SPARA 2020 – Benchmarking of airports

Report, WPA7.4

Svein Bråthen, Molde University College, Norway

Rico Merkert, University of Sydney, Australia

Northern Periphery and Arctic Programme 2014–2020

EUROPEAN UNION
Investing in your future
European Regional Development Fund

Preface

The SPARA2020 project proposal has underlined that peripheral and remote airports face a range of very special challenges. These airports are usually loss making and subject to low traffic volumes. The cost of safe and regulatory compliant airports augment with practices and procedures, e.g. ICAO regulations, that they often have in common with major airports with very different sets of challenges. Peripheral and remote airports have the additional challenge of changeable and extreme weather, low temperatures and challenging terrain.

All remote and peripheral communities recognise that transport connectivity is critical for their wellbeing and future viability. Public funds and/or cross subvention from larger airports under common ownership are usually required to support and maintain airport infrastructure for smaller airports with thin routes, of which many are supported through a PSO scheme.

In order to assess the efficiency of the airports in the Northern Periphery, SPARA2020 Work Package A7.4 contains a benchmarking study. This report presents the main findings from this study. Data collection proved to be a challenge. Unfortunately, the number of respondents remained too low, despite up to 10 rounds of calls and reminders. The final sample allowed for indicative findings only. Thus, the study has limited value for transferring knowledge between pairs of 'learners' and 'peers' in the sample of airports.

The authors would like to thank the Norwegian airport operator, Avinor ASA, represented by Senior Advisor Eigil U. Olsen, for their help in providing detailed data for all their regional airports. We would also like to thank the other responding airports for their kind assistance. The authors are solely responsible for the content of the report, including any remaining mistakes.

This report represents joint work, although Rico Merkert has been responsible for running the PPM and DEA models.

Molde and Sydney, 2018

Svein Bråthen and Rico Merkert

Contents

1. Introduction

The main objective for a comprehensive benchmarking study is to reveal areas in which the Northern Periphery (NPP) airports can improve, by comparing themselves with best practice airports that enjoy a competitive advantage. This study reports the content of Work Package A7.4.

At the outset, the two main objectives were:

1. Identification of cost, technological and organizational leadership among NPP airports, and the characteristics of the best practice of airport management (what drives the efficiency and advanced leadership of top and under performers).
2. Dissemination of best practice from the leading peers to other comparable airports.

Partial Productivity Measures (PPM) and Data Envelopment Analysis (DEA) were applied in order to identify the efficiency of each airport within the NPP area. The ambition was to have a large sample of airports in order to assess airport efficiency comprehensively for the NPP airports. Unfortunately, this ambition was met only partially.

The second step in the DEA analysis (truncated regressions based on the DEA results) address in a bit more detail issues that can explain the differences in terms of relative efficiency. The variables in the final assessment included airport governance, degree of outsourcing, whether the airports serve offshore operations and accessibility with public transport.

Ideally, for interpretation of the results we also envisage to incorporate technical and operation aspects of airport management In a third step. For that matter a mapping process will be required. However, the results became too uncertain to carry this out. Such mapping was carried out by Avinor (who control most airports in Norway), comparing a number of smaller airports in Finland, Iceland, Norway and Sweden (Avinor As (2010), and we refer to this work if such a closer investigation of the airport operations are to be considered for individual airports.

2. Research approach

This section briefly presents the two main research methods that we find useful for SPARA2020. The methods are Partial Productivity Measures (PPM) and Data Envelopment Analysis (DEA). Both methods are well known from numerous benchmark studies within the airport industry as well as within other service industries. We refer to Appendix 3 for further elaboration.

2.1. The Partial Productivity Measures (PPM) approach and its output

Applying partial indicators of performance is the traditional and most commonly used method to compare airports. Typically, these studies focus on the following dimensions of airport performance:

- Cost efficiency
- Productivity
- Revenue generating capability
- Profitability

For each of these dimensions, measures have been developed which relate in some way the airport's inputs and outputs. The major inputs in an airport system are: labour and capital. Depending on the performance measure used, the inputs are measured in either physical or financial terms. For example, labour can be expressed in terms of number of employees or in terms of total labour costs incurred by the airport. Capital is usually measured in physical terms and can be represented by for example, the capacity of the runways or the amount of terminal space allocated to retail activities.

As far as output is concerned various measures can be used. Traffic represents the key output of an airport and there are typically three dimensions; passengers, freight or aircraft movements. For the majority of civil airports, the most important output is passenger traffic. However, some airports have substantial freight activity. The challenge for researchers in the partial performance field has been to devise robust and reliable measures of output that cover the different types of traffic. The Work Load Unit (WLU) was devised to solve this problem as it essentially combines passenger and freight volume into one aggregate measure of airport output. The use of WLU is particularly relevant for larger airports where freight volumes are substantial. Hence, WLU is not used in this study.

Given that data is accessible, partial measures are intuitively very easy to compute, understand and interpret. In this report, passengers and cargo per aircraft movement are used as operational PPMs whereas passengers/terminal m^2 and per full-time employee are used as productivity PPMs. We have also calculated financial and unit cost PPMs that are described in the analysis below.

One significant limitation with partial measures is that they are less effective in providing a robust assessment of an airport's overall performance especially within the context of measuring the performance of the airport in relation to its optimum potential performance. Techniques like DEA are able to do this. However, the advantage of the PPM approach lies in the relative simplicity of the calculations and analysis of the results. Furthermore, PPM allows the researcher to investigate and compare performance at a disaggregated level; for example, measuring the non-aeronautical performance of different airports.

2.2. The DEA approach and its output

Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) are appropriate methods for assessing airport efficiency. The advantages of SFA over other methods until recently, are that it builds on econometrics and therefore it is able to capture noise in the data more adequately as compared to other methods. DEA has recently been developed to account for noise in data and hence SFA is no longer advantageous over it. We will have our main focus on the use of DEA.

The main advantages of DEA over other methods and which are relevant for the assessment are as follows:

- DEA is easy to grasp and understand for managers; the benchmark is other service providers providing the same type of services using the same types of inputs and, these other providers are observable and not derived from some assumed production function.
- DEA readily incorporates multiple inputs and outputs and, it does not require price data to calculate technical efficiency. This makes it especially suitable for analysing the efficiency of service production, where it is often difficult to assign prices to many of the outputs.

- It determines sources of inefficiency and efficiency levels and provides a means of decomposing economic (cost) efficiency into technical and allocative efficiency. Furthermore, technical efficiency is decomposed into scale and non-scale effects.
- DEA identifies the “peers” for units (airports) that are not efficient. It thus provides a set of role models that the inefficient units can look to in order to improve its operations. This makes DEA a very useful tool for benchmarking compared to other methods, but it puts demands on the sample size and representativeness.
- DEA can be extended to study efficiency over time using the Malmquist productivity index (MPI). Thus its advantages over other methods are maintained even when efficiency is being studied over time. Here, the time span became too short for using MPI.

Like any assessment method, DEA is based on a number of assumptions and hence has some weaknesses that one needs to acknowledge. The main ones are follows:

- DEA is a deterministic rather than a statistical approach. Its results would therefore be sensitive to measurement errors. However, recently it has been proven that applying DEA together with bootstrapping takes account of statistical noise adequately.
- DEA scores are sensitive to the number of inputs and outputs, and the sample size. An Increase of the sample size will tend to reduce the average efficiency score because including more observations provides greater scope for DEA to find a comparison partner. Conversely, fewer observations relative to the number of inputs and outputs can inflate the efficiency scores. There are however ways of dealing with this problem.
- The method is also sensitive to outliers. We have paid close attention to this and to justify very carefully and with transparency if any airport should become omitted from the study for this reason.

The input factors used are:

- *Labour,*
  - Length of longest runways
  - Terminal size

The following output factors used are:

- *Passengers*
  - Cargo
  - Aircraft movements

In sum, the most beneficial aspect of DEA is that it combines several inputs and outputs to produce a single overall efficiency measure. These factors have proven to be obtainable from various airport operators.

