Performance review body of the single european sky

EU-wide target ranges for RP 3

Annex 2. Air Navigation Service Providers: Advice on benchmarking of ANSPs and EU-wide cost targets

03 June 2018

Air Navigation Service Providers: Advice on benchmarking of ANSPs and EU-wide cost targets

Academic Group
Nicole Adler, Peter Bogetoft, and Nicola Volta

Final Report

2018-06-03

Executive Summary

1. 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.

2. Benchmarking can provide a useful substitute for such settings. 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 cost in a competitive setting. Hence, instead of competing on the market, we create pseudo competition via benchmarking based regulation, where the ANSPs compete via a model.

3. 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. We calculate such cost targets for the individual ANSPs operating different parts of the Single European Sky.

4. We analyze en-route and terminal activities separately rather than gate-to-gate provision. En-route provision has remained a monopolistic service provided by a single ANSP in each Member State. Competition for the market in terminal provision has begun in Sweden, Germany, the UK and Spain. Competition for the market, which auctions the services to the most competitive bidder for a specific, pre-specified timeframe, could replace the need for economic regulation. Indeed, countries in which the auctioning process has been initiated are not required to provide data on the services to the European Commission.

5. We find that the ANSPs could save between 25 and 30% of total costs on average by adjusting to best practices. 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 to capture technological progress, but also to work with additional individual requirements encouraging the less efficient ANSPs to catch-up to best practices. In this report, we focus on these individualized requirements after accounting for the environment including variability and complexity.

6. The results of the two, very different, modeling approaches are subsequently combined. Several approaches are possible including the most conservative, benefit of the doubt approach, in which the highest efficiency estimate is assigned to the individual ANSP. According to this approach, potential cost savings of one billion euros was possible in 2016 in the en-route sector and another 300 million in terminal provision.

7. In summation, the results suggest that the determined unit cost rate for the beginning of reference period 3, were the ANSPs to provide their services efficiently from a cost perspective, could be in the range of 38 standardized 2016 euros and that this should drop by 5% to 36 euros in 2024, after accounting for expected demand growth.

8. We note that the data shows large heterogeneity suggesting that:

   i) standardized accounting practices need to be strengthened such that the cost categories are clearly defined.

   ii) there would appear to be a lack of distinction between en-route and terminal activities.

   These issues lead to substantial uncertainty and, in some cases, differences in the solutions of the models. In the SFA approach, the noise in the data explains 30% of the variation in the en-route sector and this is substantially higher in the terminal sector. Consequently, we would advise the PRB to follow through on reported information in order to ensure that economic information is improved in the near future. It is important to continue to estimate the efficiencies with new data once it becomes available and it would be helpful to speed up the reporting process since a year and a half is unnecessarily long.

9 Finally, given the substantial difference in findings across the ANSPs, after accounting for environmental externalities, it is clear that the price caps set for reference period 3 ought to be individual. A distinguishable set of ANSPs are working a long distance from best practice whereas others do not require so much guidance.

Contents

1. Introduction 6
2. Regulatory benchmarking 8
3. Data and data standardizations 20
4. DEA and SFA models 29
5. Saving potentials 43
6. Final remarks 58
Relevant literature 62

List of figures and tables

Figure 2-1 Systems view of the ANSP....10
Table 2-1 State-of-the-art frontier methods....10
Figure 2-2 Different estimation methods....11
Figure 2-3 Efficiency measurement....12
Table 2-2 Alternative models of the cost structure....17
Table 3-1: Four Indices of Complexity....24
Figure 3-2: Growth in Output and Contextual Variables from 2006 to 2016....26
Table 3-1: Average Cost and Production Variables from 2006 to 2016....27
Table 3-2: Data issues (2016 standardized values)....28
Figure 4-1 An outlier criterion....33
Table 4-1 En-route model in DEA runs....35
Table 4-2 TER model in DEA runs....36
Table 4-3 Variables in stochastic frontier cost function....37
Table 4-4 Results of the en-route stochastic cost function....39
Table 4-5 Results of the terminal stochastic cost function....41
Table 5-1: DEA Efficiency estimates for en-route provision....43
Table 5-2: DEA Efficiency estimates for terminal provision....45
Table 5-3: SFA Efficiency estimates for en-route provision....46
Table 5-4: SFA Efficiency estimates for terminal provision....48
Table 5-5: Potential savings for en-route provision (%)....50
Table 5-6: Potential savings for en-route provision (million euros)....51
Table 5-7: Potential savings for terminal provision (%)....52
Table 5-8: Potential savings for terminal provision (million euros)....53
Table 5-9: Determined Unit Cost Rate per ANSP for RP3 using DEA....55
Table 5-10: Determined Unit Cost Rate per ANSP for RP3 using SFA....56
Table 5-11: Highest Efficient Determined Unit Cost Rate per ANSP for RP3....57

1. Introduction

Background

1.01 Ernst & Young Special Business Services CVBA, Belgium, has been engaged by The European Commission – DG MOVE to assist in the implementation of performance and charging schemes for air navigation services. E&Y has subsequently engaged the Academic Group to help in provision of certain services within the scope of its contract with DG MOVE.

1.02 The task of the Academic Group is to provide advice to the Performance Review Board (PRB) on target setting for cost efficiency in view of RP3.

Overall objective of the Academic Group assignment

1.03 The Academic Group has been asked to provide a report and a model to the PRB with meaningful and scientifically robust EU-wide targets on cost efficiency and on first-best local targets based on a benchmarking of ANSPs from the efficient cost frontier based on a proven models.

Using the information in the regulation process

1.04 This project does not include support for the selection of the regulatory framework. It suffices therefore to note that the benchmarking undertaken supports many types of regulatory frameworks, including cost plus, revenue and price caps, yardstick competition and even concession based auctions (Bogetoft, 2012; Deliverable 5 of COMPAIR, 2018).

1.05 We note that auctions have begun in terminal provision in some countries, namely Germany, Spain, Sweden and the United Kingdom. For greater detail, refer to Deliverable 3 of the COMPAIR project[1], 2018.

1.06 This project also does not include the transformation of benchmarking results into specific requirements within a given regulatory framework. We are not missioned to guide in the selection of a specific rate by which ANSPs can reasonably catch up to best practices.

The report

1.07 In this report, we present the results of our analysis. The presentation is structured as follows:

[1] https://www.compair-project.eu/

- Regulatory Benchmarking
- Data and data standardizations
- DEA and SFA models
- Saving Potentials
- Final Remarks

2. Regulatory benchmarking

2.01 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 / 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.

2.02 We note that the Performance Review Unit (PRU) of Eurocontrol has been collecting data systematically on ANSP services since 2002. We note that substantial work on verifying data is undertaken leading to the likelihood that the information for the timeframe analyzed (2006 to 2016 inclusive) should be reasonably reliable.

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

Benchmarking

2.04 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.

2.05 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.

KPI based

2.06 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.

2.07 Unfortunately, the use of KPIs has its limits.

2.08 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.

2.09 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 the substitution between inputs and between outputs.

2.10 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

2.11 For these – and other – 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 systemic 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, as illustrated in Figure 2-1 below. 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.

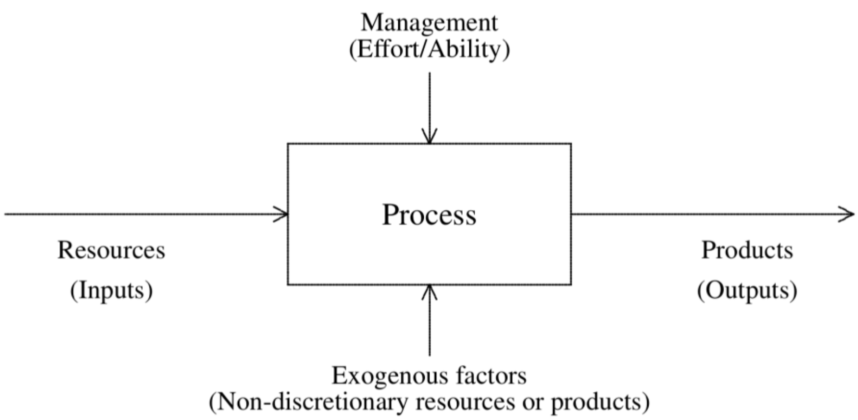

Figure 2-1 Systems view of the ANSP

Frontier methods

2.12 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 2-1 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 2-1 State-of-the-art frontier methods

2.13 The different estimation methods used for benchmarking are basically suggestions for how to compare individual observations, as illustrated by the dots in Figure 2-2 below, given the relationships between input costs and outputs.

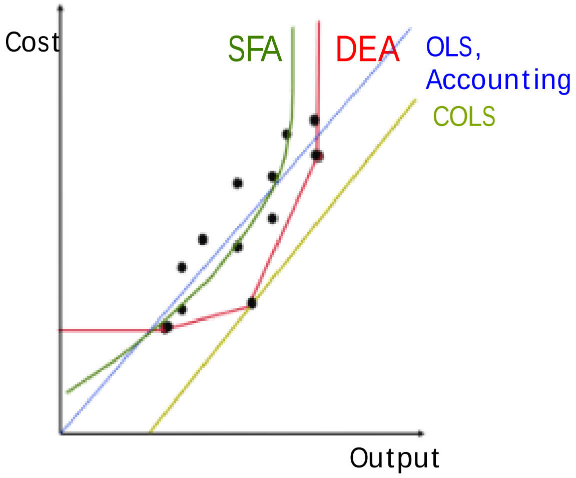

Figure 2-2 Different estimation methods

2.14 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 pros and cons. In this project, we therefore apply both.

Efficiency measures

2.15 Once a benchmarking model linking costs to services, cost drivers and prices has been established, the measurement of efficiency is simple in principle. The cost efficiency of an ANSP can, for example, be defined as:

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

2.16 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.

2.17 The relationship to potential savings is illustrated in Figure 2-3.

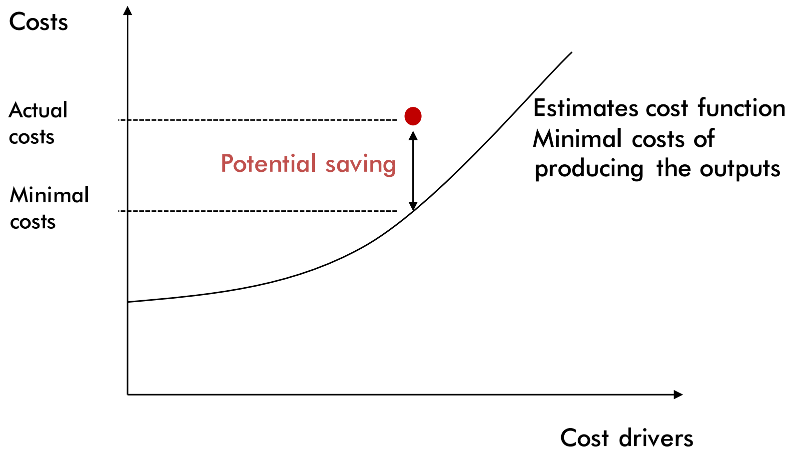

Figure 2-3 Efficiency measurement

The benchmarking process

2.18 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.

2.19 *Choice of variable standardizations*: Choice of accounting standards, cost allocation rules, in/out of scope rules, assets definitions, operating standards etc. is necessary to ensure a good data set from ANSPs with different internal practices.

2.20 *Choice of variable aggregations*: Choice of aggregation parameters, like interest and inflation rates, for the calculation of standardized capital costs, and the search for relevant combined cost drivers, using, for example, engineering information, is necessary to reduce the dimensionality of possibly relevant data.

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

2.22 *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.

2.23 *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 shall be investigated, taking into account the outlier detecting 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. The 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.

