be disadvantages of being large but no disadvantages from being small.

c) Non-Decreasing Returns to Scale (NDRS), sometimes referred to Increasing Returns to Scale (IRS), means that there may be disadvantages of being small but no disadvantages of being large.

d) Variable Return to Scale (VRS) means that there are likely disadvantages of being too small and too large.

5.9. Both conceptual reasoning and statistical tests aid in determining the appropriate scale assumption. The CRS assumption is the most stringent and results in the lowest efficiency scores. To align with the SFA translog model, we have chosen the VRS convex DEA model.

5.10. Finally, we analyze the ANSPs on an annual basis in order to minimize the impact of noise in the data.

Outliers

5.11. Outlier analysis consists of screening extreme observations. Depending on the approach chosen (DEA or SFA), outliers may have a different impact. In DEA, particular emphasis is put on the quality of observations that define best practice. In SFA, outliers may distort the estimation of the curvature and affect the magnitude of the idiosyncratic error term.

5.12. There are several possible outlier detection techniques that are relevant for DEA models, c.f. Bogetoft and Otto (2011) and Wilson (1993). One approach is to identify the number of times a DMU serves as a peer unit for other DMUs, peer counting. If a DMU is the peer for an extreme number of units, it is either a very efficient unit or there may be mistakes in the reported numbers. An alternative approach is the super efficiency criterion (Andersen and Petersen, 1993; Banker and Chang, 2006). The idea is to eliminate ANSPs that are far outside the technology spanned by the other ANSPs.

5.13. Applying multiple approaches, we identified MUAC as an outlier and have removed the ANSP from the analyses across all years, simply assuming that it is consistently relatively efficient.

DEA Variables

5.14. In the en-route model, we define five cost drivers as shown in Table 4. The total IFR flight hours controlled is a direct measure of workload (Flight Hours), the total hours that the sectors are open is the measure of the size of the operation and the actual and potential workload (Sector Opening Hours), the complexity index multiplied by IFR flights controlled is a workload measure that is corrected by complexity, the variability index multiplied by IFR flights controlled is a measure of the capacity for handling a large workload at least temporarily, and delays include the total minutes of delay specifically attributed to the ANSP.

Model	Variables
Inputs	
Total Costs	Total expenses PPP corrected
Outputs	
Flight hours	Total IFR flight hours controlled en-route
Sector opening hours	Total hours that sum of sectors open
Complexity*Flight hours	Complexity Index * flight hours controlled
Variability*Flight hours	Variability Index * flight hours controlled
Delays	Total minutes of delay annually ascribed to ANSP
Estimation Approach	
	Variable returns-to-scale
	Outlier MUAC eliminated

Table 4 - Variables in DEA model

Stochastic Frontier Analysis

5.15. The econometric approach to efficiency estimation is concerned with measuring the performance of firms and institutions in converting inputs to outputs. SFA may be applied to either cross-sectional or panel data at the firm level in order to estimate the relationship between inputs and outputs whilst accounting for exogenous factors. The latter may impact the production relationship however the management of the firm in general may have little to no control.

5.16. A firm is deemed cost efficient if it minimizes the total production cost of a given output, which requires technical efficiency but also a mix of inputs that makes more intensive use of the relatively cheaper variables. After testing both Cobb-Douglas and the more flexible translog cost function approaches, we chose the latter due to the higher log likelihood function values.

5.17. Due to the existence of panel data and potential externalities, we apply the Battese and Coelli (1995) model, which accounts for potential heteroscedasticity in the decomposed error terms and the estimation of the impact of externalities on the inefficiency distribution. Consequently, the Battese and Coelli model considers environmental variables twice if necessary, namely within the cost function and as an explanation for the average level of inefficiencies (Hattori, 2002).

5.18. From the dataset, we apply the model to the set of variables described in Table 5, where the cost of operation index equals the producer price index (PPI). Total costs and prices are normalised by one of the prices in order to meet the homogeneity condition and we have chosen the purchase price parity (PPP) index accordingly.