As mentioned in the introduction, the DEA efficiency scores have been assessed by means of a second stage analysis. The DEA efficiency scores are regressed on external factors such as governance (whether the airport is member of a group or not), whether it has offshore operations (e.g. a heliport), the degree to which airport operations are outsourced to third parties and finally whether there are public transport access to the airport or not. Some aspects, like military activities, public or private ownership were considered but data representativeness did not allow for any further assessment.

2.3. The engineering-based approach

In Section 1 we referred to the benchmark study by Avinor (2010). The study is very detailed, and it advises some specific steps to improve the airport operations efficiency with respect to detailed technical and organizational aspects of airport operations. The ambitions in this benchmarking study was to test this approach as a third step, once a selection of 'learners' and 'peers' were identified. However, the characteristics of the sample did not allow this step to be taken. Hence, we only provide a brief description of the approach.

Ideally, DEA identifies the efficient 'peers' for units that are not efficient ('learner' airports). The engineering-based approach in Avinor (2010) appears as convenient for looking into the inefficient units and comparing them with their 'peers' in order to improve their operations. We still consider this approach to be a viable type of approach for individual airports that decides to carry out an efficiency improvement program. Detailed technical insight into airport operations is required. Examples of elements that need to be included in such an assessment are:

- Different organizational structures
- Details on staff multi-tasking
- Automation of operations, current state and potential for development
- Collaboration and coordination with the airlines
- Off-site development, like paid parking
- Emergency preparedness (like ambulance flights) and its consequences for the operations

Snap-shots of the current state should be made in the beginning of the project period, as basis for ex post evaluation. We refer to Avinor (2010) for details.

3. Data

The data were collected by means of the questionnaire that is presented in Appendix 1. The questionnaire was e-mailed to the respondents in MS Word format. In addition, an identical online questionnaire was made available to the respondents.

At the outset, we were looking for a balanced number of airports from Norway, Sweden, Finland, Iceland, Ireland and relevant parts of the UK. This proved to be difficult. We were left with a rich sample of Norwegian airports but a rather thin sample if none at all from the other countries.

Table 3.1 shows the airports in the resulting sample.

Table 3.1: Sample of 52 NPA airports

Norway	Rest of NPA
Ålesund Lufthavn	Vágar Airport (Faroe Islands, Denmark)
Alta Lufthavn	Ireland West Airport, Knock (Ireland)
Båtsfjord Lufthavn	Cork International Airport (Ireland)
Berlevåg Lufthavn	Donegal Airport (Ireland)
Brønnøysund Lufthavn	Gander International Airport (Newfoundland, Canada)
Florø Lufthavn	Coda Derry Airport (Northern Ireland)
Førde Lufthavn	Benbecula (Scotland)
Hammerfest Lufthavn	Wick John O' Groats Airport (Scotland)
Harstad Narvik Evenes Lufthavn	Stornoway Airport (Scotland)
Hasvik Lufthavn	Argyll & Bute Council (Oban Airport, Coll, Colonsay) (Scotland)
Haugesund Lufthavn	Dundee Airport (Scotland)
Honningsvåg Lufthavn	Åre Östersund Airport (Sweden)
Kirkenes Lufthavn	Höga Kusten Airport (Sweden)
Kristiansand Lufthavn	Norrköping Airport (Sweden)
Kristiansund Lufthavn	Jonkoping Airport (Sweden)
Lakselv Lufthavn	Sveg Airport (Sweden)
Leknes Lufthavn	Hemavan Tarnaby Airport (Sweden)
Mehamn Lufthavn	
Mo I Rana Lufthavn	
Molde Lufthavn	
Mosjøen Lufthavn	
Namsos Lufthavn	
Ørsta Volda Lufthavn	
Røros Lufthavn	
Rørvik Lufthavn	
Røst Lufthavn	
Sandane Lufthavn	
Sandnessjøen Lufthavn	
Sogndal Lufthavn	
Sørkjosen Lufthavn	
Stokmarknes Lufthavn	
Svalbard Lufthavn	
Svolvær Lufthavn	
Vadsø Lufthavn	
Vardø Lufthavn	

Norway is represented with 35 airports. Sweden has 6, Scotland 5, Ireland 3, Canada 1, Northern Ireland 1 and the Faroes 1 airport, respectively. Canada was included because the remoter regions' airports have operations quite similar to the ones in the NPP area.

The thin data for most countries makes cross-country comparisons difficult, and we will have to resort to a rather limited analysis where the numbers (apart from Norway) are not representative for the countries within the NPP area.

4. Results and discussion

4.1. Partial productivity measures

A few commonly used PPM measures are presented. Again, a comparison between countries gives no support for conclusions to de drawn because of sample bias.

Table 4.1: Selected operational Partial Productivity Measures (PPM)

	pax/ATM 2015	pax/ATM 2016	Cargo (kg) /ATM 2015	Cargo (kg)/ATM 2016
Norway	19.01	19.39	18.43	23.50
Faroe Islands	58.75	59.56	0.16	0.15
Ireland	57.83	57.42	0.00	0.00
Newfoundland	7.21	6.73	0.02	0.02
Northern Ireland	143.45	151.39	0.00	0.00
Scotland	6.70	6.51	0.29	0.28
Sweden	18.05	16.59	0.08	0.07

The main findings are:

- Northern Ireland is an outlier as the airports are too large.
- Scotland and Canada/Newfoundland either has small aircraft or low load factor but in any case with the lowest score on passengers/air traffic movement.
- Freight seems to be relatively important in Norway and to some extend in Scotland and Faroe Islands. Remoteness and island location appear to matter.

Table 4.2: Selected productivity PPMs

	pax/RWY length	pax/RWY length	pax/ terminal m²	pax/ terminal m²	pax/ FTE	pax/ FTE
	2015	**2016**	**2015**	**2016**	**2015**	**2016**
Norway	111.83	110.03	87.17	86.80	5484.56	5751.90
Faroe Islands	153.42	162.31	69.00	73.00	5750.00	5959.18
Ireland	255.41	276.59	73.21	80.66	5713.46	6088.89
Newfoundland	30.14	29.38	9.65	9.41	5939.70	5790.10
Northern Ireland	144.62	147.77	101.59	103.81	4444.66	4541.73
Scotland	16.27	18.17	37.11	39.69	969.84	1002.62
Sweden	34.82	35.01	36.65	36.86	2515.93	2507.65

The productivity measures indicates that Scottish and the Swedish airports in the sample are neither labour productive nor capital productive. Please note that a full DEA may affect this picture and that the sample size affects the robustness negatively. The airport in the sample from Canada/Newfoundland has a long runway to act as a backup for transatlantic flights, and is hence an outlier in the calculations of these productivity PPMS.

Table 4.3: Selected financial PPMs

	AR/ pax €	AR/ pax €	non-AR/ pax €	non-AR/ pax €	EBIT/ pax €	EBIT/ pax €
	2015	**2016**	**2015**	**2016**	**2015**	**2016**
Norway	9.02	8.50	3.72	3.76	-51.38	-45.74
Faroe Islands	7.47	5.82	2.61	3.77	5.23	1.47
Ireland	32.64	27.47	5.68	5.10	5.90	6.90
Newfoundland	26.31	22.33	12.16	13.67	9.74	6.26
Northern Ireland	5.77	5.47	0.70	0.66	-3.99	-3.78
Scotland	23.75	17.12	5.42	4.70	-23.35	-30.92
Sweden	17.68	38.73	27.14	33.16	-3.97	-5.04

Note: all financial data presented in 2015 and 2016 €

The main findings are

- Aeronautical revenues (AR) are highest in Canada/newfoundland and lowest in Northern Ireland.
- Non-aeronautical revenues (NAI) are relatively highest at Swedish airports and lowest at Northern Ireland airports.
- Low numbers of NAI is probably related to no or few additional services.
- There is a mixed picture in terms of profitability but airports in the main loss making. Given that these results are correct then Norwegian airports lose on average around 50 Euros per passenger and Scottish airport around 30 Euros per passenger.