2.24 *Model validation*: If data permit, it is also useful to undertake second stage analyses to see is important variables have been excluded. Similarly, second stage analyses can be useful to evaluate the impact of possible facilitating or complicating factors that cannot be represented by good so-called volume measures, e.g. ratio measures like the degree of density of the airspace. With small samples and heterogeneous ANSPs, it may be useful to include also a formal process of submitting, evaluating and resubmitting claims of special circumstances that have not been accounted for by the model.

2.25 It is worthwhile emphasizing that model development is not a linear process but rather an iterative one. During the frontier model estimation, for example, one may identify extreme observations that have resulted from data error not captured by the initial data cleaning or the econometric analyses.

Choosing a good model

2.26 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.

2.27 *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.

2.28 *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 a large unexplained variation. At the same time, there must be a proper balance between the complexity of the model used and the sample size on which we estimate it. 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 well 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. As a measure of the statistical goodness-of-fit of a model, we can for example use the sums of squared inefficiencies, i.e. the sum of squared deviations from the full efficiency score of 1. Formally, this “deviation” measure can be expressed as:

$$D = \sum_{i=1}^{n}(1 - E(M)^{i})^{2}$$

2.29 In this formula, n is the number of ANSPs and $E(M)^i$ is the efficiency of ANSP^i when evaluated in the model M. Clearly, small values of D are an indication that the model M provides a good fit in the sense that the ANSPs in general are evaluated to be close to fully efficient, which is the natural null hypothesis.

2.30 *Intuition and experience*: Intuition and experience is a less stringent but important safeguard against false model specifications and the over- or underuse of data to draw false conclusions. It is important that the models produce results that are not that different from the results one would have found in other studies, countries or related industries. Of course, in the usage of such criteria, one also runs the risk of mistakes. We may screen away extraordinary but true results (Type 1 error) and we may go for a more common set of results based on false models (Type 2 error). The intuition and experience must therefore be used with caution.

2.31 *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. If the model were not good, a high-powered incentive scheme, for example, would not be attractive since it would allocate too much risk to the ANSPs. Lastly, let us mention the trivial but very important requirement to comply with the specific conditions laid out in the regulatory directives of the individual jurisdictions.

2.32 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 or price based costs function

2.33 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.

2.34 We can distinguish two types of cost functions.

2.35 Output based costs function explain cost directly as a function of the services provided and the contexts in which they are provided:

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

2.36 Price based cost functions explain costs by the outputs provided, the prices of input factors, and the contexts:

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

2.37 Both approaches have their pros and cons in a practical, regulatory context.

2.38 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.

2.39 On the other hand, the output case approach also 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 labour are very different across the ANSPs, we would expect them to use different factor combinations – one relying more on labour 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.

2.40 Since both approaches clearly have pros and cons, we have chosen in this project 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 can capture some of the methodological uncertainty of any benchmarking study.

Alternative modelling approaches

2.41 In Table 2-2 below we summarize the different modelling approaches we have considered.

Approach	Details	Est. Method	Data required
A: Direct cost estimation	Totex target = func(Outputs)	DEA and SFA	Totex Outputs
B: Cost estimation via optimization of cost shares	(Labor Cost, Material Cost, Capex) = func(Outputs) defines technology T Totex target = min Labor Cost + Material Cost + Capex in T	DEA and SFA	Totex split Outputs
C: Cost estimation based on technology and prices	(Labor, Material, Capital) = func(Outputs) defines technology T Totex target = min Labor*PriceLabor + Materials*PriceMaterials + Capital*PriceCapital in T	DEA and SFA	Physical inputs Totex Prices
D: The economist notion of cost function	Totex target = func(Outputs,Prices)	SFA	Totex Prices

Table 2-2 Alternative models of the cost structure

2.42 The different approaches to modelling a cost structure, A-D, is likely to lead to different saving potentials.

2.43 Conceptually, one might see the first, A, as the most favorable to the companies and the last, D, as the harshest approach.

2.44 In A, no-one may have picked optimal combination of production factors, i.e. no-one may have been technically and allocatively efficient. Hence estimated best practice costs may be too high. If for example all firms tend to spend too much on capital equipment given then price ratio between labour and capital, we will not see this in A when we estimate best practice.

2.45 In D, we presume that all firms minimize costs and pick the best combination given prices and underlying unknown technology. Hence we do in theory keep the firms responsible for not only being technical efficiency, i.e. for using the least possible of the production factors. We also require that they are allocatively efficient in the sense of picking the right cost-minimizing mix of production factors given the factor prices.

2.46 In practice, however, this ranking may not hold. There are several reasons for this. One of the reasons is that prices are not exogenously given competitive prices as it is implicitly assumed in the D approach. Rather, in our study, they are calculated based on information of cost shares and the use of physical inputs, cf. below. Hence a firm having a small book value, for example, will get a high capital price since depreciation is relatively large compared to the book value, and a firm having a large labour force will for the given labour costs have a small estimated labour price. In other words, the calculated prices will to some degree rationalize the actual input choices of the firms, and hereby make the evaluations fall short of requiring allocative efficiency.

2.47 Moreover, the DEA and SFA estimates will give different results by the nature of the two approaches. DEA has a more flexible mean structure tending to make DEA less harsh, but on the other hand SFA allows for noise in the outcome and does not interpret all model deviations as inefficiency, which tends to make SFA less harsh.

2.48 In summary, we cannot ex ante rank the different approaches A-D using DEA and SFA. We can however see the spectrum of approaches as providing an idea of the possible variation in outcomes that result from best practice methods and which should be taken into account in regulatory applications of the results.

Combining DEA and SFA results

2.49 The DEA and SFA results may be combined in different ways and there are examples of all types of aggregation applied in different regulatory practices across Europe.

2.50 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.

2.51 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.

2.52 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.

2.53 Calculating the average score of the results of the two models, mean(DEA,SFA), would balance the pros and cons of each model equally. This would lead to results similar to that of the interval estimates.

Predicting best practice cost levels

2.54 As part of the applications, regulators may expect changes in the supply of services, and be interested in best practice costs levels in one or more future scenarios.

2.55 If we assume that good predictions of future service levels are available as y^{FUTURE}, such predictions can be established by simply inserting the projected service levels in the estimated cost function.

2.56 In the DEA (A) approach and the SFA (A) approach where no prices are applied, the formula for predicted future costs is:

$$Totex^{FUTURE} = f^{DEA}(y^{FUTURE})$$

2.57 In the SFA (D) approach, the prediction depends not only on future outputs, y^{FUTURE}, but also on future factor prices, p^{FUTURE}. The predicted future best practice costs are:

$$Totex^{FUTURE} = f^{SFA(D)}(y^{FUTURE}, p^{FUTURE})$$

2.58 The best practice costs functions, f^{DEA} and f^{SFA}, are estimated based on best practice historically, which we can denote NOW to be distinguished from FUTURE.

2.59 The advantage of these approaches is that they take into account the structure of the estimated cost functions. When the cost functions are not linear, it does not make conceptual sense to summarize them as for example an expected cost per output. The cost per unit of one output will vary with the scale of the firms as well as with the mix of services to be produced.

2.60 The above approaches could be refined by including also a general improvement of best practices from the last year of data used in the estimations, i.e. NOW, till the year of the projected future demand, i.e. FUTURE. In theory, one could even make such general productivity improvements firm specific using for example a Malmquist approach.

2.61 This would lead to predictions using formula like

$$Totex^{FUTURE} = f^{SFA}(y^{FUTURE}, p^{FUTURE})(1 - TP)^{FUTURE-NOW}$$

$$Totex^{FUTURE} = f^{DEA}(y^{FUTURE})\ (1 - TP)^{FUTURE-NOW}$$

2.62 Here, TP measures the rate of technological progress, i.e. the yearly improvement in best practice. This could for example be 1% in an industry with heavy infrastructure investments, and 3-4% in a more rapidly changeable service sector.

3. Data and data standardizations

3.01 In this chapter, we will briefly discuss the available data, including the cost measures, the main service dimensions and the main contextual variables, on which our models rely.

3.02 The collection, cleaning and to a large extent the standardization of data is not part of this project. It relies principally on ATM Cost Effectiveness (ACE) data that are collected and validated by the Performance Review Unit (PRU) of Eurocontrol in cooperation with ANSPs. The data is publicly available on the Eurocontrol website[2].

3.03 To measure the inputs or resource usage of entities like an ANSP, it is common to use monetary cost values.

3.04 Moreover, in regulation, it is common to recommend the so-called Totex benchmarking approach.

$$Totex = Opex + Capex$$

3.05 Totex represents the total expenditures and it is the sum of the operating expenditure, Opex, and capital expenditure, Capex. Economists think of Opex as an indication of the variable costs, VC, and of Capex as an indication of the fixed costs, FC.

3.06 Using Totex as the input of our system model, the ANSPs are held responsible for the task of finding the optimal balance between operating and capital expenses. Of course, regulatory usage and regulatory history may call for other approaches. If Capex has been partly dictated or recognized by a regulator, for example, it may sometimes be interested to focus on Opex efficiency alone. Ideally, in such cases, one should still include Capex as a separate, non-controllable input to make sure that possible trade-offs between Opex and Capex are taken into account.

3.07 Unfortunately, raw accounting data from the ANSPs might not be ideal, since depreciation periods differ, rules for activations vary, the scope of the accounting differs, etc. Considerable cost standardization is therefore needed before the accounting costs can be compared.

3.08 As explained above, the data standardizations are largely the responsibility of the PRU. However, we have introduced additional

[2] http://www.eurocontrol.int/prc/publications

standardization of cost to make these comparable across service providers and to calculate appropriate production factor prices.

3.09 Standardization ensures that the econometric cost function is homogeneous and in alignment with the underlying economic theory on cost functions (Coelli et al., 2005). In this process, we have relied on a series of indices, including purchase power parity, exchange rates for countries outside the euro, inflation rates, the producer price index and the cost of borrowing index.

3.10 Purchase Power Parity (PPP) indicates how many currency units a particular quantity of goods and services costs in different countries. PPPs can be used as currency conversion rates to convert expenditures expressed in national currencies into an artificial common currency thus eliminating the effect of price level differences across countries.

3.11 Exchange Rates are needed to make cross-country comparisons where purchase price parity is not applied and countries utilise different currencies. Since 19 of the 28 EU countries are within the euro zone, the remainder will require the data to be PPPd or the exchange rate needs to be applied.

3.12 Inflation rates are necessary to account for time differences, i.e. to make cost data from different periods comparable. The PRU applies the historical inflation figures from EUROSTAT[3] or from the International Monetary Fund[4], for countries not included within the Eurostat database. In this analysis, the base year is 2016 and the inflation rates are applied to earlier years accordingly.

3.13 In general, producer price indices measure the average movements of prices paid by domestic producers for goods and services sold on the domestic or/and on the export markets between one time period and another. The Producer Price Index is used to represent the cost of purchasing materials and supplies from local producers.

3.14 The Capital Goods Price Index is based on the cost of borrowing indicator which is constructed following the European Commission's Joint Research Center guidelines (JRC, 2008). We focus on the cost of loans as a proxy for the ANSP cost of capital. Since the ANSPs are regulated entities subject to price-cap regulation, in practice the cost of equities is generally in line with the cost of loans. We collect the cost of borrowing from statistics publicly available at the European Central Bank. We use the interest rate data on "new business" to non-financial

[3] http://ec.europa.eu/eurostat/web/main/home
[4] http://www.imf.org/external/pubs/ft/weo/2017/01/weodata/index.aspx

corporations over 5 years initial rate fixation. The index is inflated to 2016 by using the Euro area inflation rate source from EUROSTAT.