Dependent Variable	
Total Cost	$\frac{\text{total cost } / \text{ PPP}}{\text{producer price index}}$
Independent Inputs	
Output	total IFR flight hours controlled en-route
Labour price	$\frac{(\text{total staff cost}/\text{ATCO hours })/\text{PPP}}{\text{producer price index}}$
Capital price	$\frac{(\text{depreciation cost } + \text{ cost of capital }) \ / \ (\text{sector openings}/ \text{ PPP})}{\text{producer price index}}$
Environmental Variables	
Airspace characteristics	variability (seasonality), complexity, sector opening hours, time trend, delays

Table 5 - Variables in Stochastic Frontier Cost Function

5.19. Given the translog nature of the analysis, which ensures a reasonably flexible cost function, all of the independent inputs are also multiplied by themselves and between each other.

5.20. We implement the estimations in STATA, using the tailor-made SFPANEL package (Belotti et al., 2012). We tested a number of alternative specifications including SFA with time decay in the inefficiency term (Battese and Coelli, 1992) and SFA with exogenous drivers affecting the distribution of the inefficiency term (Battese and Coelli, 1995) and chose the latter based on the log likelihood values. We also note that all variables were subsequently standardized by dividing them by their geometric mean prior to logging the data.

5.21. The SFA model applied to en-route air traffic control provision is presented below:

$$
\begin{aligned}
\ln\left(\frac{Total\ Cost_{it}}{PPI_{it}}\right) \\
&= \beta_0 + \beta_1 \ln(IFR\ flight\ hours_{it}) + \beta_2 \ln\left(\frac{Labor\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_3 \ln\left(\frac{Capital\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_4 \frac{1}{2} \ln(IFR\ flight\ hours_{it}) \ln(IFR\ flight\ hours_{it}) \\
&+ \beta_5 \frac{1}{2} \ln\left(\frac{Labor\ Price_{it}}{PPI_{it}}\right) \ln\left(\frac{Labor\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_6 \frac{1}{2} \ln\left(\frac{Capital\ Price_{it}}{PPI_{it}}\right) \ln\left(\frac{Capital\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_7 \ln(IFR\ flight\ hours_{it}) \ln\left(\frac{Labor\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_8 \ln(IFR\ flight\ hours_{it}) \ln\left(\frac{Capital\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_9 \ln\left(\frac{Labor\ Price_{it}}{PPI_{it}}\right) \ln\left(\frac{Capital\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_{z1} \ln(Complexity_{it}) + \beta_{z2} \ln(Variability_{it}) \\
&+ \beta_{z3} \ln(Sectors_{it}) + \beta_{z3} \ln(time_t)\, \beta_{z3} + v_{it} + u_{it}
\end{aligned}
$$

where $v_{it} \sim N(0, \sigma_v^2)$ and $u_{\text{it}} \sim N(\delta_1 ln(complexity)_{it} + \tau_{it}, \sigma_u^2)$

5.22. The results of the stochastic cost function models with respect to en-route services are presented in Section 6. Two models are analysed, namely without and with a time trend variable that estimates market level changes. All cost elements are PPP to allow for international comparisons. All variables are logarithm transformed and normalized by the geometric mean.

6. Results

6.1. In this section of the report, we present the estimates of the Union-wide ANSPs cost efficiency. Two models were applied in order to estimate the efficiency of the 29 ANSPs, namely radial, variable returns-to-scale DEA and translog SFA models.

DEA cost efficiency

6.2. The DEA model includes the variables specified in Table 4. The DEA cost frontier model includes a single input, total costs and four outputs: flight hours controlled, sector opening hours, complexityhours, and variabilityhours. We also estimate a second model that includes delays.

6.3. The AG performed a systematic outlier analysis prior to applying the DEA models, following the method in Bogetoft and Otto (2011). The evidence is that in most of the annual analyses, MUAC has been identified as outlier, therefore it has been classified accordingly. MUAC has been assigned an efficiency score equal to 1, and the remaining ANSPs have been investigated and assigned an efficiency score based on a DEA analysis limited to 28 annual observations.

6.4. The results of the DEA-VRS model without and with delays are presented in Table 6. The scores are cost efficiency measures, ranging from 0 to 1. For example, the estimated median score is equal to 0.85 in year 2019, which means that the estimated inefficiency score is 15%. The scores presented in this table are Union-wide annual median scores. Without considering delays, the efficiency increased in the system, moving from 61% in year 2012 to 85% in year 2019. During RP1, the efficiency levels remained relatively constant over the observed period. After a drop between the two regulatory periods, the relative improvement consistently increases over the five years of RP2.