Table 4.5: Unit cost results

	Staff unit cost (per FTE)	Staff unit cost (per FTE)	Longest RWY cost (per m)	Longest RWY cost (per m)	Terminal space unit cost (in m2)	Terminal space unit cost (in m2)
	2015	**2016**	**2015**	**2016**	**2015**	**2016**
Norway	67869.56	62011.65	1497.14	1421.55	1762.52	1702.55
Faroe Islands	36497.88	36751.20	744.68	722.95	334.92	325.15
Ireland	32052.11	30937.76	275.87	297.41	382.14	411.97
Newfoundland	62391.90	67934.24	1202.96	1170.16	208.57	202.88
Northern Ireland	33151.47	32116.51	434.29	420.73	305.09	295.56
Scotland	42667.20	40986.78	743.07	733.92	1343.89	1413.14
Sweden	51889.36	49194.77	412.30	483.72	935.54	926.54

Note: all financial data presented in 2015 and 2016 €

The main unit cost PPMs are:

- Unit labour cost in Scotland are coming down which is useful given the relatively low labour productivity
- Norway as expected highest unit labour cost because of the high cost level in general
- The terminal space unit costs are high in Norway and Scotland.

4.2. The DEA analysis

A two stage DEA model is used to identify the production (efficiency) frontier and inefficiency scores (distance of individual observation from efficient frontier) for each airport of our sample. A key advantage of DEA compared to parametric approaches such as Stochastic Frontier Approach (SFA) is that it performs better with small data samples and its application has been popular in the transport literature. The DEA approach is flexible, allowing the selection of multiple inputs and outputs, in our case staff numbers (FTE), terminal size (in m2) and length of the longest runway (in m, the former as a proxy for labour and the latter as proxies for deployed capital) and passenger numbers, cargo tonnes lifted and air traffic movements (ATM) as outputs. It also allows for non-constant returns to scale (Banker et al., 1984), which is important in the context of transport (in particular when one aims to identify differences between the urban and regional context), where variable returns to scale are prevalent. Moreover, when historical data are available, DEA can identify changes in efficiency arising from institutional reform or technical change, thus providing an understanding (for operators and policy-makers alike) as to whether the operating strategy has improved efficiency performance. Given the focus on cost efficiency, we us an input oriented model (that evaluates the efficiency of using various inputs for a given level of output). As typical in transportation we use a two-stage DEA efficiency approach in which, as part of the second stage, efficiency scores estimated in the first stage DEA models are used as dependent variables in a truncated (truncated at 0) second stage regression model to provide an understanding of the determinants of efficiency. Bootstrapping (with 2000 iterations) of the first stage DEA scores is used in order to minimise bias and unreliable second stage regression results (Simar and Wilson, 2000, 2008). From a policy point of view, the regression results that identify the determinants of inefficiencies form the evidence base to provide recommendations for the improvement of cost efficiency. This approach not only can identify technical efficiency (productivity), but also allocative (optimal mix of multiple inputs) and cost efficiency. The properties of DEA and truncated regression are elaborated further in Appendix 3.

Table 4.6: First-stage DEA results

		DEA results 2015 (incl cargo)				DEA results 2016 (incl cargo)			
	#	TE	TE_{cor}	AE	CE	TE	TE_{cor}	AE	CE
Norway	35	0.930	0.854	0.909	0.848	0.915	0.835	0.914	0.839
Faroe Islands	1	0.557	0.526	0.880	0.490	0.570	0.534	0.898	0.512
Ireland	3	0.749	0.680	0.839	0.648	0.776	0.697	0.852	0.676
Newfoundland	1	1.000	0.869	0.463	0.463	1.000	0.860	0.486	0.486
Northern Ireland	1	0.603	0.573	0.762	0.459	0.639	0.607	0.739	0.472
Scotland	5	0.695	0.630	0.829	0.595	0.702	0.632	0.832	0.606
Sweden	6	0.583	0.551	0.835	0.484	0.601	0.565	0.826	0.498
Total	52	0.845	0.776	0.877	0.748	0.84	0.766	0.88	0.747

Note: These are average DEA results. TE: Technical efficiency (original); TE: Technical efficiency (bias corrected/bootstrapped); AE: Allocative efficiency; CE: Cost efficiency

As shown in Section 2, the inputs of the DEA study were:

- Runway length of longest runway (RWYL),
- Terminal size in m^2,
- The number of full time equivalents.

Outputs were:

- Amount of cargo lifted,
- The number of passengers,
- The number of air traffic movements (ATM).

Variable Returns to Scale (VRS) was applied in this study.

The main results were:

- Sweden is on average the least efficient performer, which is mainly due to their airports being very small in terms of outputs, but please bear in mind the sample size.
- Newfoundland is the best performer in Technical Efficiency but relatively poor in Allocative Efficiency (too long runway for what they do in their ordinary services, but please bear in mind their role as a backup in transatlantic services) and hence their have a low score on Cost Efficiency.

- Overall majority of airports seems to be fairly inefficient (i.e. they could do better), which may be related to a couple of relatively large airports (in terms of outputs) being part of the sample. Again, the sample is thin and hence the results may be biased.

For the truncated regressions we had to drop some of our initially proposed explanatory variables such as ownership (only 2 out of our 52 sampled airports in private ownership) and military use. There are only 2 airports in the sample with military involvement, as all other military airports had to be omitted due to lack of cost and other data. Furthermore, ambulance flights on a regular basis was omitted as a variable because only four of our sampled airports did not provide such services. Finally, sufficient access to capital market was excluded as only three airports in the sample did not think they had sufficient access. This is in itself an interesting finding that we find worthwhile mentioning.

The variables that turned out to have sufficient variation in the sample were:

- Air services for offshore oil and gas fields (yes/no),
- Outsourcing of operations (in %),
- Public transport access or not,
- Group governance or stand-alone airports.

2015/2016 average has been used, as there are almost no variation in these explanatory variables over the two years.

Table 4.7 shows the truncated regression model results (truncated at 0).

Table 4.7: Second-stage DEA results

	TE	**TE$_{cor}$**	**AE**	**CE**
Governance (Group=1)	0.194***	0.167***	0.082**	0.243***
Offshore service (yes=1)	-0.072	-0.073	0.021	-0.57
Outsourcing (in %)	0.0031	0.0026	0.0033***	0.0056**
Public transport access (yes=1)	-0.042	-0.059	0.018	-0.018

The governance structure of regional airport management has a statistically significant impact on all types of airport efficiency. Most notably cost efficiency is higher for airports managed through a group/holding rather than a stand-alone entity. This is contrasting the findings of GAP (2012) who arrived at the opposite conclusion. On average, group governance airports are 24% more cost efficient than their counterparts. Outsourcing is the other variable with a significant impact on cost efficiency. One % of additional outsourcing leading on average to a cost efficiency improvement of 5.6%. This is entirely due to the positive effect that outsourcing has on allocative efficiency while it is insignificant in terms of having an impact on technical efficiency. From a theoretical perspective, this makes sense as allocative efficiency would naturally (i.e. how will the mix of resources be used in order to achieve a given level of output) be the efficiency that benefits the most from outsourcing.

5. Summary and conclusions

This study reports is carried out under Work Package A7.4. At the outset, the two main objectives were:

1. Identification of cost, technological and organizational leadership among NPP airports, and the characteristics of the best practice of airport management (what drives the efficiency and advanced leadership of top and under performers).
2. Dissemination of best practice from the leading peers to other comparable airports.

Partial Productivity Measures (PPM) and a 2-stage Data Envelopment Analysis (DEA) with bootstrapping were applied in order to identify the efficiency of each airport within the NPP area. The ambition was to have a large sample of airports in order to assess airport efficiency comprehensively for the NPP airports. This ambition was met only partially.

The second step in the DEA (truncated regressions based on the DEA results) address in a bit more detail issues that can explain the differences in terms of relative efficiency. The variables in the final assessment included airport governance, degree of outsourcing, whether the airports serve offshore operations, and accessibility with public transport.