3.15 Consequently, we apply a cost of capital that is in line with recommendations from the theory (Coelli et al., 2005).

Physical production factors

3.16 In the estimation of price-based cost functions, we also include measures of the physical amount of different production factors utilized by the ANSPs, as well as the factor prices.

3.17 In general, we can think of the production factors as the set of resources required for the generation of the services provided, often classified into four specific groups; land (including all natural resources), labour (including all human resources), capital (including all man-made resources), and management (which brings all the previous resources together for production).

3.18 Capital is estimated according to the net book value as reported by ANSPs and assembled in the ACE database. The net book value of the equipment and other capital assets utilized for en-route operations has been separated from the terminal operations. This is the only information that is not published openly in the ACE reports but was provided to the group for the purposes of this analysis.

3.19 Two categories of staff work for the air navigation service providers, namely air traffic controllers (ATCOs) and support staff. Air traffic controllers' hours and full time equivalent numbers are reported separately for en-route and terminal activities.

3.20 Support staff, which consist of many types of employees, ranging from technicians to meteorologists to management, are reported in full time equivalents but only on a gate-to-gate basis. Consequently, it is not possible to separate the support staff between en-route and terminal activities.

Pricing factors

3.21 The price of capital is based on the net book value, cost of capital and depreciation costs as recorded in the ACE database, and the capital goods price index based on the cost of borrowing as described in paragraph 3.14. The price of capital is created separately for en-route and terminal operations using net book value data for each operation (this is the only information that is not published openly in the ACE reports but was provided to the group for the purposes of this analysis). The price of capital is measured as follows:

$$Price\ of\ Capital = \left(\frac{Depreciation\ Costs + Cost\ of\ Capital}{\frac{NBV}{Cost\ of\ Borrowing\ Index}} \right)$$

3.22 The price of staff is based on the total cost of labour divided by ATCO hours, eith er en-route or terminal, since it was not possible to separate such costs across ATCOs and support staff individually.

$$Price\ of\ Staff = \left(\frac{Cost\ of\ Staff}{ATCOs\ hours\ in\ operation} \right)$$

3.23 The price of energy, supplies, materials and outsourcing are determined by the Member States' producer price index (PPI), as published in Eurostat.

Service measures (Outputs)

3.24 ANSP output involves the safe separation of flights. The network manager estimates the total IFR flight hours controlled according to the entry and exit times of an aircraft through the specific airspace. The en-route movements are controlled through air control and approach centers.

3.25 The terminal activities are undertaken in airport towers and terminal approach centers, depending on the complexity of the airspace. The terminal output is measured in IFR flight movements.

Contextual characteristics (Exogenous)

3.26 Contextual variables in general refer to the environment in which the ANSP must work but over which the management has little to no control. In the case of ANSPs, this refers to seasonality and complexity for the most part.

3.27 To capture the workload provided by the ANSPs, it is necessary to also consider the complexity of the flight paths being handled. The ACE working group on complexity produced a definition in 2006 and the data has been subsequently estimated on an annual basis. The reason for creating the index was precisely to enable such benchmarking as reported here. The complexity index is created based on four indices as depicted in Table 3-1.

Complexity Dimension	Indicator	Description
Traffic density	Adjusted density	A measure of the potential number of interactions between aircraft in a given volume of airspace.
Traffic in evolution	Potential vertical interactions (VDIF)	Captures the potential interactions between climbing, cruising and descending aircraft.
Flow structure	Potential horizontal interactions (HDIF)	Provides a measure of the potential interactions based on the aircraft headings.
Traffic mix	Potential speed interactions (SDIF)	Assesses the potential interactions based on the aircraft speeds.

Table 3-1: Four Indices of Complexity

3.28 To create a single index, the three potential interactions are summed and subsequently multiplied with the adjusted density.

3.29 Although the complexity index is mostly describing the en-route traffic control workload, it also includes an element relevant to terminal activities, in particular traffic evolution. Consequently the index has been applied to both en-route and terminal analysis.

3.30 Variability of the traffic load over time is also likely to impact the relative cost base for the ANSPs. Variability is computed by dividing traffic levels in the peak month by the average monthly traffic. Variability is likely to impact overall annual costs if it is not possible to employ ATCOs on a seasonal basis thus leading to higher annual costs as compared to an ANSP with similar output spread evenly over the year.

3.31 Complexity and Variability are characteristics of the air traffic controlled by the ANSPs and can be included in the benchmarking analysis in different ways. Most commonly, they are included as explanatory variables in the inefficiency model (parametrizing the inefficiency distribution in SFA or used as explanatory variables in second stage analysis in DEA) or they are used to construct additional volume based output measures that can be thought of directly as outputs along TotalIFRhours and TotalIFRmovements. To enable the latter approach in the DEA models, we have constructed the following additional variables:

- Complexity_Volume_ IFRhours= Complexity* TotalIFRhours
- Variability_Volume_IFRhours= Variability* TotalIFRhours
- Complexity_Volume_IFRmovements= Complexity* TotalIFRmovements
- Variability_Volume_IFRmovements= Variability* TotalIFRmovements

3.32 One can think of the variability volume measures, for example, as measures of the capacity of the individual ANSPs. Capacity to handle a large traffic load is an important service and therefore a natural output (in energy network companies, we similarly find that peak-load is a natural cost driver– indeed capacity is a more important cost driver in general that the actual flow of energy).

3.33 The size of the airspace may also be considered as a contextual variable over which the ANSPs have no control. An alternative measure of capacity may be the number of area control sectors open at maximum configuration, which is relevant for the en-route activities. This variable may also be managed to some extent, as with variability.

More information about the outputs and context variables

3.34 Table 3-1 summarizes the average data over time, including the contextual variables, comprising variability, complexity and the maximum number of sectors. All cost data is presented in standardized 2016 euro values.

3.35 As can be seen in Figure 3-2, complexity has grown the most over the decade by around 22% alongside growth in output flight hours controlled of around 11%. Terminal IFR movements decreased over the period by around 5%.

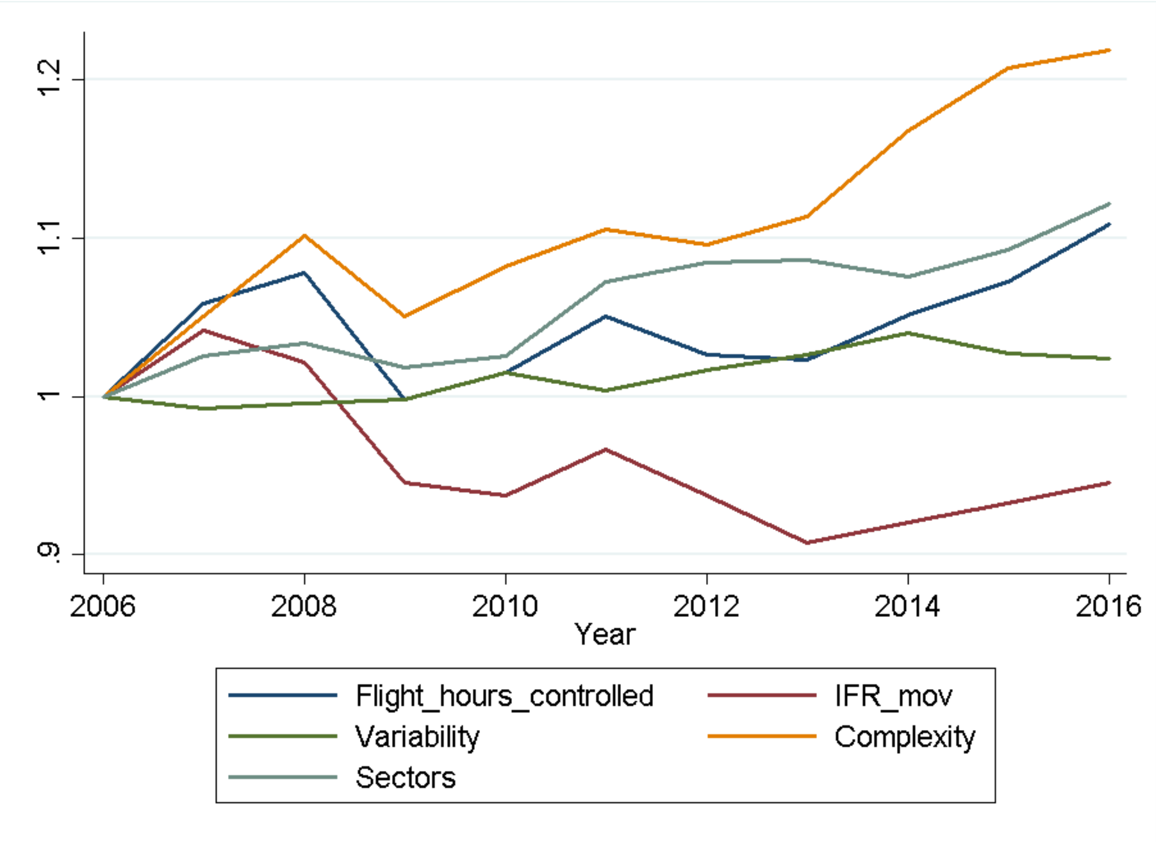

Figure 3-2: Growth in Output and Contextual Variables from 2006 to 2016

		2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
En-route	NBV (Euros)	167,856,120	165,154,414	160,200,592	153,159,142	149,221,868	148,849,798	146,388,474	142,868,893	143,755,439	139,274,876	139,496,847
	Staff-costs (Euros)	118,024,620	124,570,102	127,762,243	128,879,675	123,791,442	123,851,978	125,223,791	124,362,337	125,206,655	127,525,530	127,091,335
	Non-staff-costs (Euros)	34,347,298	33,855,872	31,862,276	33,029,001	31,665,598	32,610,898	30,921,362	29,865,333	29,996,630	30,042,340	28,457,799
	Cost of capital (Euros)	12,220,599	12,504,408	12,442,102	12,259,086	11,033,004	11,920,567	12,748,718	11,996,014	12,482,269	12,786,070	12,142,084
	Depreciation (Euros)	25,721,855	24,755,421	22,632,517	23,851,851	23,974,857	23,628,625	23,505,642	22,876,505	23,440,384	23,749,882	23,829,115
	Flight hours controlled	408,826	432,827	440,639	408,111	414,966	429,383	419,732	418,474	429,970	438,333	453,326
TER	NBV (Euros)	84,877,304	82,356,262	63,736,015	69,280,928	68,893,993	60,757,827	56,782,361	53,350,460	49,390,353	47,662,115	45,174,730
	Staff-costs (Euros)	40,524,654	44,636,555	43,836,726	44,130,451	39,458,645	39,570,081	40,207,416	38,531,869	38,175,690	39,630,905	38,666,396
	Non-staff-costs (Euros)	10,142,549	11,208,383	10,312,589	10,688,962	11,253,831	10,940,422	9,786,282	9,177,079	9,625,431	9,020,058	8,763,518
	Cost of capital (Euros)	2,987,467	3,154,435	3,260,512	3,145,377	2,789,870	2,790,252	2,307,203	2,229,887	2,262,883	2,400,838	2,381,154
	Depreciation (Euros)	6,024,274	6,458,495	6,246,094	7,000,902	6,815,982	6,456,858	5,721,002	5,643,075	5,433,931	5,154,351	4,715,202
	IFR movements	505,682	526,938	516,438	477,990	473,878	488,846	473,930	458,906	465,286	471,641	478,184
	Variability	1.200	1.191	1.195	1.198	1.218	1.204	1.220	1.232	1.248	1.232	1.228
	Complexity	4.683	4.920	5.158	4.922	5.069	5.180	5.133	5.216	5.470	5.654	5.707
	Maximum number of Sectors	19.700	20.200	20.367	20.067	20.200	21.133	21.367	21.400	21.200	21.533	22.100

Table 3-1: Average Cost and Production Variables from 2006 to 2016

3.36 The relative percentages of the four cost categories have remained relatively constant over the decade analyzed.

3.37 When analyzing en-route data, staff costs are around 63%, operating costs are 19%, cost of capital is 6% and depreciation is 12%.