Year	Without delays	With delays
2012	0.61	0.73
2013	0.59	0.77
2014	0.62	0.84
2015	0.59	0.59
2016	0.66	0.71
2017	0.71	0.81
2018	0.79	0.85
2019	0.85	0.90

Table 6 - DEA cost efficiency estimates without and with delays

6.5. The scores without delays in Table 6, are lower than those with delays by definition. This is simply due to the additional output dimension, which enables ANSPs with low delay levels to improve their relative performance. If delays are included, the ANSPs Union-wide DEA-VRS efficiency scores increased during the time interval 2012-2019, rising from 73% in year 2012 to 90% in year 2019. As in the previous case, there is a large drop in year 2015, at the beginning of RP2. The efficiency rose during RP1 (2012-14), then dropped by 25% in 2015. Subsequently, performance improves consistently until the end of RP2 (year 2019).

6.6. Figure 10 presents the distribution of the estimated efficiency scores for the 29 ANSPs per year of observation using the DEA-VRS model with delays. The graph in each year is a box plot, and the bottom line of the rectangular box is the efficiency score located at the 25th percentile of the distribution: for example, in year 2012 the efficiency of the first quartile of the distribution of efficiency scores is equal to 40%. The line in the middle of the box is the median and is the efficiency score exactly in the middle of the distribution. In year 2012, it lies at 73%. The upper line of the rectangular box is the efficiency score of the upper quartile of the distribution. In year 2012, it is 100% (and it is the same in all following years).

6.7. We note that the interquartile range, i.e., the vertical distance between the bottom and the upper line of the rectangular box has reduced across the timeframe. The inter-quartile range is about 60% in year 2012, and just above 50% in year 2019. Hence, the Union-wide ANSP system has reduced the dispersion in the efficiency scores by the end of RP2.