The main results from the DEA study were:

- Sweden is on average the least efficient performer, which is mainly due to their airports being very small in terms of outputs, but please bear in mind the sample size.
- Newfoundland is the best performer in Technical Efficiency but relatively poor in Allocative Efficiency (too long runway for what they do in their ordinary services, but please bear in mind this airport's role as a backup in transatlantic services). Hence, they have a low score on Cost Efficiency.
- Overall majority of airports seems to be fairly inefficient (i.e. they could do better), which may be related to a couple of relatively large airports (in terms of outputs) being part of the sample. Again, the sample is thin and hence the results may be biased.

The second-stage regression made in order to explain the differences in efficiency scores gave the following main results:

- The governance structure of regional airport management has a statistically significant impact on all types of airport efficiency. Most notably cost efficiency is higher for airports managed through a group/holding rather than a stand-alone entity. This is contrasting the findings of GAP (2012) who arrived at the opposite conclusion. On average, group governance airports are 24% more cost efficient than their counterparts.
- Outsourcing is the other variable with a significant impact on cost efficiency. One % of additional outsourcing leading on average to a cost efficiency improvement of 5.6%. This is entirely due to the positive effect that outsourcing has on allocative efficiency while it is insignificant in terms of having an impact on technical efficiency. From a theoretical perspective, this makes sense as allocative efficiency would naturally (i.e. how will the mix of resources be used in order to achieve a given level of output) be the efficiency that benefits the most from outsourcing.

An important remark here is that the characteristics of the sample call for a very careful interpretation of the results. Nevertheless, one robust finding is that there are obvious elements of inefficiency among NPP airports. Another interesting finding is the indication that group ownership seems more efficient.

Ideally, a follow-up of any robust DEA results including an identification of 'learners' and 'peers' would entail a thorough review of various technical and operation aspects of airport management in a third analytical step. In order to do this properly, a mapping process will be required. However, the results became too uncertain to carry this out. Such mapping was performed by Avinor (who control most airports in Norway), comparing a number of smaller airports in Finland, Iceland, Norway and Sweden (Avinor As (2010), and we refer to this work if such a closer investigation of the airport operations are to be considered for individual airports.

References

Avinor (2010). Project benchmarking—Review benchmarking of small and remote regional airports.

Banker, R. D., Charnes, A. and Cooper, W. W. (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, p. 1078-1092.

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

Farrell, M. J. (1958). The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society*, A, 253-281.

GAP (2012). Comparative study (benchmarking) on the efficiency of Avinor´s airport operations.

Merkert R, R Pagliari, J Odeck, S Bråthen, N Halpern and J Husdal (2011). Benchmarking Avinor's Efficiency - a Prestudy. Report 1066, Møreforsking Molde AS

Simar, L. and Wilson, P. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, Vol.136, pp. 31-64.

Appendix 1: The questionnaire

Benchmarking of Airports in Europe's Northern Peripheral and Remote areas

Introduction

SPARA 2020 is a 3 Year €2.4m Project co funded by the Northern Periphery and Artic Programme 2014 – 2020, designed to address some of the challenges facing Peripheral and Remote Airports.

Sustaining, future-proofing and delivering community and regional resilience and connectivity with medium term, proactive responses to the special challenges of, and threats to, remote and peripheral airports/aviation is the aim of this project.

The project intends to maximise revenues in challenging settings, significantly control costs, increase self-sufficiency and resilience, and create a forum for raising standards and improved interactions with each other, with common challenges, with suppliers, and with regulators and government transport policy planners. For more information, please see http://spara2020.eu/.

The benchmarking study

To ensure fulfilment of the aforementioned objectives, SPARA2020 need to explore how the airports are performing and to identify the main factors for improvement with respect to cost efficiency. This part of SPARA2020 is lead by Professor Rico Merkert (ITLS, The University of Sydney Business School) and Professor Svein Bråthen (Molde University) and is part of the SPARA 2020 project. Other parts of the project will look into the airports' services as seen from the users' point of view, among other things.

All responses are completely voluntary and hence they will be made anonymous and reported at an aggregated level only. Please answer all questions to the best of your knowledge. Thank you for your valuable support with this study!

Please state the name of your airport and country:

____

Northern Periphery and Arctic Programme 2014–2020
EUROPEAN UNION
Investing in your future
European Regional Development Fund

Benchmarking and evaluation of NPP airports

Data collection form

This research is carried out by Professor Rico Merkert (ITLS, The University of Sydney Business School) and Professor Svein Bråthen (Molde University) and is part of the SPARA 2020 project. All responses are completely voluntary, will be made anonymous and reported at aggregated level only. Please answer all questions to the best of your knowledge. Thank you for your valuable support with this study!

Airport and country: ____

1. Quantitative measures

Variable	31 Dec 2015	31 Dec 2016
Quantitative measures I		
Number of runways in regular use		
Length of Runway 1 (in metres)		
Length of Runway 2 (in metres)		
Airport area (total area within the airport fence in m^2)		
Terminal size (total useable area in m^2)		
Other airport buildings' size (total useable area in m^2, e.g. for firefighting, cargo handling, aircraft hangars, flights schools, general aviation, for vehicles))		
Size of carpark for customers (in parking spaces including off-site)		
Size of carpark for staff (in parking spaces)		
Is there public transport access to the airport (yes/no)?		
Is there a flight school located at the airport (yes/no)?		

Variable	31 Dec 2015	31 Dec 2016
Quantitative measures II		
Number of employees (in full-time equivalents [FTE]):		
Number of employees (FTE) in the following areas (if the employees are working with more than one of the listed tasks, please distribute the FTE on the relevant tasks):		
a. Runway and taxiway operations		
b. Terminal operations excl. handling		
c. Handling		
d. Air traffic control		
e. Flight school		
f. Firefighting		
g. Overall maintenance		
h. Security services		
i. Cafés, shops		
j. Other non-airside activities		
k. Management and administration		
Number of passengers, departures (if in doubt include transfers)		
Number of passengers, arrivals (if in doubt include transfers)		
Number of passengers connecting with other flights (transfers)		
Number of passengers in transit		
Cargo shipped out+in (tons)		
Number of aircraft movements out+in		
Number of regular helicopter flights		
% of cancellations		
% of scheduled flights departing on time		
The duration of high season(s) (in months per year, high season defined as having longer opening hours and/or traffic more than 10 % above the average month.)		
Airport opening hours in high season (opening and closing hours. Please also state the periods if closed some time during the day.)		
Airport opening hours off season		
Approximate number of days per year with need for snow clearance		

2. Qualitative measures

Variable	31 Dec 2015	31 Dec 2016
Qualitative measures (see declarations below the table)		
Ownership (public/private)		
Governance (group vs. stand-alone)		
Military involvement (yes/no)		
Cooperation with the local community on operations (yes/no)		
Outsourcing of operations (in %)		
Ambulance flights on a regular basis (yes/no)		
Air services for offshore oil and gas fields (yes/no)		
Sufficient access to capital market, if relevant (yes/no/don't know)		

Below, we ask you to elaborate a bit more on the responses given in the table above. Please state the answers below for the current financial year:

Ownership of the airport (one cross-mark + sub-questions on alternative 3):

☐ 100 % private

☐ 100 % public (please state whether the public ownership is national, regional or local)

☐ Public/private

- Please state the share of private ownership

  ______ %

- Please state whether the public ownership is national, regional or local (one cross-mark)

  ☐ National

  ☐ Regional

  ☐ Local

Governance of the airport (one cross-mark):

☐ Stand-alone airport

☐ Part of a network of airports owned by the private sector

☐ Part of a network of airports owned by the public sector

☐ Part of a network of airports with public/private ownership

Cooperation with the local public community on operations (several cross-marks possible)

☐ No, all operations are done by the airport staff

☐ Yes, cooperation on firefighting

☐ Yes, cooperation on snow clearance

- ☐ Yes, cooperation on general maintenance
- ☐ Yes, cooperation on other general operations

Outsourcing of operations to the private sector (several cross-marks possible)

- ☐ No, all operations are done by the airport staff
- ☐ Yes, firefighting is outsourced to private operators
- ☐ Yes, snow clearance is mainly outsourced to private operators
- ☐ Yes, security is mainly outsourced to private operators
- ☐ Yes, air traffic control is mainly outsourced to private operators
- ☐ Yes, general maintenance is mainly outsourced to private operators
- ☐ Yes, other general operations are mainly outsourced to private operators

Would you say that your airport finds it difficult to raise sufficient funds (e.g. for terminal upgrades) from capital markets?