3.38 When analyzing terminal data, staff costs are around 69%, operating costs are 17%, cost of capital is 4% and depreciation is 10%.

3.39 We note that some of the data shows substantial spikes, for the most part in the capital related expenditures as demonstrated in some examples in Table 3.2.

Finland	**en-route NBV**
2013	20,887,714
2014	16,831,462
2015	23,488,580

Finland	**terminal NBV**
2006	656,558,404
2007	663,107,180
2008	40,579,390
2009	5,037,920
2010	43,692,932

Greece	**en-route NBV**
2013	100,422,098
2014	98,998,900
2015	5,250,129
2016	5,849,108

Netherlands	**en-route staff cost**
2008	82,806,449
2009	101,582,552
2010	84,097,171

UK	**en-route non staff operating cost**
2013	112,197,243
2014	97,574,282
2015	117,241,471

Poland	**en-route cost of capital**
2013	4,674,790
2014	474,688
2015	9,936,848
2016	11,086,963

Table 3-2: Data issues (2016 standardized values)

4. DEA and SFA models

Data Envelopment Analysis

4.01 The non-parametric DEA approach uses linear programming to evaluate the performance of the firms or organization.

4.02 In the DEA literature is common to refer to the evaluated as Decision Making Units (DMUs). A DMU can be an observation of (inputs, outputs) for a firm at a given time (cross section) or at other time periods (panel data).

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

4.04 In the 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, (x^k,y^k), k = 1,...,K and satisfies certain technological assumptions specific to the given approach.

4.05 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 speculations 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 — or, as consultants might put it, “we bend ourselves backwards to make you look as good as possible”.

4.06 Apart from the sales talk, it is important to understand 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 chocks, bad reporting practice or ambiguity in accounting practices, for example — the result may not be an inner approximation of the true possibilities. If there is considerable noise in the data, one can even argue that firms may be evaluated against the hardest possible standards (possibly the luckiest firms) and not against a cautious standard.

4.07 The basic DEA models mainly differ in the assumptions that they make about the technology T. The most important assumptions are:

- A1 Free disposability: We can produce less with more.
- A2 Convexity: A weighted average of feasible production plans is feasible.
- A3 Scaling: Production can possibly be scaled.
- A4 Additivity, replicability: The sum of two feasible production plans is feasible.

Returns to scale

4.08 In words, these assumptions about returns to scale can be interpreted as follows:

- Constant Return to Scale (CRS) means that we do not believe there to be significant disadvantage of being small or large
- Non-Increasing Return to Scale (NIRS), sometimes referred to as Decreasing Return to Scale (DRS), means that there may be disadvantages of being large but no disadvantages of being small
- Non-Decreasing Return to Scale (NDRS), sometimes referred to Increasing Return to Scale (IRS), means that there may be disadvantages of being small but no disadvantages of being large
- Variable Return to Scale (VRS) means that there are likely disadvantages of being too small and too large.
- Free Disposability Hull (FDH) means that we make no ex ante assumptions about the impact on size and that we even do not assume that we can make linear interpolation (convex combinations) between two points
- Free Replicability Hull (FRH) is like FDH except that we also assume that we can combine (replicate) individual units, such that we can compare for example a large TSO with the sum of a small and a medium sized TSO.

4.09 The strongest assumption is CRS. It leads to the lowest efficiency scores.

4.10 Conceptual reasoning as well as statistical tests can be used to decide what the correct scale assumption is.

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

4.12 Conceptually, one can also argue that it is difficult to justify disadvantages of being large since reorganization can in such cases often be introduced to avoid the problems of coordination in large entities. In the case of air control, this is less obvious, since flight paths have to be coordinated among different parts of the European sky.

4.13 In our case, therefore, we have largely relied on statistical testing to decide on the proper assumption. Starting from VRS, we have tested if it is acceptable to assume NDRS or even CRS, and in we have found that this is the case. Our maintained assumptions in the DEA models is therefore CRS

Outliers

4.14 Outlier analysis consists of screening extreme observations to make sure that they can actually be considered natural parts of the data set, i.e. that the data has been generated by the data generation process.

4.15 Depending on the approach chosen (DEA or SFA), outliers may have 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.

4.16 In non-parametric methods, extreme observations are such that dominate a large part of the sample directly or through convex combinations. Usually, if erroneous, they are fairly few and may be detected different programming and statistical methods. The outliers are then systematically reviewed in all input and output dimensions to verify whether the observations are attached with errors in data. The occurrence and impact of outliers in non-parametric settings in mitigated with the enlargement of the sample size.

4.17 There are several possible outlier detection techniques that can be relevant for DEA model, c.f. also Bogetoft and Otto (2011) and Wilson (1993).

4.18 Here we will restrict attention to the techniques that have commonly been used in regulatory benchmarking. We will draw on our experience from EuropeanDistribution System Operator (DSO) benchmarking.

4.19 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 units – or there may be some mistakes in the reported numbers.

4.20 The other approach is to investigate the impact on specific firms from sequentially eliminating DMUs using the notion of efficiency ladders. If the elimination of one DMU leads to a significant increase in the efficiency of sufficiently number of units, there are again good reasons to check this unit more carefully.

4.21 Thirdly, one can use shell analysis where the idea is to study the impact of groups of DMU, like the ones in the first shell, the second shell etc, c.f. also Agrell and Bogetoft (2000). As the cost function is peeled this way, one shall check the shells with a significant impact on efficiency while there is less reason to continue the controls when the average efficiency is only improving slightly when a shell is eliminated. The Danish water regulation is using a peeling approach by simply eliminating the first shell before doing the estimation that counts.

4.22 A fourth criterion, and one of the criteria used in the German regulation, is that a single DSO should not have a too large of an impact on the average efficiency. We can operationalize the average efficiency impact criterion analogous to the tests used above. Specifically, let I be the set of n DSO is the data set and i be a potential outlier. Also let E(k,I) be the efficiency of k when all DSO are used to estimate the technology and let E(k;I\i) be the efficiency when DSO i does not enter the estimation. We can therefore evaluate the impact on the average efficiency by:

$$\sum_{k \neq i} \frac{(E(k; I \backslash i) - 1)^2}{(E(k; I) - 1)^2}$$

4.23 Small values of this as evaluated in a F(n-1,n-1) distribution, c.f. Banker(1996), will be an indication that *i* is an outlier.

4.24 The last and final criterion is the extreme super efficiency criterion. The idea is to eliminate DSO that are far outside the technology spanned by the other DSOs. Again, this is used in the German regulation where the criteria is operationalizes as follows: A DSO I is classified as an outlier if:

$$E(i; I \backslash i) > q(0.75) + 1.5 \cdot [q(0.75) - q(0.25)]$$

where q(a) is the a fractile of the distribution of super-efficiencies, such that e.g. q(0.75) is the super-efficiency value that 75% has a value below. We see that the difference q(0.75)-q(0.25) is also what is sometimes called the Inner Quartile Range IQR and the cut-off is what is sometimes called the inner fence in box plots. To get an idea of the cut-off levels in a standard normal distribution, N(0,1), see c.f. Figure 4-1 below

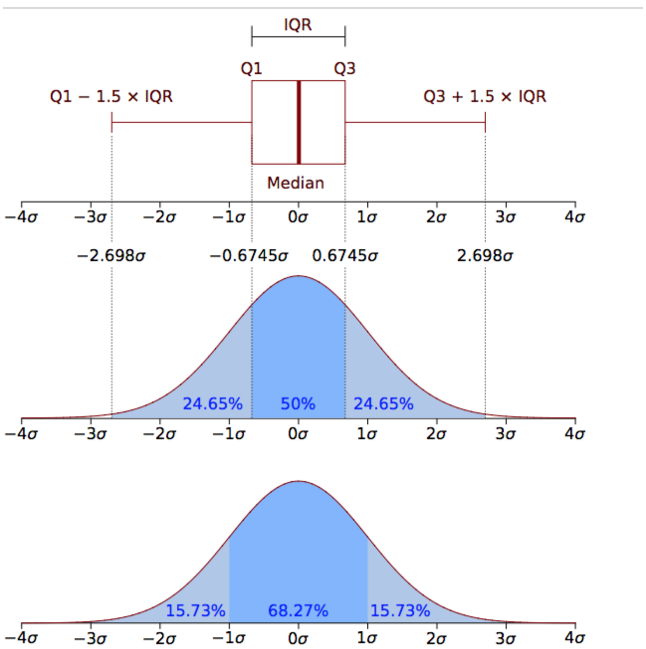

Figure 4-1 An outlier criterion

Reference set

4.25 A DMU can be an observation of (inputs, outputs) for a firm at a given time (cross section) or at other time periods (panel data). When panel data is availed, there are several possibilities to construct the reference set used to estimate the technology.

4.26 One possibility is to pool the data from all periods.

4.27 In that case, care should be taken to correct the economic measures for price changes during the period.

4.28 The advantage of using yearly data as different observation is that more data points are available to support the estimation.

4.29 From a theoretical perspective, a disadvantage of using yearly data in this way is that firms may end up “competing” against themselves in previous years, which in turn leads to so-called Ratchet effects, i.e. firms avoid performing too well in some years not to face too harsh conditions in the future. In practice this is probably not a significant issue, except perhaps for the most efficient DSOs that have very peculiar combinations of cost drivers.

4.30 From an applied perspective, a much more relevant disadvantage of the use of yearly data is that yearly data may be more noisy than average data. This happens for example if there are difficulties getting the periodization of costs entirely right.

4.31 In summary, it is important to understand that the estimation of a DEA model on several years of data does not come with the same benefits as in the SFA approach.

4.32 Another possibility is to analysis year by year, i.e. to only compare a DMU in one year with the other DMUs from that year. This is the approach taken here since we avoid the problem of mistakes and noise in data accumulating by the minimal costs decreasing monotonously in the size of the reference set.

4.33 There are other approaches as well.

4.34 One can for example take as observations the average of a DMUs data from the last three year. This avoids some of the problems by noise in the data in a given year, but obviously raises the question of which years to average and how much weight to put on recent compared to more distant years.

4.35 Also, and this is the approach we have used, one can do yearly analysis but investigate how the efficiency of the DMUs varies from one year to the next.

En-route and Terminal model specifications

4.36 Using the modelling principles above, we have determined our proposed En-route and Terminal models. The model specifications are summarized in Table 4-1 and Table 4-2 below.

Model	ACC base model
Inputs	
TC_ACC	Cost efficiency estimated on Totex PPP corrected
Outputs	
Flight_hours_controlled_ACC	Total flight hours
Complexity_Volume_hours	Complexity* Flight_hours_controlled_ACC
Variability_Volume_hours	Variability* Flight_hours_controlled_ACC
Sectors	Number of Sectors
Reference set	
Data from same year	Yearly analysis 2006-2016
Estimation approach	
	Constant returns to scale
dea_crs_ex_output	Outliers eliminated using super-efficiency and average impact criteria

Table 4-1 En-route model in DEA runs

4.37 In the en-route model (ACC), we have four cost drivers. The first, Flight_hours_controlled_ACC, is a direct measure of the workload, the second, Complexity_Volume_hours, is a workload measure that is corrected for the complexity of controlling, the third, Variability_Volume_hours, is a measure of the capacity for handling a large workload at least temporarily, and the fourth, Sectors, again is a measure of the size of the operation and the actual and potential workload.