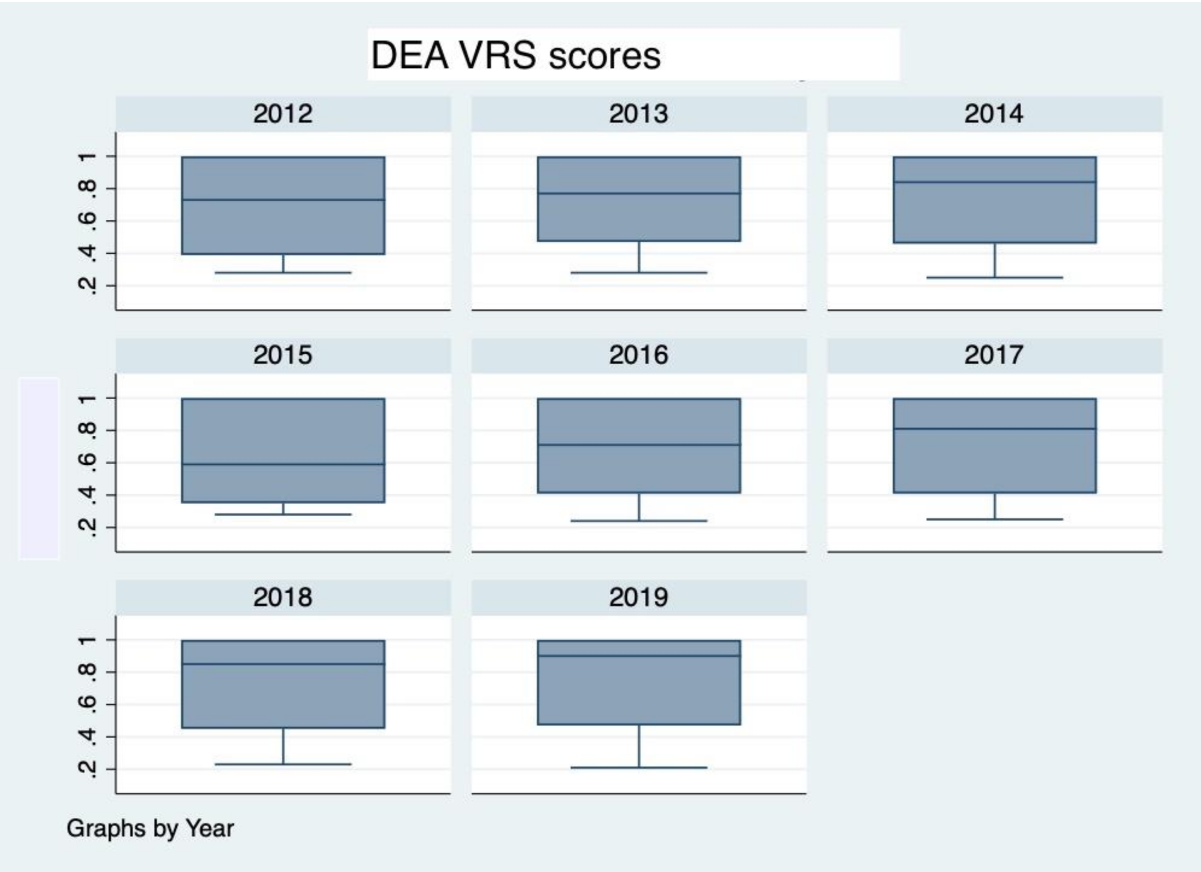

Figure 10 - Box plot distribution of DEA efficiency scores with delays

SFA cost efficiency

6.8. Cost efficiency with SFA is estimated with the translog, Battese and Coelli (1995) model. Total costs are explained by flight hours controlled, price of labor, price of capital, sector opening hours, complexity and variability together with a time trend in Model (1). Models (2) and (3) include delays. Model (3) treats complexity as a determinant of inefficiency rather than an explanatory variable. The estimates are presented in Table 7.

6.9. Models (1) to (3) show that flight hours controlled, and the prices of capital and labor, all explain total costs. Regarding input prices, the most significant values are labor, followed by capital in explaining overall costs.

6.10. We note that Model (3) appears to be preferable from a statistical perspective because the log-likelihood value is higher. Furthermore, it has a statistically significant (at the 5% level) estimated coefficient for the standard deviation of the inefficiency error component, σ_u, as well as the coefficient related to the standard deviation of the shock error component, σ_v. Hence, inefficiency is an important component of the cost function error term, as required for the adoption of SFA. We therefore refer to Model (3) for the rest of this section.

Dependent variable: *Total Costs*						
	Model 1		Model 2		Model 3	
Output and input prices	Coefficient	S.E.	Coefficient	S.E.	Coefficient	S.E.
Flight hours	0.40***	(0.11)	0.35***	(0.05)	0.33***	(0.11)
Capital price	0.35***	(0.06)	0.32***	(0.03)	0.29***	(0.04)
Staff price	0.53***	(0.06)	0.54***	(0.04)	0.56***	(0.06)
(Flight hours) ^{2}	-0.002	(0.07)	0.02	(0.04)	-0.003	(0.05)
(Capital price) ^{2}	0.02	(0.12)	0.02	(0.03)	0.01	(0.03)
(Staff price) ^{2}	0.47	(0.64)	0.66***	(0.18)	0.60***	-0.17
(Flight hours)(Capital price)*	-0.02	(0.05)	-0.003	(0.04)	-0.05	(0.03)
(Flight hours)(Staff price)*	-0.20***	(0.01)	0.21***	(0.04)	-0.13	(0.08)
(Capital price)(Staff price)*	-0.18	(0.26)	-0.23***	(0.08)	-0.23***	(0.08)
Shifters of total costs						
Sector opening hours	0.59***	(0.09)	0.60***	(0.06)	0.65***	(0.09)
Delays			0.002	(0.004)	0.001	(0.01)
Complexity			0.07**	(0.03)		
Variability (seasonality)					0.87***	(0.35)
Time trend	-0.006	(0.01)	-0.01**	(0.004)	-0.01**	(0.004)
Constant	12.3	(22.41)	16.67**	(7.17)	15.21**	(7.76)
Inefficiency (m)						
Complexity					0.09	(0.95)
Constant	-0.02	(0.36)	-4.41	(1.02)	-1.25	(1.67)
$\boldsymbol{\sigma_u}$ *(Inefficiency error component)*	0.26***	(0.95)	0.32	(0.21)	0.45**	(0.25)
$\boldsymbol{\sigma_v}$ *(shock error component)*	0.001	(2.97)	0.05***	(0.02)	0.05*	(0.03)
$\boldsymbol{\lambda}$	262.79***	(0.13)	6.97***	(0.21)	9.82***	(0.26)
Log-likelihood	149.61		151.18		168.3100	
Observations	232				232	
Legend: Battese and Coelli (1995) SFA models. *** = 1% significance level;						
** = 5% significance level; * = 10% significance level						