- ☐ Yes
- ☐ No
- ☐ Don’t know

3. Supplementary questions

It would be fantastic if you could also submit the annual reports of your airport to rico.merkert@sydney.edu.au and falko.muller@himolde.no for each year showing the operations costs and revenues. The operations costs could preferably be broken down into the categories listed as categories 2-6 in the question on "Outsourcing of operations to the private sector" above – or at least into Operation costs, Staff costs and Materials costs. In addition, and only if possible please state the Annual Insurance Expenses for the buildings. Are you aware of any relevant reports on the economic role of the airport from recent years (or reports in progress)? If so, we would be grateful for further details:

Thank you very much for your response and cooperation. If you are interested in a copy of the aggregated results (we guarantee full anonymity for all respondents), please leave your email address:

……….

Appendix 2: The original study proposal

A generic methodology for benchmarking and evaluation of NPP airports – Identifying key drivers of cost efficiency

Introduction

The main objective for a comprehensive benchmarking study is to reveal areas in which the NPP airports can improve, by comparing themselves with best practice airports that enjoy a competitive advantage.

We suggest that this part of SPARA2020 is organized as a WP that has two main objectives:

3. Identification of cost, technological and organizational leadership among NPP airports, and the characteristics of the best practice of airport management (what drives the efficiency and advanced leadership of top and under performers).
4. Dissemination of best practice from the leading peers to other comparable airports.

The first objective can be achieved by applying a number of different procedures. First, the best practice airports can be identified by means of Partial Productivity Measures (PPM) and/ or Data Envelopment Analysis (DEA). These methods can provide a robust identification of the efficiency of each airport within the NPP area. Efficiency development/changes over time can also be assessed. When the efficiency scores of different airports in the sample of around 60 partner airports are identified a second step (truncated regressions based on the DEA results) would address in more detail issues that can explain the differences in terms of relative efficiency. Those variables should include airport organization (incl. ownership), to what degree are the staff doing multi-tasking, opening hours, emergency preparedness issues, management structures and so forth. This is done in order to compare less efficient airports with their peers (the best practitioners within a relevant group of fairly similar airports) and look for areas of improvement. Our up to date and more NPP tailored sample as well as state-of-the-art econometrics will add significant value to existing knowledge at relatively small costs. For interpretation of the results we also envisage to incorporate technical and operation aspects of airport management. For that matter a mapping process will be required. Such mapping was carried out by Avinor (who control most airports in Norway), comparing a number of smaller airports in Finland, Iceland, Norway and Sweden (Avinor As (2010)[1]. The project leader at Avinor, Eigil Ulvin Olsen has been extensively consulted when writing this proposal and is willing to share the results and his insights in terms of methodology and how such benchmarking has improved the operation of a number of Avinor's airports.

This WP aims to cooperate closely with the other WPs.

[1] Avinor AS (2010). Project Benchmarking – Review benchmarking of small and remote regional airports.

Research methods

This section briefly presents the two main research methods that we find useful for SPARA2020. The methods are Partial Productivity Measures (PPM) and Data Envelopment Analysis (DEA). Both methods are well known from numerous benchmark studies within the airport industry as well as within other service industries. In addition, we will supplement our findings with a more engineering-based approach, inspired by an internal benchmarking study undertaken by Avinor, Norway in 2010.

The PPM approach

Applying partial indicators of performance is the traditional and most commonly used method to compare airports. Typically these studies focus on the following dimensions of airport performance:

- Cost efficiency
- Productivity
- Revenue generating capability
- Profitability

For each of these dimensions, measures have been developed which relate in some way the airport's inputs and outputs. The major inputs in an airport system are: labour and capital. Depending on the performance measure used, the inputs are measured in either physical or financial terms. For example, labour can be expressed in terms of number of employees or in terms of total labour costs incurred by the airport. Capital is usually measured in physical terms and can be represented by for example, the capacity of the runways or the amount of terminal space allocated to retail activities.

As far as output is concerned various measures can be used. Traffic represents the key output of an airport and there are typically three dimensions; passengers, freight or aircraft movements. For the majority of civil airports, the most important output is passenger traffic. However, some airports have substantial freight activity. The challenge for researchers in the partial performance field has been to devise robust and reliable measures of output that cover the different types of traffic. The Work Load Unit (WLU) was devised to solve this problem as it essentially combines passenger and freight volume into one aggregate measure of airport output.

Given that data is accessible, partial measures are intuitively very easy to compute, understand and interpret. Judgment would need to be made on whether to normalise the data as done in previous studies in order to take into account differences in the degree of outsourcing between airports. However, as the benchmarking exercise itself is measuring the outcome of managerial decision-making i.e. cost-efficiency, out-sourcing services will have an effect on performance and should ideally be incorporated in the analysis.

One significant limitation with partial measures is that they are less effective in providing a robust assessment of an airport's overall performance especially within the context of measuring the performance of the airport in relation to its optimum potential performance. Techniques like DEA are able to do this. However, the advantage of the PPM approach lies in the relative simplicity of the calculations and analysis of the results. Furthermore, PPM allows the researcher to investigate and compare performance at a disaggregated level; for example, measuring the non-aeronautical performance of different airports.

Northern Periphery and Arctic Programme 2014–2020
EUROPEAN UNION
Investing in your future
European Regional Development Fund

The DEA approach

It appears that Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA) are appropriate methods for assessing efficiency for Avinor. The reasons for this are discussed in the PS. The advantages of SFA over other methods until recently, are that it builds on econometrics and therefore it is able to capture noise in the data more adequately as compared to other methods. DEA has recently been developed to account for noise in data and hence SFA is no longer advantageous over it. We will have our main focus on the use of DEA instead of SFA.

The main advantages of DEA over other methods and which are relevant for the assessment of Avinor are as follows:

- DEA is easy to grasp and understand for managers; the benchmark is other service providers providing the same type of services using the same types of inputs and, these other providers are observable and not derived from some assumed production function.
- DEA readily incorporates multiple inputs and outputs and, it does not require price data to calculate technical efficiency. This makes it especially suitable for analysing the efficiency of service production, where it is often difficult to assign prices to many of the outputs.
- It determines sources of inefficiency and efficiency levels and provides a means of decomposing economic (cost) efficiency into technical and allocative efficiency. Furthermore, technical efficiency is decomposed into scale and non-scale effects.
- DEA identifies the "peers" for units (airports) that are not efficient. It thus provides a set of role models that the inefficient units can look to in order to improve its operations. This makes DEA a very useful tool for benchmarking compared to other methods.
- DEA can be extended to study efficiency over time using the Malmquist productivity index. Thus its advantages over other methods are maintained even when efficiency is being studied over time.

Like any assessment method, DEA too is based on a number of assumptions and hence has some weaknesses that one needs to acknowledge. The main ones are follows:

- DEA is a deterministic rather than a statistical approach. Its results would therefore be sensitive to measurement errors. However, recently it has been proven that applying DEA together with bootstrapping takes account of statistical noise adequately.
- DEA scores are sensitive to the number of inputs and outputs, and the sample size. An Increase of the sample size will tend to reduce the average efficiency score because including more observations provides greater scope for DEA to find a comparison partner. Conversely, fewer observations relative to the number of inputs and outputs can inflate the efficiency scores. There are however ways of dealing with this problem, as described in the PS. In this case, we would probably deal with somewhere between 7 and 15 inputs and outputs, which raises the need for data from between 25 and 50 airports. Below, we will provide a list of airport service providers that will be used in this study, which will satisfy this sample size.
- The method is also sensitive to outliers. We will pay close attention to this and to justify very carefully and with transparency if any airport should become omitted from the study for this reason.