Model	TER base model
Inputs	
TC_TER	Cost efficiency estimated on Totex PPP corrected
Outputs	
IFR_mov_TER	Number of movements
Complexity_Volume_Mov	Complexity* IFR_mov_TER
Variability_Volume_Mov	Variability* IFR_mov_TER
Reference set	
Data from same year	Yearly analysis 2006-2016
Estimation approach	
	Constant returns to scale
	Outliers eliminated using super-efficiency and average impact criteria
dea_crs_ex_output	

Table 4-2 TER model in DEA runs

4.38 In the TERminal model, we have used similar cost drivers as in the En-route model, except that Sectors have not been introduced. The relevance of Variability is again to account for peak workload. The complexity measure contains a part that is related to airport movements, but is of course less directly relevant here than in the enroute model.

4.39 Indeed, as our results will show, the rationale for the TERminal model is weaker that the rationale for the en-route model. This explains also why the DEA and SFA models give more deviant results in the TERminal case compared to the en-route case. When the cost drivers are less obvious, a DEA approach will tend to interpret the deviations as large inefficiencies, while a SFA approach will interpret the outcome as governed by a large amount of noise.

Stochastic Frontier Analysis

4.40 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.

4.41 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. We apply a translog functions. 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).

4.42 From the dataset, we apply the model to the set of variables described in Table 4-3, where the cost of operation index equals the Producer price index.

4.43 Total costs and prices are normalised by one of the prices in order to ensure homogeneity condition.

Dependent Variable	
	$\frac{\text{total cost}}{\text{producer price index}}$
Independent Inputs	
Output	total IFR flight hours controlled (en-route) and total IFR airport movements (terminal)
Labour price	$\frac{\text{total staff cost/ATCO hours}}{\text{producer price index}}$
Capital price	$\frac{\text{(depreciation cost + cost of capital) / (NBV/ capital goods price index)}}{\text{producer price index}}$
Environmental Variables	
Airspace characteristics	seasonality, complexity, maximum number of sectors, time trend

Table 4-3 Variables in stochastic frontier cost function

4.44 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.

4.45 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). We present the results of Battese and Coelli (1995) because this model provided the most reasonable estimations according to the log likelihood estimates.

4.46 The SFA model applied to en-route air traffic control provision is presented in equations 4-1. The producer price index is referred to as PPI.

$$
\begin{aligned}
&\ln\left(\frac{Total\ Cost_{it}}{PPI_{it}}\right) \\
&= \beta_0 + \beta_1 \ln(IFR\ flight\ hours_{it}) + \beta_2 \ln\left(\frac{Labor\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_3 \ln\left(\frac{Capital\ Price_{it}}{PPI_{it}}\right) + \beta_4 \frac{1}{2} \ln(IFR\ flight\ hours_{it}) \ln(IFR\ flight\ hours_{it}) \\
&+ \beta_5 \frac{1}{2} \ln\left(\frac{Labor\ Price_{it}}{PPI_{it}}\right) \ln\left(\frac{Labor\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_6 \frac{1}{2} \ln\left(\frac{Capital\ Price_{it}}{PPI_{it}}\right) \ln\left(\frac{Capital\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_7 \ln(IFR\ flight\ hours_{it}) \ln\left(\frac{Labor\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_8 \ln(IFR\ flight\ hours_{it}) \ln\left(\frac{Capital\ Price_{it}}{PPI_{it}}\right) \\
&+ \beta_9 \ln\left(\frac{Labor\ Price_{it}}{PPI_{it}}\right) \ln\left(\frac{Capital\ Price_{it}}{PPI_{it}}\right) + \beta_{z1} \ln(Complexity_{it}) \\
&+ \beta_{z2} \ln(Variability_{it}) + \beta_{z3} \ln(Sectors_{it}) + \beta_{z3} \ln(time_t)\, \beta_{z3} + v_{it} + u_{it}
\end{aligned}
\tag{4.1}
$$

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

4.47 The results of the stochastic cost function models with respect to en-route services are presented in Table 4-4. Two models are analysed, namely without and with a time trend variable that estimates market level changes. All cost elements are PPPd to allow for international comparisons. All variables are logarithm transformed and normalized by the geometric mean.

En-route		
Variable	**Model (i)**	**Model (ii)**
Output	0.44 (0.04) ***	0.44 (0.04) ***
Staff Prices	0.60 (0.03) ***	0.60 (0.03) ***
Capital input price	0.18 (0.02) ***	0.14 (0.02) ***
Staff Prices^2	0.49 (0.08) ***	0.50 (0.08) ***
Capital input price^2	0.12 (0.06) **	0.12 (0.06) **
Output^2	0.11 (0.02) ***	0.11 (0.02) ***
Staff Prices * Capital input price	-0.13 (0.04) ***	-0.13 (0.04) ***
Output* Staff Prices	-0.20 (0.03) ***	-0.18 (0.03) ***
Output* Capital input price	0.11 (0.02) ***	0.11 (0.02) ***
Complexity	0.55 (0.16) ***	0.58 (0.17) ***
Sectors	0.53 (0.04) ***	0.54 (0.05) ***
Variability	2.14 (0.29) ***	2.47 (0.30) ***
Time		-0.01 (0.00) **
Constant	-3.73 (0.40) ***	-3.97 (0.39) ***
Inefficiency (m)		
Complexity	-0.81 (0.15) ***	-0.81 (0.17) ***
Constant	0.89 (0.21) ***	0.93 (0.22) ***
σ_u (inefficiency error component)	0.20 (0.01) ***	0.19 (0.01) ***
σ_v (stochastic noise error component)	0.11 (0.02) ***	0.11 (0.02) ***
λ	1.73 (0.03) ***	1.69 (0.03) ***
Log-Likelihood	59.6	63.4

Table 4-4 Results of the en-route stochastic cost function

*Notes: The analysis is based on 327 observations. Indicators: *significant at the 10 per cent confidence level; ** significant at 5 per cent confidence level; *** significant at 1 per cent confidence level.*

4.48 The results of the two models are consistent and the three independent variables, namely output, staff and capital, are positive and significant as expected. In other words, the higher any of these three elements, the higher is the expected total cost.

4.49 Time is both significant and negative in model (ii). This indicates that there has been a slight decrease in costs over the timeframe analysed. Given the higher log-likelihood, we analyse the efficiency results of model (ii).

4.50 The size variable that appears the most relevant is the maximum number of sectors as compared to airspace, which in many models was insignificant. In other words, the larger the number of maximum sectors required, the greater is the expected annual costs.

4.51 Variability creates substantial additional costs. This variable suggests that ANSPs handling the same number of controlled hours as another but with a substantial peak in demand over a month or more will require more resources than an ANSP with more constant yet equivalent demand over the year.

4.52 Complexity also impacts the costs of an ANSP but in two different directions. First, complexity increases workload above and beyond the hours controlled i.e. creates more work than estimated directly. On the other hand, the management of ANSPs with higher complexity are apparently better able to generate an efficient system which thus explains the efficiency score u_{it}. This is in line with previous analyses (c.f. CEG, 2011; Adler et al., 2017).

4.53 Both the inefficiency and noise components are statistically significant and clearly separable in the analysis, with lambda greater than 1.5. We note that, based on the values of lambda, the inefficiencies explain approximately 70% of the variation in the data and the rest is random noise. This clearly suggests that there is substantial inefficiency in the system being analysed.

Terminal Stochastic Cost Function

4.54 The results of the stochastic cost function models with respect to terminal services are presented in Table 4-. Two models are analysed, namely without and with a time trend variable that estimates market level changes. All cost elements are PPPd to allow for international comparisons. All variables are logged and normalized by the geometric mean.

TERMINAL		
Variable	**Model (iii)**	**Model (iv)**
Output	0.84 (0.01) ***	0.84 (0.01) ***
Staff Prices	0.60 (0.03) ***	0.59 (0.03) ***
Capital input price	0.10 (0.02) ***	0.08 (0.03) **
Staff Prices^2	0.12 (0.07)	0.12 (0.07)
Capital input price^2	0.05 (0.03)	0.03 (0.03)
Output^2	0.08 (0.02) ***	0.07 (0.02) ***
Staff Prices * Capital input price	0.06 (0.04)	0.05 (0.04)
Output* Staff Prices	-0.09 (0.03) **	-0.09 (0.03) **
Output* Capital input price	0.02 (0.02)	0.01 (0.02)
Variability	4.34 (0.40) ***	4.53 (0.41) ***
Time		-0.01 (0.00) *
Constant	-3.34 (0.32) ***	-3.44 (0.33) ***
Inefficiency (m)		
Complexity	-0.27 (0.13) ***	-0.26 (0.11) **
Constant	0.46 (0.16) ***	0.49 (0.17) ***
σ_u (inefficiency error component)	0.16 (0.05) ***	0.16 (0.05) ***
σ_v (stochastic noise error component)	0.23 (0.02) ***	0.22 (0.03) ***
λ	0.70 (0.07) ***	0.72 (0.08) ***
Log-Likelihood	-37.5	-35.5

Table 4-5 Results of the terminal stochastic cost function

*Notes: The analysis is based on 307 observations. Indicators: *significant at the 10 per cent confidence level; ** significant at 5 per cent confidence level; *** significant at 1 per cent confidence level.*

4.55 The results of both models are consistent and the three independent variables, namely output, staff and capital, are positive and significant as expected. In other words, the higher any of these three elements, the higher is the expected total cost.

4.56 The time trend (model iv) is negative but with a small significance level. This indicates that almost no cost reductions have been achieved over time in terminal air traffic control provision.

4.57 As with the en-route stochastic cost function, variability increases costs substantially.

4.58 The complexity index for the most part describes en-route traffic provision but one element does also explain airport complexity. Complexity did not prove significant in the SFA cost function but it does explain efficiency (*u*). Since the management of the ANSPs are in control of both en-route and terminal provision in most countries today, it is reasonable that complexity is relevant in explaining managerial efficiency in models (iii) and (iv) as it was in models (i) and (ii). Complexity would appear to consistently help management ensure more efficient use of resources.

4.59 Both the inefficiency and noise components are statistically significant and clearly separable in the analysis. The model assigns the larger part of variance to the random noise (as highlighted by a lambda lower than 1). Consequently, it would appear that it is necessary to work on ensuring comparable data across countries in terms of terminal provision.

4.60 Despite the (slightly) higher log-likelihood of model (iii), we analyse the efficiency estimated through model (iv) to ensure consistency with the en-route provision assessment.