Table 7 - ANSPs cost function estimated with SFA translog model

SFA cost elasticities

6.11. Since the variables are in logs and standardized by their geometric means, the estimated coefficients of the first degree variables, namely flight hours, capital price and labor price, represent elasticities. Hence, the cost elasticity of flight hours is equal to about 0.33%, which suggests that were the volume of hours of traffic controlled by an ANSP to increase by 1%, the total costs will rise by 0.33%.

6.12. The cost elasticity of the price of capital is 0.29%, while the cost elasticity of the price of labor is higher and equal to 0.56%. An increase by 1% in the price of capital gives rise to an increase in total ANSP costs equal to 0.29%, whereas a 1% increase in the cost of labor will increase total costs by 0.56%.

6.13. Longer sector opening hours also contribute significantly to the overall costs. The estimated coefficient is equal to 0.65 and it is statistically significant.

6.14. Delays do not significantly impact cost efficiency for the ANSPs. The two estimated coefficients in Models (2) and (3) are not statistically significant. However, the estimated coefficient is positive. This implies that the lower the en-route delays, the higher is the cost to the ANSPs. In order to minimize delays, ANSPs may need to incur higher costs.

6.15. Higher complexity in ASNP operations implies higher total costs. The estimated coefficient of complexity in Model (2) is positive and significant, equal to 0.07. However, Model (2) suffers from a not statistically significant, coefficient for the standard deviation of the inefficiency error component, σ_u. Hence, after showing that complexity is a positive shifter in the cost frontier, in Model (3) we move complexity to be a determinant of the inefficiency error component and add variability as a cost function explanatory variable. The estimated coefficient of σ_u becomes positive and statistically significant, as required by SFA.

6.16. Variability (i.e. seasonality) contributes to higher costs and is statistically significant at the 1% level. In Model (3), the estimated coefficient is equal to 0.87.

6.17. The negative time trend, whilst not significant in Model (1), is significant at the 5% level in Models (2) and (3). This suggests that costs have decreased over the two reference periods by on average 1% annually.

SFA efficiency distributions

6.18. Figure 11 presents the distribution of cost efficiency scores over the period 2012 to 2019 using the SFA Model (3) estimates. Across all years, the inter-quartile range is smaller than that of the results of the DEA model. Consequently, SFA generates less dispersion in the efficiency scores, although the 75th percentile is always lower than 100%.

6.19. The SFA estimates yield Union-wide cost efficiency estimates of 83% if we do not consider delays (Model 1) and 88% if we include delays (Model 3).