Its few weaknesses can be easily overcome by applying both state of the art techniques (like using bootstrapping techniques) in the analysis and common sense during sampling and interpretation of the results. DEA has proven useful for investigating the efficiency of government service providers such as airports.

The four most commonly used input factors are:

- *Labour,*
- *Length and number of runways*
- *Terminal size*
- *Airport area*

The following output factors are commonly used:

- *Passengers*
- *Cargo*
- *Aircraft movements*

In sum, the most beneficial aspect of DEA is that it combines several inputs and outputs to produce a single overall efficiency measure. These factors have proven to be obtainable from various airport operators.

Our DEA approach will also be expanded by a second stage analysis where the DEA efficiency scores are regressed on external factors such as ownership, regulation, weather conditions etc., to infer how these factors influence the efficiency scores measured in the first stage.

We would recommend a study that takes the individual airports as the units of analysis. A further breakdown would probably leave a significant amount of uncertainty with respect to both data availability and comparison possibilities/robustness. Overhead costs could be allocated to each airport in line with what is described in Cranfield University (2006). For example, the allocation of central overhead costs to each airport on the basis of its proportion of traffic is one of several intuitive and valid methods of accounting for administration expenditure.

A study could be done in two parts, and these parts are complementary in nature, also because they will rely on the same data set:

1. A 2-stage DEA analysis and PPM on a larger set of NPP airports. Depending of the data, both technical and cost efficiency can be examined.
2. Econometric benchmarking and PPM analysis of clusters of airports, where we will select the largest/busiest airport of a given country compared with the largest/busiest airports in the peer countries; and then the same with the smallest airports. In that way we aim at separating the impact of size. In general, it will be fairly easy to address various groupings of airports once the data set is established.

In both parts of the study the DEA models will be conducted in two stages, where in the second stage the DEA efficiency scores are regressed in a second stage on external factors such as

- *Ownership,*
- *Governance (e.g. part of and largely managed by a larger airport group),*
- *Size of airport,*

- *Military involvement (shared cost/responsibilities),*
- *Degree of outsourcing,*
- *Type of regulation,*
- *Weather conditions etc.,*

to infer how these factors influence the efficiency. Furthermore, DEA will be conducted together with bootstrapping to certain confidence intervals for the efficiency scores derived. Finally, DEA's extension to the Malmquist Productivity Index (MPI) will be used to study the developments in efficiency over time.

The engineering-based approach

In Section 1 we referred to the benchmark study by Avinor (2010). The study is very detailed, and it advises some specific steps to improve the airport operations efficiency with respect to detailed technical and organizational aspects of airport operations.

An approach in line with the Avinor study is likely to provide valuable input to SPARA 2020. Because DEA identifies the "peers" for units (airports) that are not efficient, the approach in Avinor's own study appears as convenient for looking into the inefficient units and comparing them with their "peers" in order to improve their operations. Since SPARA2020 is likely to have around 60 airports as partners, data for such a detailed comparison should be available. However, the amount of work will be a bit uncertain, and to a certain extent dependent upon the result from the DEA study (in terms of the number of airports that might need closer examination) and also_on very specific insights into the details in airport operations. We still consider this approach to be a viable complementary step which will add significant value. The detailed technical insight into airport operations which will be necessary to do the more detailed approach as suggested by the Avinor study, will be facilitated through close contact with Avinor's project team. Examples of elements that need to be looked into in this part of the study are:

- Regulatory issues. An update on regulation and their impacts
- Different organizational structures
- Outsourcing
- Staff multi-tasking
- Automation of operations, current state and potential for development
- Collaboration and coordination with the airlines
- Off-site development, like paid parking
- Emergency preparedness (like ambulance flights)

Snap-shots of the current state should be made in the beginning of the project period, as basis for ex post evaluation.

Data availability

We have focused on a comparison between operators in the NPP countries with a fairly similar governance structure for airport operations. In terms of governance structures; Avinor, Swedavia, Finavia and Isavia are very similar. They are government-owned airport operating companies (and limited companies). They tend to have an overall corporate structure and a centralised administration but with divisions that are responsible for groups of airports. Their largest airport (the capital city airport) is typically owned and operated as a subsidiary of the airport operator. HIAL is operated in a slightly different way.

Norway: Avinor have good data on all relevant input and output factors for a DEA study. Some additional data would have to be collected for the two-stage regression analysis (like e.g. weather conditions and outsourcing of operations), but this data is easy to obtain. In addition to Avinor's airports, operational data (for a DEA technical efficiency study) from Sandefjord/Torp (TRF) will be available. Data from the newest private airport, Moss/Rygge (RYG) are not available. The airport has been operating only since 2008, which makes it less suitable for comparison.

Sweden. Traffic data is available from 1997-2009 for 41 airports (passengers, freight and mail, and aircraft movements). 14 of the airports are operated by Swedavia and financial information is available from 2004-2009 for each of those airports (e.g. capital expenditure, labour cost, turnover, operating result). Operational data is also available (e.g. on terminal space, runways, staff). Staff numbers may have to be converted to full time equivalents. For non-Swedavia airports, contact is established and data can be provided upon request.

Finland. Traffic data from 1998-2009+ is available for 27 airports. 25 airports are operated by Finavia. Financial accounts for Finavia airports are consolidated in their annual report. They can provide 10 years of financial information and human resources data for each of their 25 airports. Operational data for each of Finavia's airports (e.g. on terminal space, runways, staff) is available.

Iceland. All of the airports are operated by Isavia (14 airports). We have good reasons to believe that data from these airports will be available.

Scotland: For the Highlands and Islands group of airports, there is access to annual accounts from 2003/04 – 2008/09+. Although their annual accounts are aggregated for the group, they report operating costs and revenue for the following airports individually: Barra, Benbecula, Campbeltown, Dundee Airport Ltd, Inverness, Islay, Kirkwall, Stornoway, Sumburgh, Tiree, Wick. It should also be possible to get suitable operational data.

We will direct our attention towards the around 60 SPARA 2020 partner airports in this study for necessary information needed for the DEA/PPM study. In addition, it should be possible to obtain more detailed technical, operational and organizational data for each of these airports for the engineering-based part of the study.

Output

The main output from this study for the airports will be

- The ability to compare themselves with relevant peers with respect to an estimated potential for efficiency improvement, and
- To get a deeper understanding with respect to the determinants for differences in efficiency, by doing more detailed comparative studies. Some of these determinants may be external factors like weather conditions whereas others (like staff multi-tasking, outsourcing and other technical and operational issues) are more internal factors that can be influenced by the airport management.
- Linking the efficiency findings with economic impact and other benchmarking (e.g. extent of non-aeronautical activities) results of other WPs.

The main purpose is to identify in some detail areas of improvement.

Appendix 3: Analytical approaches to productivity, effectiveness and efficiency

Assessing performance of production units is a complex task and there are many terms used to refer to it. Efficiency for instance is a term used to refer to how well a unit of production is performing. Other terms frequently used interchangeably with efficiency includes productivity, Total factor Productivity(TFP) and effectiveness It is therefore important to make a distinction between these terms especially in the way we intend to use them here; in order to avoid misunderstandings.

Productivity refers to the ratio between output(s) and input(s). If for instance, one is comparing two units of production, the one with the highest output/input ratio is considered as being more productive than the other. Note that this is irrespective of whether outputs/inputs are measurable in monetary terms or not. Note further that, productivity defined in this way has no upper limit such as how high the output/input ratio must be in order to be efficient.

Effectiveness, unlike productivity refers to the degree to which the outputs of a service provider achieve the stated objectives of that service. For example, the Ministry of Transport and Communication may state objectives to be met by the airports such as maximum waiting time for luggage. Effectiveness is thus a measure of how these objectives are met.