5. Saving potentials

Data Envelopment Analysis

5.01 The efficiency results of the DEA model for en-route air traffic control provision are presented in Table 5-1.

		All years	**1st period**	**2nd period**	**Latest 3 years**
ANSP	**Country**	**2006-2016**	**2006-2010**	**2011-2016**	**2014-2016**
Austro Control	Austria	0.73	0.79	0.68	0.67
Belgocontrol	Belgium	0.52	0.53	0.52	0.53
BULATSA	Bulgaria	0.36	0.27	0.45	0.50
Croatia Control	Croatia	0.58	0.55	0.60	0.59
DCAC Cyprus	Cyprus	0.87	0.75	0.97	1.00
ANS CR	Czech Republic	0.57	0.56	0.58	0.55
NAVIAIR	Denmark	0.92	0.93	0.91	0.90
EANS	Estonia	0.92	0.97	0.88	0.84
Finavia	Finland	0.96	1.00	0.92	0.94
DSNA	France	0.86	0.88	0.84	0.82
DFS	Germany	0.95	1.00	0.90	0.86
HCAA	Greece	0.88	0.75	0.98	0.99
HungaroControl	Hungary	0.59	0.68	0.51	0.53
MUAC	International	1.00	1.00	1.00	1.00
IAA	Ireland	0.85	0.83	0.86	0.96
ENAV	Italy	0.67	0.68	0.65	0.64
LGS	Latvia	0.78	0.63	0.91	0.82
Oro Navigacija	Lithuania	0.40	0.34	0.45	0.45
MATS	Malta	0.82	0.65	0.96	1.00
LVNL	Netherlands	0.57	0.61	0.54	0.53
Avinor (Continental)	Norway	1.00	0.99	1.00	1.00
PANSA	Poland	0.51	0.51	0.51	0.47
NAV Portugal (Continental)	Portugal	0.56	0.45	0.65	0.73
ROMATSA	Romania	0.34	0.32	0.36	0.37
LPS	Slovak Republic	0.41	0.40	0.42	0.43
Slovenia Control	Slovenia	0.66	0.64	0.67	0.65
ENAIRE	Spain	0.53	0.45	0.60	0.64
LFV	Sweden	0.90	0.85	0.94	1.00
Skyguide	Switzerland	0.99	0.98	1.00	1.00
NATS (Continental)	UK	0.93	0.94	0.92	0.99

Table 5-1: DEA Efficiency estimates for en-route provision

5.02 Overall, the results of the DEA suggest that there is an average level of inefficiency in the region of 25% given the current market organization.

5.03 According to the average latest three year assessment (last column), MUAC, Switzerland, Cyprus, Norway, Malta and Sweden are relatively efficient together with Greece, Switzerland and the UK. It should be noted that at least one ANSP must receive a score of one in each year due to the nature of the DEA model. Furthermore, relatively specialised ANSPs, for example with high seasonality, are likely to be deemed relatively efficient and receive an efficiency score of one. The least efficient include Romania, the Slovak Republic, Lithuania and Poland.

5.04 Comparing the average efficiency estimates from 2006 to 2010 (period 1 prior to the price capping regulation) and from 2011-2016 (period 2), the majority of ANSPs show an efficiency improvement. Notable improvements include Bulgaria, Cyprus, Latvia, Malta, Portugal and Spain. On the other hand, Hungary, Austria, the Netherlands and Germany show declines between the two periods on average.

5.05 The results for the DEA model applied to the terminal activities are presented in Table 5-2. The terminal efficiency results do not include Sweden (data only available to 2007), MUAC (only provides upper airspace air traffic control provision) and the earliest years for Greece (2006-2009).

		All years	1st period	2nd period	Latest 3 years
ANSP	**Country**	**2006-2016**	**2006-2010**	**2011-2016**	**2014-2016**
Austro Control	Austria	0.86	0.85	0.87	0.85
Belgocontrol	Belgium	0.79	0.87	0.72	0.71
BULATSA	Bulgaria	0.26	0.24	0.27	0.31
Croatia Control	Croatia	0.57	0.69	0.47	0.50
DCAC Cyprus	Cyprus	0.62	0.50	0.72	0.82
ANS CR	Czech Republic	0.35	0.31	0.39	0.38
NAVIAIR	Denmark	0.87	0.89	0.86	0.92
EANS	Estonia	0.78	0.55	0.97	0.98
Finavia	Finland	0.53	0.46	0.59	0.64
DSNA	France	0.63	0.60	0.66	0.64
DFS	Germany	0.97	0.98	0.95	0.90
HCAA	Greece	0.61	0.46	0.74	1.00
HungaroControl	Hungary	0.27	0.28	0.26	0.28
IAA	Ireland	0.73	0.81	0.66	0.77
ENAV	Italy	0.66	0.70	0.62	0.61
LGS	Latvia	0.43	0.29	0.56	0.53
Oro Navigacija	Lithuania	0.36	0.37	0.35	0.37
MATS	Malta	0.51	0.36	0.64	0.68
LVNL	Netherlands	0.78	0.59	0.94	0.93
Avinor (Continental)	Norway	0.52	0.50	0.55	0.60
PANSA	Poland	0.49	0.49	0.48	0.47
NAV Portugal (Continental)	Portugal	0.47	0.42	0.51	0.62
ROMATSA	Romania	0.22	0.26	0.19	0.20
LPS	Slovak Republic	0.25	0.27	0.24	0.23
Slovenia Control	Slovenia	0.66	0.68	0.64	0.61
ENAIRE	Spain	0.38	0.29	0.45	0.48
Skyguide	Switzerland	0.70	0.66	0.74	0.67
NATS (Continental)	UK	0.93	0.87	0.97	1.00

Table 5-2: DEA Efficiency estimates for terminal provision

5.06 Overall, the results of the DEA suggest that there is an average level of inefficiency in the region of 40% given the current market organization.

5.07 According to the latest three year average, the UK and Greece are deemed relatively efficient, together with Estonia and the Netherlands. The least efficient include Romania, the Slovak Republic, Hungary and Bulgaria.

5.08 Comparing the average efficiency estimates between periods 1 and 2, it would appear that there are substantial changes over time. The majority of ANSPs show an average efficiency improvement. The largest positive changes have been made by Latvia, Malta, Estonia, Greece, the Netherlands, Spain and Cyprus. On the other hand, Croatia, Romania, Ireland, Belgium and Italy show greater cost inefficiencies in terminal provision over time.

Stochastic Frontier Analysis

5.09 The efficiency results of the SFA models (ii) for en-route air traffic control provision are presented in Table 5-3.

		All years	**1st period**	**2nd period**	**Latest 3 years**
ANSP	**Country**	**2006-2016**	**2006-2010**	**2011-2016**	**2014-2016**
Austro Control	Austria	0.89	0.90	0.88	0.89
Belgocontrol	Belgium	0.88	0.90	0.87	0.87
BULATSA	Bulgaria	0.63	0.52	0.73	0.81
Croatia Control	Croatia	0.83	0.78	0.87	0.86
DCAC Cyprus	Cyprus	0.50	0.40	0.58	0.59
ANS CR	Czech Republic	0.74	0.70	0.77	0.80
NAVIAIR	Denmark	0.60	0.65	0.55	0.53
EANS	Estonia	0.74	0.78	0.70	0.70
Finavia	Finland	0.56	0.60	0.53	0.51
DSNA	France	0.84	0.84	0.84	0.80
DFS	Germany	0.92	0.94	0.91	0.89
HCAA	Greece	0.65	0.57	0.67	0.72
HungaroControl	Hungary	0.65	0.68	0.63	0.66
MUAC	International	0.94	0.94	0.95	0.94
IAA	Ireland	0.54	0.54	0.54	0.57
ENAV	Italy	0.66	0.67	0.66	0.62
LGS	Latvia	0.56	0.52	0.60	0.56
Oro Navigacija	Lithuania	0.48	0.43	0.52	0.53
MATS	Malta	0.44	0.40	0.47	0.46
LVNL	Netherlands	0.85	0.87	0.84	0.86
Avinor (Continental)	Norway	0.70	0.64	0.75	0.70
PANSA	Poland	0.71	0.69	0.72	0.72
NAV Portugal (Continental)	Portugal	0.59	0.59	0.60	0.62
ROMATSA	Romania	0.47	0.40	0.53	0.57
LPS	Slovak Republic	0.77	0.68	0.85	0.90
Slovenia Control	Slovenia	0.80	0.73	0.86	0.87
ENAIRE	Spain	0.58	0.51	0.64	0.63
LFV	Sweden	0.56	0.56	0.57	0.58
Skyguide	Switzerland	0.89	0.87	0.90	0.90
NATS (Continental)	UK	0.88	0.90	0.86	0.85

Table 5-3: SFA Efficiency estimates for en-route provision

5.10 Overall, the results of the SFA suggest that there is an average level of inefficiency in the region of 30% given the current market organization.

5.11 According to the average efficiency estimates of the most recent three years, MUAC (international), Switzerland, the Slovak Republic, Germany and Austria are deemed relatively more efficient. The least efficient include Malta, Finland, Denmark and Lithuania.

5.12 Comparing the average efficiency estimates across the two periods, it would appear that the majority of ANSPs either remained consistent or improved. In particular, Cyprus, Bulgaria, Romania, Spain, and the Slovak republic have achieved improvements of 25% and above. However, Denmark, Estonia, and Finland all appear to have reduced in efficiency between the two periods by 10% and above.

5.13 The results for model (iv) applied to the terminal activities are presented in Table 5-4. The terminal efficiency results do not include Sweden (data only available to 2007), MUAC (only provides upper airspace air traffic control provision) and the earliest years for Greece (2006-2009).

		All years	1st period	2nd period	Latest 3 years
ANSP	**Country**	**2006-2016**	**2006-2010**	**2011-2016**	**2014-2016**
Austro Control	Austria	0.81	0.82	0.81	0.81
Belgocontrol	Belgium	0.82	0.83	0.82	0.81
BULATSA	Bulgaria	0.68	0.68	0.67	0.68
Croatia Control	Croatia	0.73	0.73	0.73	0.74
DCAC Cyprus	Cyprus	0.73	0.70	0.75	0.76
ANS CR	Czech Republic	0.74	0.75	0.72	0.72
NAVIAIR	Denmark	0.78	0.77	0.79	0.78
EANS	Estonia	0.70	0.75	0.65	0.62
Finavia	Finland	0.65	0.66	0.65	0.64
DSNA	France	0.72	0.71	0.73	0.73
DFS	Germany	0.87	0.88	0.87	0.86
HCAA	Greece	0.77	0.66	0.79	0.88
HungaroControl	Hungary	0.72	0.71	0.73	0.74
IAA	Ireland	0.76	0.76	0.75	0.77
ENAV	Italy	0.75	0.75	0.74	0.73
LGS	Latvia	0.72	0.62	0.80	0.79
Oro Navigacija	Lithuania	0.60	0.60	0.60	0.59
MATS	Malta	0.68	0.63	0.72	0.73
LVNL	Netherlands	0.82	0.82	0.82	0.80
Avinor (Continental)	Norway	0.69	0.68	0.70	0.69
PANSA	Poland	0.67	0.66	0.68	0.66
NAV Portugal (Continental)	Portugal	0.69	0.69	0.70	0.72
ROMATSA	Romania	0.60	0.58	0.61	0.60
LPS	Slovak Republic	0.71	0.72	0.71	0.72
Slovenia Control	Slovenia	0.86	0.86	0.86	0.86
ENAIRE	Spain	0.72	0.66	0.77	0.78
Skyguide	Switzerland	0.85	0.86	0.85	0.84
NATS (Continental)	UK	0.81	0.81	0.82	0.82

Table 5-4: SFA Efficiency estimates for terminal provision

5.14 Overall, the results of the SFA suggest that there is an average level of inefficiency in the region of 27% given the current market organization.

5.15 Based on the latest three year period (2014 to 2016 inclusive), Germany, Slovenia, Greece and Switzerland are relatively more efficient. The least efficient include Lithuania, Cyprus, and Romania are located at the bottom of the ranks.

5.16 Comparing the average efficiency estimates across the two periods, it would appear that the majority of ANSPs remained consistent. However, major improvements are estimated for Cyprus, Latvia, Greece and Spain. On the opposite side, Estonia has the greatest reduction in efficiency.

Combining the results

5.17 The suggested savings are based on the average cost efficiencies for the latest three years, 2014 to 2016, and are computed as follows:

$$1 - Efficiency\ score\ =\ Potential\ saving$$

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

1) Potential savings as the maximum value resulting from the DEA and SFA models

2) Potential savings as the minimum value resulting from the DEA and SFA estimates

3) Potential savings as the average of the two sets of results

5.19 Table 5-5 presents the potential savings related to the en-route provision. Values can be read as percentage reductions in costs necessary to achieve efficient production levels.