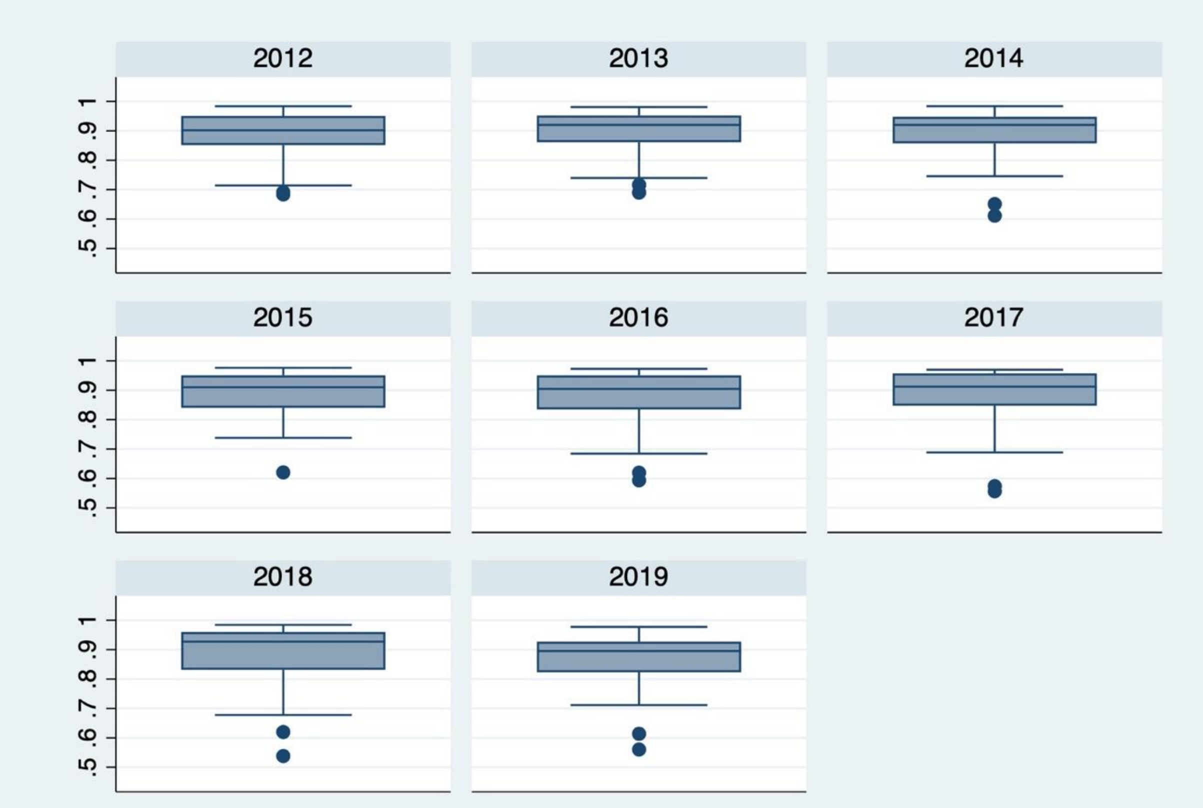

Figure 11 - Box plot distribution of SFA efficiency scores including delays

6.20. The SFA findings indicate that, on average over the observation period, the Union-wide cost efficiency is 83% without accounting for delays (Model 1), which rises to 88% when delays are incorporated (Model 3).

Combining the results

6.21. The suggested savings are based on the average cost efficiencies for the entire period, 2012 to 2019, and are computed as follows:

Potential cost saving = 1 – average efficiency score.

6.22. We combine the potential savings obtained by the DEA and the SFA models following three approaches:

- Potential savings as the maximum value resulting from the DEA and SFA models
- Potential savings as the minimum value resulting from the DEA and SFA estimates

Potential savings as the average of the two sets of results

6.23. Table 8 reports the DEA and the SFA Union-wide average estimated cost efficiency scores. We report two average measures: the simple average and the weighted average. In this case, the weight is set by the share of each ANSP's total costs on the Union-wide total costs, thus taking account of relative size. The results of the two models are reported across the two reference periods (RP) included in the observations: RP1 covering 2012 to 2014, i.e., 3 years, and RP2 covering 2015 to 2019, i.e., 5 years.

6.24. Table 9 presents the average potential savings for the overall period, i.e., 2012-19. The weighted average Union-wide ANSP inefficiency score offers a more accurate measure for this indicator in contrast to the arithmetic mean. The weighted average considers the varying sizes of the 29 ANSPs, ensuring a balanced representation. Conversely, the arithmetic mean distorts the measure by equally weighing all ANSPs, a misleading estimate if the target is a comprehensive Union-wide inefficiency score.

6.25. By considering the weighted average and the middle point between DEA and SFA, we obtain a potential savings equal to 1 – 0.84 = 0.16, i.e., a 16% cost reduction. For the last year only, i.e., 2019, taking into account the total costs for each ANSP in 2019, the average potential savings suggest that approximately one billion euros in costs could have been saved, of the 5.7 billion spent by the ANSPs in the dataset based on the 2019 PPP/PPI-adjusted costs.