Efficiency builds on the concept of productivity defined above. It refers to the degree to which productivities defined above -including all inputs and outputs that matters - matches the optimal productivities. It is thus a relative measure where productivities are related to some production frontier; either constructed from best practices or constructed econometrically using some known or accepted functional forms.

From the definitions above it is clear that productivities either total or partial does not measure efficiency for the simple reason that they do not relate to a given standard i.e., how large they should be in order for a unit of production to strive to be efficient according to some rule of measure. Efficiency measurement is considered as the appropriate measure of performance.

Different efficiency concepts

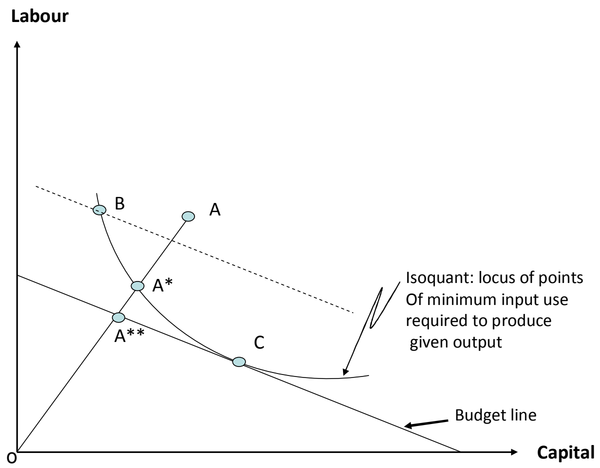
The most common efficiency concept is *technical efficiency* which refers to the conversion of physical inputs- such as labour and capital- into outputs relative to best practice. Airports or airport authorities that operates at best practice are said to be 100% technically efficient i.e., they do not waste resources. Any airport that operates below best practice is said to be technically inefficient and the airports technical efficiency is expressed as a percentage of best practice. Technical efficiency will be affected by scale or size of operations since it is based on engineering relationships but not on prices and costs.

The second efficiency concept is *allocative efficiency* which refers to whether inputs, for a given level of output and set of input prices, are chosen to minimize the cost of production assuming that the airport being examined is fully efficient. As opposed to technical efficiency, which shows excess use of inputs, allocative efficiency shows whether the right mix of inputs are chosen. It is also expressed as a percentage of scores where 100% indicates that airport is using its inputs in the proportions which minimize its costs; it is allocative efficient. It should be noted that an airport that is operating at best practice in terms of technical efficiency could still be allocatively inefficient since it is not using its inputs in the proportion that minimizes its costs, given the relative input prices.

The last efficiency concept is the *cost efficiency* which refers to the combination of technical and allocative efficiency. An airport will be cost efficient if it is both technical and allocative efficient. Cost efficiency is calculated as the product of technical and allocative efficiency scores expressed as a percentage. Thus, an airport can only achieve a 100% cost efficiency score if it is has achieved 100% scores in both technical and allocative efficiency.

The three efficiency concepts are illustrated in figure 1. The inputs we consider are labour and capital that are required to produce airports services, e.g., number passengers handled. The curve plotted in the figure is the isoquant: it plots the minimum amounts of labour and capital required to produce a given output quantity (given number of passengers handled). It is thus also called *technical efficiency frontier*. If an airport (or organization in question) is producing at a point on the curve such as B, A* or C, then that airport is technically efficient as opposed to A which is technically inefficient. The straight line denoted budget line plots the combination of labour and capital that have the same cost. The slope of this budget line is given by the negative of the ratio of capital price to the labour price. Budget line closer to the origin implies lower total cost. It follows that the cost of producing a given output is minimized at the point where the budget line is tangent to the isoquant (efficiency frontier); at this point, technical and allocative efficiency are attained.

Figure 1: illustrating efficiency concepts



Now consider the point of operations marked A, A*, A**, B and C in the Figure. An airport operating at point A would be technically inefficient because it uses more inputs than required to produce at the frontier (isoquant). An airport at point B would on the other hand be technically efficient but not cost efficient. B could for example, maintain its level of production while producing at a less cost by moving to point C. An airport operating at C is both technically and allocatively efficient and hence also cost efficient. There is a way of calculating the efficiency scores for individual units of production such as those plotted in the figure. Suppose an airport situated in A moved to point C in order to be cost efficient; by how much would its cost efficiency increase? Because it is both technically and allocatively inefficient would have to improve in both to be cost efficiency. In terms of technical efficiency it will have to increase efficiency by the distance $\frac{(OA-OA^{*})}{OA}$ to reach the frontier. In terms of allocative efficiency it will have to increase efficiency by $\frac{(OA^{*}-OA^{**})}{OA^{*}}$. Since the cost efficiency is the product of technical and allocative efficiency, its cost efficiency will increase by

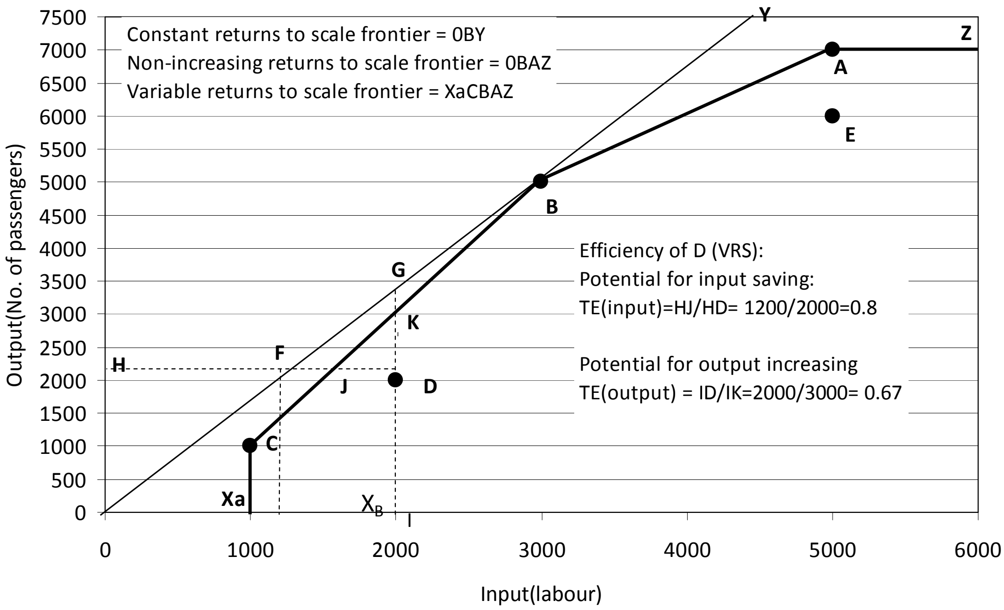
$$\frac{(OA-OA^{*})}{OA}\times\frac{(OA^{*}-OA^{**})}{OA^{*}}=\frac{(OA-OA^{**})}{OA}.$$

Returns to scale is another important concept of efficiency measurement. It refers to changes in ouput subsequent to a proportional change in all inputs (where all inputs increase by a constant factor). If output increases by that same proportional change then there are constant returns to scale (CRS). If output increases by less than that proportional change, there are decreasing returns to scale (DRS). If output increases by more than that proportion, there are increasing returns to scale (IRS). Variable returns to scale occur when there is a mix of all the aforementioned. The example illustrated in Figure 1 above was based on the assumption of constant returns to scale. This assumption essentially means that the size of production units is considered irrelevant when measuring relative efficiency; while in practice size may matter. If for example it is assumed that all airports operate with constant returns to scale, it means that doubling of inputs will double output irrespective of the size of the airport. This will imply that there are no economies or diseconomies of scale in the production of airport services. Obviously such an assumption would be unrealistic. Some airports would be too small and therefore operates with increasing returns to scale while other are too large and therefore operates with decreasing returns to scale.

It would be to the airports advantage to operate with optimal scale; neither too small if there are increasing returns scale and neither too large if there are decreasing returns to scale. Thus, when assessing efficiency, both technical and scale efficiency should be examined.