ANSP	Country	Maximum potential saving	Minimum potential saving	Average potential saving
Austro Control	Austria	0.33	0.11	0.22
Belgocontrol	Belgium	0.47	0.12	0.29
BULATSA	Bulgaria	0.50	0.20	0.35
Croatia Control	Croatia	0.41	0.14	0.27
DCAC Cyprus	Cyprus	0.42	0.00	0.21
ANS CR	Czech Republic	0.45	0.18	0.32
NAVIAIR	Denmark	0.47	0.10	0.29
EANS	Estonia	0.32	0.16	0.24
Finavia	Finland	0.51	0.06	0.29
DSNA	France	0.23	0.18	0.21
DFS	Germany	0.14	0.13	0.13
HCAA	Greece	0.23	0.01	0.12
HungaroControl	Hungary	0.47	0.36	0.41
MUAC	International	0.06	0.00	0.03
IAA	Ireland	0.43	0.04	0.23
ENAV	Italy	0.40	0.36	0.38
LGS	Latvia	0.44	0.18	0.31
Oro Navigacija	Lithuania	0.55	0.48	0.51
MATS	Malta	0.55	0.00	0.28
LVNL	Netherlands	0.47	0.11	0.29
Avinor (Continental)	Norway	0.32	0.00	0.16
PANSA	Poland	0.53	0.27	0.40
NAV Portugal (Continental)	Portugal	0.38	0.27	0.33
ROMATSA	Romania	0.63	0.45	0.54
LPS	Slovak Republic	0.57	0.11	0.34
Slovenia Control	Slovenia	0.35	0.13	0.24
ENAIRE	Spain	0.37	0.36	0.37
LFV	Sweden	0.42	0.00	0.21
Skyguide	Switzerland	0.09	0.00	0.05
NATS (Continental)	UK	0.16	0.01	0.08

Table 5-5: Potential savings for en-route provision (%)

5.20 Table 5-6 presents the potential savings in monetary values (million Euros) without PPP.

ANSP	Country	Millions euros		
		Maximum potential savings	Minimum potential savings	Average potential savings
Austro Control	Austria	51.6	17.5	34.5
Belgocontrol	Belgium	44.3	11.2	27.7
BULATSA	Bulgaria	41.0	15.8	28.4
Croatia Control	Croatia	30.7	10.4	20.6
DCAC Cyprus	Cyprus	14.6	0.0	7.3
ANS CR	Czech Republic	45.7	18.6	32.2
NAVIAIR	Denmark	38.3	8.4	23.3
EANS	Estonia	5.2	2.6	3.9
Finavia	Finland	20.0	2.6	11.3
DSNA	France	231.5	186.2	208.8
DFS	Germany	111.3	103.4	107.4
HCAA	Greece	28.2	1.4	14.8
HungaroControl	Hungary	38.3	29.8	34.1
MUAC	International	8.8	0.0	4.4
IAA	Ireland	38.3	3.5	20.9
ENAV	Italy	213.5	195.2	204.3
LGS	Latvia	7.6	3.2	5.4
Oro Navigacija	Lithuania	11.4	9.8	10.6
MATS	Malta	8.2	0.0	4.1
LVNL	Netherlands	58.7	14.1	36.4
Avinor (Continental)	Norway	24.2	0.0	12.1
PANSA	Poland	70.9	36.0	53.5
NAV Portugal (Continental)	Portugal	36.0	25.6	30.8
ROMATSA	Romania	84.0	60.6	72.3
LPS	Slovak Republic	30.2	5.9	18.0
Slovenia Control	Slovenia	10.1	3.7	6.9
ENAIRE	Spain	215.5	210.0	212.7
LFV	Sweden	64.0	0.0	32.0
Skyguide	Switzerland	17.6	0.2	8.9
NATS (Continental)	UK	98.0	7.6	52.8
Total		1,697.7	983.3	1,340.4

Table 5-6: Potential savings for en-route provision (million euros)

5.21 The estimated total, annual, potential savings for en-route provision for the market vary from one billion euros (minimum potential) equivalent to an average of around 17% reduction in costs to 1.7 billion euros (maximum potential) equivalent to around 30% savings.

5.22 Table 5-7 presents the potential savings related to the terminal provision. Values can be read as percentage reductions in costs necessary to achieve efficient production levels.

ANSP	Country	Maximum potential savings	Minimum potential savings	Average potential savings
Austro Control	Austria	0.19	0.15	0.17
Belgocontrol	Belgium	0.29	0.19	0.24
BULATSA	Bulgaria	0.69	0.32	0.51
Croatia Control	Croatia	0.50	0.26	0.38
DCAC Cyprus	Cyprus	0.24	0.18	0.21
ANS CR	Czech Republic	0.62	0.28	0.45
NAVIAIR	Denmark	0.22	0.08	0.15
EANS	Estonia	0.38	0.02	0.20
Finavia	Finland	0.36	0.36	0.36
DSNA	France	0.36	0.27	0.31
DFS	Germany	0.14	0.10	0.12
HCAA	Greece	0.12	0.00	0.06
HungaroControl	Hungary	0.72	0.26	0.49
IAA	Ireland	0.23	0.23	0.23
ENAV	Italy	0.39	0.27	0.33
LGS	Latvia	0.47	0.21	0.34
Oro Navigacija	Lithuania	0.63	0.41	0.52
MATS	Malta	0.32	0.27	0.29
LVNL	Netherlands	0.20	0.07	0.13
Avinor (Continental)	Norway	0.40	0.31	0.35
PANSA	Poland	0.53	0.34	0.43
NAV Portugal (Continental)	Portugal	0.38	0.28	0.33
ROMATSA	Romania	0.80	0.40	0.60
LPS	Slovak Republic	0.77	0.28	0.53
Slovenia Control	Slovenia	0.39	0.14	0.27
ENAIRE	Spain	0.52	0.22	0.37
Skyguide	Switzerland	0.33	0.16	0.24
NATS (Continental)	UK	0.18	0.00	0.09

Table 5-7: Potential savings for terminal provision (%)

5.23 Table 5-8 presents the potential savings in monetary values (million Euros) without PPP.

ANSP	Country	Millions euros		
		Maximum potential saving	Minimum potential saving	Average potential saving
Austro Control	Austria	6.72	5.47	6.09
Belgocontrol	Belgium	15.45	9.91	12.68
BULATSA	Bulgaria	5.84	2.73	4.28
Croatia Control	Croatia	4.71	2.43	3.57
DCAC Cyprus	Cyprus	0.93	0.70	0.82
ANS CR	Czech Republic	14.03	6.26	10.14
NAVIAIR	Denmark	6.61	2.38	4.50
EANS	Estonia	0.67	0.04	0.35
Finavia	Finland	8.74	8.67	8.71
DSNA	France	86.43	65.80	76.12
DFS	Germany	32.22	21.50	26.86
HCAA	Greece	2.67	0.00	1.34
HungaroControl	Hungary	11.05	3.99	7.52
IAA	Ireland	4.84	4.77	4.80
ENAV	Italy	55.49	38.21	46.85
LGS	Latvia	2.39	1.04	1.71
Oro Navigacija	Lithuania	3.18	2.03	2.61
MATS	Malta	1.00	0.84	0.92
LVNL	Netherlands	11.08	3.79	7.44
Avinor (Continental)	Norway	34.13	26.12	30.13
PANSA	Poland	13.73	8.83	11.28
NAV Portugal (Continental)	Portugal	9.04	6.75	7.89
ROMATSA	Romania	22.75	11.19	16.97
LPS	Slovak Republic	5.16	1.89	3.52
Slovenia Control	Slovenia	1.40	0.51	0.96
ENAIRE	Spain	88.01	38.00	63.00
Skyguide	Switzerland	37.40	18.38	27.89
NATS (Continental)	UK	29.59	0.00	14.79
Total		515.26	292.23	403.74

Table 5-8: Potential savings for terminal provision (million euros)

5.24 The estimated annual, system total, potential savings for terminal provision vary from 290 million euros (minimum potential) equivalent to around 18% of the 2016 terminal provision costs to 515 million euros (maximum potential) equivalent to 33% savings.

Determined Unit Cost Forecast

5.25 We forecast the determined unit cost for the third reference period based on the model total cost estimates. In order to estimate costs for 2020 to 2024, we have relied on the base case scenario output estimates recently published by STATFOR[5]. The seven year forecast estimates flight movements and service units for all the countries analysed in this report except for MUAC. However, the en-route analyses undertaken in this report are based on flight hours controlled. Consequently, we calculated the average time per flight per ANSP in 2016 and assume that this remains constant over the timeframe. Subsequently, we were able to adapt the flight movement estimates to flight hours controlled for reference period 3 (RP3). Furthermore, we assume that all other variables remain constant, including the price of labour, price of capital, maximum number of sectors, complexity and seasonality.

5.26 After estimating the efficient total costs per ANSP for RP3 according to the two models, we then divided the values by the service unit forecasts in order to create Tables 5-9 and 5-10. The tables specify the projected, best practice, determined unit costs for each ANSP in 2016 standardized PPP euros.

[5] https://www.eurocontrol.int/sites/default/files/content/documents/official-documents/forecasts/seven-year-flights-service-units-forecast-2017-2023-Feb2017.pdf

Country	2020	2021	2022	2023	2024
Austria	32.99	32.99	32.99	33.00	33.01
Belgium	15.49	15.37	15.26	15.15	15.03
Bulgaria	23.95	23.52	23.09	22.66	22.23
Croatia	42.73	42.52	42.31	42.13	41.94
Cyprus	26.85	27.52	28.16	28.84	26.09
Czech Republic	29.55	29.20	28.87	28.59	28.34
Denmark	35.06	34.76	34.48	34.24	34.00
Estonia	21.02	20.91	20.83	20.75	20.66
Finland	33.01	32.73	32.45	32.25	32.03
France	36.88	36.69	36.49	36.32	36.20
Germany	40.71	40.15	39.56	39.10	38.61
Greece	30.57	30.51	30.47	30.44	30.40
Hungary	26.04	25.68	25.33	25.00	24.69
Ireland	18.07	17.86	17.73	17.60	17.45
Italy	39.67	39.51	39.33	39.18	39.02
Latvia	23.83	23.80	23.80	23.82	23.84
Lithuania	27.43	27.40	27.26	27.14	27.07
Malta	20.76	20.17	19.65	19.10	18.55
Netherlands	18.79	19.40	19.96	20.58	21.24
Norway	39.16	39.14	39.12	39.05	38.91
Poland	24.68	24.05	23.53	23.15	22.83
Portugal	25.76	25.82	25.91	25.93	25.92
Romania	21.83	21.70	21.57	21.45	21.33
Slovak Republic	27.87	27.59	27.30	27.01	26.73
Slovenia	39.35	39.16	38.93	38.81	38.63
Spain	41.29	40.91	40.54	40.15	39.77
Sweden	33.34	33.16	33.01	32.89	32.79
Switzerland	83.72	83.72	83.57	83.47	83.32
UK	46.75	46.40	46.13	45.88	45.64

Table 5-9: Determined Unit Cost Rate per ANSP for RP3 using DEA

Country	2020	2021	2022	2023	2024
Austria	41.69	41.11	40.53	40.03	39.51
Belgium	26.86	26.43	26.00	25.64	25.28
Bulgaria	38.08	36.92	35.81	34.74	33.68
Croatia	61.49	60.24	59.02	57.91	56.77
Cyprus	13.14	13.23	13.31	13.42	11.95
Czech Republic	39.32	38.26	37.31	36.52	35.77
Denmark	17.47	17.21	16.93	16.69	16.44
Estonia	18.18	17.86	17.49	17.16	16.82
Finland	17.72	17.50	17.25	17.05	16.82
France	43.69	43.41	43.13	42.90	42.71
Germany	52.20	51.83	51.43	51.12	50.80
Greece	23.99	23.73	23.50	23.27	23.04
Hungary	24.79	24.11	23.44	22.84	22.24
Ireland	10.63	10.47	10.33	10.21	10.07
Italy	39.77	39.58	39.38	39.21	39.03
Latvia	15.33	15.20	15.04	14.91	14.78
Lithuania	33.99	33.52	32.96	32.47	32.00
Malta	8.83	8.52	8.22	7.93	7.63
Netherlands	27.82	27.65	27.38	27.16	26.93
Norway	17.54	17.49	17.42	17.32	17.20
Poland	39.26	38.22	37.29	36.57	35.91
Portugal	18.48	18.29	18.14	17.98	17.80
Romania	24.90	24.33	23.78	23.28	22.78
Slovak Republic	58.68	56.78	54.87	53.15	51.50
Slovenia	61.56	60.23	58.88	57.76	56.53
Spain	44.98	44.60	44.25	43.85	43.46
Sweden	18.72	18.54	18.37	18.23	18.10
Switzerland	70.96	70.32	69.58	68.99	68.33
UK	41.56	41.19	40.90	40.61	40.33

Table 5-10: Determined Unit Cost Rate per ANSP for RP3 using SFA

5.27 Tables 5-9 and 5-10 show the large differences in costs across the multiple countries with the efficient average determined unit cost dropping from 33 standardized euros to just under 31 over the five year timeframe despite the estimated increases in traffic.