	DEA-VRS			SFA-TL		
	overall period	RP1	RP2	overall period	RP1	RP2
Average	0.71	0.71	0.71	0.88	0.88	0.88
Weighted average	0.79	0.80	0.78	0.89	0.89	0.90

Table 8 - Estimated average ANSP cost efficiency

	Maximum	Minimum	Median
Average	29%	12%	21%
Weighted average	21%	11%	16%

Table 9 - Potential cost savings Union-wide

7. Conclusions and recommendations

Regulatory Benchmarking

7.1. Benchmarking methods, and in particular Data Envelopment Analysis (DEA), and Stochastic Frontier Analysis (SFA), have become well-established and informative tools for purposes of economic regulation. DEA and SFA are now routinely used by European regulators to set reasonable revenue / price caps for energy transmission and distribution system operators for example.

7.2. The cost efficiency of Air Navigation Service Providers (ANSPs) is an important element in the creation of an efficient Single European Sky. Each ANSP serves an individual airspace and in so doing is a natural monopoly. Since there is little direct competition in the market, efficiency is not encouraged by sound competitive pressure.

7.3. Benchmarking allows us to identify best practices, and if ANSPs are asked over time to adjust to best-practice cost, their cost efficiency will converge towards the cost levels of a competitive setting. Hence, instead of competing in the market, we create pseudo competition via benchmarking based regulation, where the ANSPs compete via a model. We note that this issue is particularly relevant in en-route provision given the clear monopolistic status of the ANSPs.

7.4. In this report, we develop two such benchmarking models, and we discuss how to combine them. One is based on data envelopment analysis (DEA) and another on stochastic frontier analysis (SFA). They can be combined in different ways (min, max, average) to determine more or less ambitious cost targets for each individual ANSP.

Methodological differences across models

7.5. Part of the variation of our results can be explained by the nature of the two approaches we have used. In the DEA models, all deviations from the model are classified as inefficiency whilst SFA uses a combination of noise and inefficiency to explain the deviations.

7.6. Furthermore, the SFA model makes more assumptions ex ante, including the structure of the cost function and the existence of competitive prices, which may also be driving some of the differences in the results.

7.7. Finally, we note that DEA, based on an envelopment frontier, has been estimated on an annual basis whereas the SFA model has used panel data and includes an estimate of changes over time. In this context, we have applied an average efficiency approach in the final results.

Results

7.8. We estimate the cost efficiency of 29 Air Navigation Service Providers (ANSPs), using two benchmarking models; the radial, variable returns-to-scale (VRS), Data Envelopment Analysis (DEA) model and the translog, Battese and Coelli (1995) Stochastic Frontier Analysis (SFA) model.

7.9. *DEA Cost Efficiency*: The estimated median efficiency score rose from 61% in 2012 to 85% in 2019, indicating an improvement in cost efficiency over time. When accounting for delays, the efficiency score increased from 73% in 2012 to 90% in 2019. We note that an artefact of all these models is that augmenting the number of variables (dimensions), will result in either a consistent or higher score for the individual ANSP. The results also reveal reduced dispersion in the efficiency scores by the end of RP2, reflecting a decrease in variability among ANSPs' performance.

7.10. *SFA Cost Efficiency*: Total costs were largely explained by flight hours controlled, and the prices of capital and labor. The model indicates that a 1% increase in flight hours, capital prices, and labor prices would lead to a rise in total costs of 0.33%, 0.29%, and 0.56% respectively. Delays did not significantly impact cost efficiency, but indicated that minimizing delays might incur higher costs for ANSPs. Additionally, higher complexity and variability (seasonality) contributed to increased costs.

7.11. *Cost Savings*: The AG finds that ANSPs could save approximately 16% of total costs on average by adjusting to best practices. Based on the 2019 PPP-adjusted costs, this amounts to potential savings of just under one billion euros on an annual basis. Additionally, the report highlights a wide distribution in the efficiency scores, indicating substantial variation in the performance of different ANSPs.

Recommendations

7.12. The large variation in the performance of the multiple ANSPs suggests that a one-size-fits-all approach, such as implementing a universal tariff reduction for all ANSPs, is insufficient. Tailored strategies are necessary to address the specific inefficiencies of each ANSP and maximize potential cost savings. It is therefore natural to work not only with a general cost reduction requirement to capture technological progress (which is around 1% annually over the eight years analyzed in this report) but also to work with additional individual requirements encouraging less efficient ANSPs to catch-up to best practices.

7.13. We suggest that the results could be strengthened over time. There are many ways to do so, including a further investigation of the cost standardization and the inclusion of additional cost drivers such as quality of services provided, including route directness.

7.14. Ideally, all ANSPs should use the same rules for allocating shared costs between en-route and terminal activities (where relevant) and across cost categories. Moreover, the ANSPs should also use standardized depreciation rules which would reduce some of the noise in the data.

7.15. Our analysis presumes that the number of ANSPs are fixed and that the deviation of air space between them remains unaltered. We hereby do not measure the possible gains or cost savings from consolidation of the Single European Sky. Including the United Kingdom, Canada and the United States may change the cost frontier and help to identify potential additional cost savings.

7.16. It is important to note that we only calculate potential savings of the less efficient European ANSPs adjusting to the practices of the more efficient European ANSPs. We do not make comparisons with air navigation services on other continents.

7.17. Reports, such as those produced by the FAA and Eurocontrol[3], seem to indicate that the US system is at least one third more efficient than Europe. In effect, an analysis looking for possible comparators outside of the EU could lead to a much higher savings potential.

7.18. Of course, it might also be that the variation in European efficiencies is larger than that of the US. If this is the case, the bias from using a European perspective only is less important. However, the real impact of economies of scale would only be possible with such a comparison.

Future Directions

7.19. It might be of interest to investigate the possibilities of introducing competition for the market rather than price regulation. In terminal provision, this exists in Sweden, the UK, Germany and Spain. It is likely that such an application to en-route services may lead to a more consolidated set of airspaces that achieve higher economies of scale.

7.20. It is clear that the environmental issues caused by the aviation industry are of growing concern. According to the European Commission, each aircraft flies 49 km longer than necessary on average[4], and this data considers only horizontal flight paths, not the vertical descent paths. The directness of a route is likely to contribute to a reduction in greenhouse gas emissions in the shorter term. Consequently, air traffic control provision could contribute to a reduction in emissions by minimizing the length of flight paths through improved pre-planning and reducing congestion and delays by better balancing demand and supply. Incentivizing such behavior through a hybrid price cap would likely reduce fuel burn in a relatively simple manner.

---

[3] U.S. - Europe continental comparison of ANS cost-efficiency trends (2006-2014) (eurocontrol.int) accessed online on the 31st July 2023

[4] Single European Sky (europa.eu) accessed online on the 31st July 2023.

References

Adler N., Bogetoft P. and Volta N. (2018). Air Navigation Service Providers: Advice on benchmarking of ANSPs and EU-wide cost targets for Reference Review Period 3. *Report for the Performance Review Board.*

Adler N., Brudner A. and Proost S. (2021). Review of transport market modeling using game theory. *European Journal of Operational Research*, 291/3, 808-829.

Adler N., Delhaye E., Kivel A. and Proost S. (2020). Motivating Air Navigation Service Provider Performance *Transportation Research part A: Policy & Pr*actice, 132, 1053-1069.

Adler N., Hanany E. and Proost S. (2021). Competition in Congested Service Networks with Application to Air Traffic Control Provision in Europe. *Management Science*, 68 (4), 2751-2784.

Adler N., Olesen O. and Volta N. (2022). Identifying Merger Opportunities: The Case of Air Traffic Control. *Operations Research forthcoming*, https://doi.org/10.1287/opre.2022.2348.

Alexander, I., and T. Irwin (1996). Price Caps, Rate-of-Return Regulation, and the Cost of Capital. The World Bank Group. N. 87.

Andreana, G., Gualini, A., Martini, G., Porta, F., and D. Scotti, (2021). The disruptive impact of COVID-19 on air transportation: An ITS econometric analysis. *Research in Transportation Economics*. 90, 101402.

Bogetoft, P., and L. Otto (2011). Benchmarking with DEA, SFA, and R. Springer, New York.

Shleifer, A. (1985). A theory of yardstick competition. *The RAND Journal of Economics*. 16, 319-327.

Vickrey, W. S. (1961). Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance*. 16, 8–37.