5.28 In light of the fact that DEA tends to favour specialists and SFA tends to reward the average, the results are different. We combine the determined unit costs using the benefit of the doubt approach (maximum cost which favours the ANSP over the consumer) leading to the results in Table 5-11.

Country	2020	2021	2022	2023	2024
Austria	41.69	41.11	40.53	40.03	39.51
Belgium	26.86	26.43	26.00	25.64	25.28
Bulgaria	38.08	36.92	35.81	34.74	33.68
Croatia	61.49	60.24	59.02	57.91	56.77
Cyprus	26.85	27.52	28.16	28.84	26.09
Czech Republic	39.32	38.26	37.31	36.52	35.77
Denmark	35.06	34.76	34.48	34.24	34.00
Estonia	21.02	20.91	20.83	20.75	20.66
Finland	33.01	32.73	32.45	32.25	32.03
France	43.69	43.41	43.13	42.90	42.71
Germany	52.20	51.83	51.43	51.12	50.80
Greece	30.57	30.51	30.47	30.44	30.40
Hungary	26.04	25.68	25.33	25.00	24.69
Ireland	18.07	17.86	17.73	17.60	17.45
Italy	39.77	39.58	39.38	39.21	39.03
Latvia	23.83	23.80	23.80	23.82	23.84
Lithuania	33.99	33.52	32.96	32.47	32.00
Malta	20.76	20.17	19.65	19.10	18.55
Netherlands	27.82	27.65	27.38	27.16	26.93
Norway	39.16	39.14	39.12	39.05	38.91
Poland	39.26	38.22	37.29	36.57	35.91
Portugal	25.76	25.82	25.91	25.93	25.92
Romania	24.90	24.33	23.78	23.28	22.78
Slovak Republic	58.68	56.78	54.87	53.15	51.50
Slovenia	61.56	60.23	58.88	57.76	56.53
Spain	44.98	44.60	44.25	43.85	43.46
Sweden	33.34	33.16	33.01	32.89	32.79
Switzerland	83.72	83.72	83.57	83.47	83.32
UK	46.75	46.40	46.13	45.88	45.64
Average max DUC	37.87	37.42	36.99	36.61	36.10

Table 5-11: Highest Efficient Determined Unit Cost Rate per ANSP for RP3

5.29 In summation, the results suggest that the determined unit cost rate for the beginning of RP3, were the ANSPs to provide their services efficiently, could be in the range of 38 standardized 2016 euros and that this should drop by 5% to 36 euros in 2024.

6. Final remarks

Regulatory Benchmarking

6.01 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.

6.02 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.

6.03 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 cost in a competitive setting. Hence, instead of competing on the market, we create pseudo competition via benchmarking based regulation, where the ANSPs compete via a model.

6.04 We note that this issue is particularly relevant in en-route provision given the clear monopolistic status of the ANSPs. For terminal provision, it would appear that there is an alternative in which an auctioning of the air traffic control service is possible and could replace the need for economic regulation.

6.05 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.

A first attempt

6.06 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.

6.07 It is important to understand that the results presented in this report represent the first set of benchmarking based cost targets. The underlying benchmarking models have been developed in a short timeframe based on pre-prepared data covering the years 2006 to 2016 inclusive. We note the time lag of a year and a half before data is published creates a more complicated regulatory process. Furthermore, we have had very little time to properly assess the dataset.

6.08 In these respects, the exercise undertaken deviates from more established regulatory benchmarking that often stretches over a longer timeframe normally one to two years, and is based on data collection that is tailored to the planned benchmarking exercise.

6.09 This explains some of the reasons for the different sets of results we obtained using DEA and SFA, and it provides a background for the proper application of our results.

Methodological differences

6.10 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 while SFA uses a combination of noise and inefficiency to explain the deviations.

6.11 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.

6.12 Finally, we note that the SFA model leads to a cost function which penalizes ANSPs that behave somewhat differently to the average. DEA, based on an envelopment frontier, is likely to estimate higher efficiency levels for those that specialize. In this context, we have applied a benefit of the doubt approach in the final results, which is to the benefit of the producer rather than the consumer, namely the airlines and passengers.

Application of our results

6.13 Given the uncertainty we experience in the establishment of good cost targets, we have applied the results cautiously leading to conservative savings estimates.

6.14 For the ANSPs that are consistently shown to be rather inefficient in all models, it makes sense to set more challenging cost saving targets than for those where the uncertainty is large.

6.15 A best-off approach to the combination of results may in this context be most appropriate. Consequently, the determined unit costs are based on this approach alone.

Future developments

6.16 We also suggest that the results should be improved over time. There are many ways to do so, including a further investigation of the cost standardization and the inclusion of other possible cost drivers such as quality of the services provided.

6.17 The cost measures need a closer investigation. We have undertaken a series of sensitivity tests, including PPPing only parts of the cost base to reflect that some capital costs may be bought on an international market. Still, further investigations would be worthwhile.

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

Quality

6.19 Another aspect that may be important but that is largely ignored here is the quality of service provided. Multiple key performance indicators have been created and some could be included in the cost analysis. For example, delays are collected and categorised according to cause, which could be included within the analysis. In other words, both capacities and flight efficiency indicators could be included in the cost based analyses.

6.20 Certain quality measures are difficult to include since they are affected by many random conditions outside the control of the ANSP but some are likely more directly affected by the choice of technology employed, which could therefore be included in a next step.

Fixed market structure

6.21 Our analysis presumes that the number of ANSPs are fixed and that the deviation of air space between them remains unaltered.

6.22 We hereby do not measure the possible gains or cost savings from consolidation of the Single European Sky.

European best practices

6.23 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.

6.24 Reports, such as those produced by the FAA and Eurocontrol[6], seem to indicate that the US system is substantially more efficient than that of Europe. In effect, an analysis looking for possible comparators outside of the EU could lead to much higher savings potential.

6.25 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.

Regulation or competition

6.26 It might also be interesting to investigate the possibilities of introducing competition for the market rather than regulating prices. In terminal provision, this possibility exists in Sweden, the UK and Germany and is being introduced in Spain. We do not have Swedish data but the German and UK efficiency estimates are among the higher in the terminal analysis.

[6] https://www.faa.gov/air_traffic/publications/media/us_eu_comparison_2015.pdf. Accessed on 23/5/2018.

Relevant literature

Adler, N., Kivel, A., Delhaye, E. and Proost, S., 2017. Ownership Form & Air Navigation Service Provider Performance. Seventh SESAR Innovation Days, https://www.sesarju.eu/sites/default/files/documents/sid/2017/SIDs_2017_paper_35.pdf.

Agrell, P.J., Bogetoft, P. 2000. Ekonomisk na ̇tbesiktning. final report stem. Tech. rep., SUMICSID AB. (In Swedish)

Aigner, D.J., Chu, S.F. 1968. On Estimating the Industry Production Function. American Economic Review, 58, 826–839.

Aigner, D.J., Lovell, C.A.K, Schmidt, P. 1977. Formulation and Estimation of Stochastic Frontier Production Function Models. Journal of Econometrics, 6, 21–37.

Banker, R.D. 1996. Hypothesis tests using data envelopment analysis. Journal of Productivity Analysis, 7, 139-159.

Battes, G.E., Coelli T.J 1992. Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. Journal of Productivity Analysis, 3, 153-169.

Battes, G.E., Coelli T.J 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empirical Economics, 20, 325-332.

Bogetoft, P., 2012. Performance Benchmarking Measuring and Managing Performance. Springer, New York.

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

CEG 2011. Econometric Cost-Efficiency Benchmarking of Air Navigation Service Providers. Competition Economists LLP for Eurocontrol.

Charnes, A., Cooper, W.W., Rhodes, E. 1978. Measuring the efficiency of decision making units. European Journal of Operational Research, 2, 429–444.

Coelli, T., Prasada Rao, D.S., Battese, G. 1998. An Introduction to Efficiency and Productivity Anal- ysis. Kluwer Academic Publishers.

Coelli,T., Rao, P., O'Donnell, C., Battese, G., 2005. An introduction to efficiency and productivity analysis, 2nd ed. Springer.

COMPAIR 2018. Competition for air traffic management. Deliverable 3 : http://www.compair-project.eu/files/2017_05_27_d3_2_report_on_economic_analysis_compair_00_02_00.pdf. SESAR Horizon 2020 research and innovation programme.

COMPAIR 2018. Competition for air traffic management. Deliverable 5 : http://www.compair-project.eu/files/2018_02_01_er_final_project_results_compair_01_00_00.pdf. SESAR Horizon 2020 research and innovation programme.

Deprins, D., Simar, L., Tulkens, H. 1984. Measuring labour efficiency in post offices. Tech. rep., In M. Marchand, P. Pestieau, and H. Tulkens, (eds.) "The Performance of Public Enterprises: Concepts and Measurements", North Holland, 243-267.

Fethi, M., Jackson, P.M., Weyman-Jones, T.G. 2001. European airlines: a stochastic dea study of effi- ciency with market liberalisation. Tech. rep., University of Leicester Efficiency and Productivity Research Unit.

Greene, W.H., 1990. A Gamma-Distributed Stochastic Frontier Model. Journal of Econometrics, 46,141–164.

Greene, W.H., 2008. Econometric Analysis, sixth edn. Pearson Prentice Hall

Hattori, T., 2002. Relative performance of US and Japanese electricity distribution: an application of stochastic frontier analysis. Journal of Productivity Analysis, 18(3), 269-284.

Land, K.C., Lovel, C.A.K., Thore, S. 1993. Chance-constrained data envelopment analysis. Managerial and Decision Economics, 14, 541–554.

Lovell, C.A.K. 1993. Production Frontiers and Productive Efficiency. In: Fried H, Lovell CAK, Schmidt S (eds) The Measurement of Productive Efficiency: Techniques and Applications, Ox- ford University Press, New York.

Joint Research Center (JRC), 2008. Moreno, D.G., Cost of capital indicator for EU member states – methodology, JRC Scientific and technical reports.

Olesen, O., Petersen, N.C. 1995. Chance constrained efficiency evaluation. Management Science, 41, 442–457.