The concept of scale and how it relates to CRS and VRS frontiers can be demonstrated using a simplified one input (labour cost) and one output (number of passengers) case as shown in Figure 2. Five airports, A,B,C,D,E are being evaluated.

Figure 2: Production frontier and returns to scale



The line 0BY represents the constant returns to scale (CRS) frontier while XaCBA represents the variable returns to scale (VRS) frontier. The distance from the respective frontiers determines technical efficiency under each assumption. Any efficient airport with respect to the CRS will naturally also be scale efficient. Thus, scale efficiency is calculated as the distance between the constant and variable returns to scale frontiers. Consequently, airport B is the only airport that is scale efficient. Airports A and C are technically efficient according the VRS frontier, but are not scale efficient. Consider now how the efficiency, using all the scale concepts, can be derived for airport D which is not efficient according to any of the frontiers. The following efficiency measures can be calculated:

- $\text{Input saving efficiency (CRS)} = \frac{HF}{HD} = \text{Output increasing efficiency (CRS)} = \frac{X_B D}{X_B G}$
- $\text{Input saving efficiency (VRS)} = \frac{HJ}{HD}$
- $\text{Output increasing efficiency (VRS)} = \frac{X_B D}{X_B G}$
- $\text{Scale efficiency} = \frac{HF}{HJ} = \frac{\frac{HF}{HD}}{\frac{HJ}{HD}} = \frac{CRS\ efficiency}{VRS\ efficiency}$

It follows that scale efficiency is calculated as the ratio of CRS to VRS efficiency scores or that inefficiency is composed of two parts; pure technical inefficiency (HJ/HD) and scale inefficiency (HF/HJ). For airports A and C, they are technically efficient but not scale efficient. A operates with decreasing returns to scale and hence could (theoretically) reduce its size to be scale efficient. C on the other hand operates with increasing returns to scale and could increase its size to be scale and fully efficient.

The Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA) is regarded as being one of the most successful techniques of efficiency assessment proposed by researchers in Management Science and Operations Research, as is evident by the diversity of its application, also in the aviation sector. The seminal work can be found in Farrel (1958). DEA is a linear programming approach to calculating the efficiency of an organisation within a group relative to best observed practice within that group. The organisations can be whole agencies (for example, airport groups), separate entities within the agency (for example, individual airports managed by an airport group) or disaggregated business units within the separate entities (for example, terminal activities of an airport).

DEA proceeds by defining the best practice frontier composed of the most efficient units. The relative efficiencies of the remaining units are measured as a distance from this frontier. The best practice frontier is non-parametric, i.e. no functional form needs to be specified or assumed, in contrast to other parametric production frontiers such as Stochastic Frontiers Analysis (SFA). The DEA method allows multiple outputs and inputs. Inputs may be variable and fixed, where the values of the variable inputs are allowed to change in the short run (in the airport industry inputs are e.g., number of employees, and number of check in desks) while the values of the fixed inputs are only allowed to change in the long run (fixed inputs are e.g. Number of runways). The DEA method can be input or output orientated, of which the former determines the minimum input for which the observed production of a unit is possible, while the latter determines the maximum output of the unit given the observed inputs.

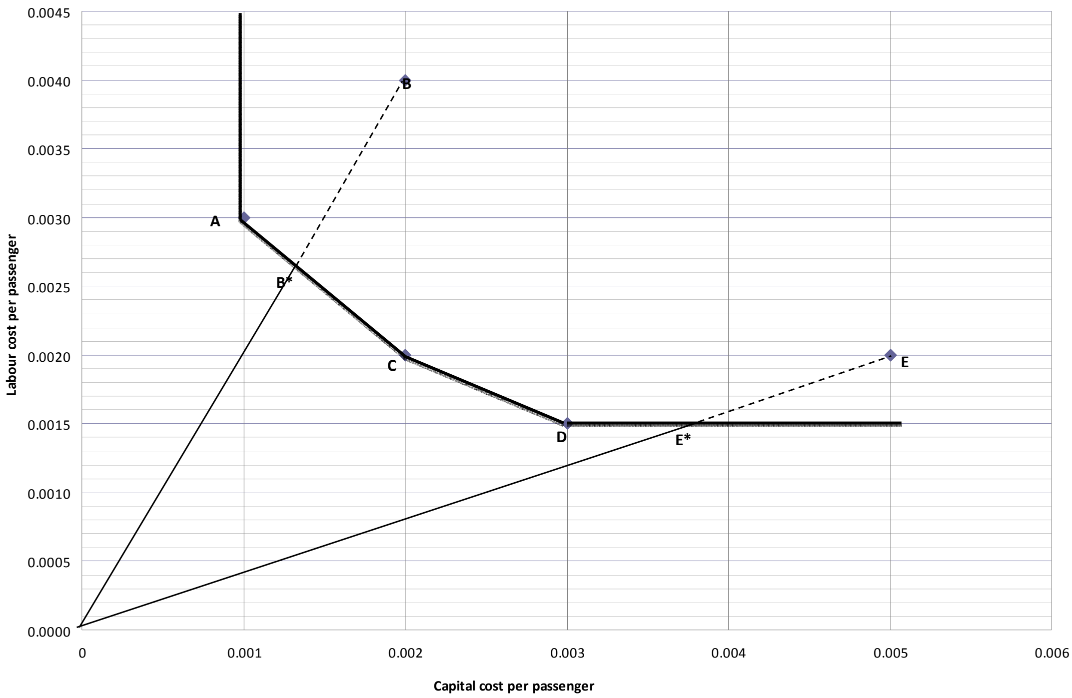
To demonstrate the workings of DEA in calculating the efficiency of comparable units in an industry such as airport or aviation industry, we consider a simple numerical example: a sample of five hypothetical airports that use two inputs- labour and capital one input - to produce one output - number of passengers handled at the airport. Obviously, the inputs and outputs of a real airport are considerably much more complex, but this simplification is a good starting point for actual as well as illustrative purpose. The data for the five hypothetical airports are presented in Table 1.

Table 1: Illustrative hypothetical data on five airports

	Output	*Inputs [in mill NOK]*		*Input/output ratios*	
	Number of passenger(O1)	*Labour costs(I1)*	*Capital costs (I2)*	*I1/O1*	*I2/O1*
Airport A	200 000	200	600	0.001	0.003
Airport B	300 000	600	1200	0.002	0.004
Airport C	100 000	200	200	0.002	0.002
Airport D	200 000	600	300	0.003	0.002
Airport E	100 000	500	200	0.005	0.002

The five airports range in size from 200 to 600 in terms of labour costs and there is similarly a large variation in capital costs. Given the large discrepancies among the five airports' characteristics it is not obvious how to compare them or which airports should be a role model for others to improve their performances. This is where the workings of DEA comes to use; the answers to the questions become clearer when the input/output ratios - Labour costs per passenger and capital cost per passenger- are plotted as in Figure 3. It is intuitively obvious that the smaller the output/input ratios, the more efficient the airport must be. Thus the airports closest to the origin and the two axes are the most efficient.

Figure 3: Illustrating airport efficiency measurement using DEA



The line A, C, D represents what is known as the "best-practice frontier". Points on this frontier are considered as being 100 % efficient. The best practice airports are therefore airports A, C and D; they all have an efficiency rating of 1.0(i., e.100 % efficient). All other airports north-east of the frontier are considered inefficient and will have an efficiency rating of less than 1.0; they use more inputs (labour and capital) to produce the same throughput as the airports on the frontier. They thus can be able to reduce their input use (labour and capital) and still maintain their output levels as compared with best practice airports.

Consider airport B, who uses more inputs than is required to be on the frontier. The efficiency rating for airport B is the ratio of the distance of the line segment from the origin to point B* and from the origin to point B (the ratio of best practice to observed inputs). The efficiency of airport B is thus calculated as:

$$E_B = \frac{0B^*}{0B}$$

In Figure 3, the above ratio is found to be $\frac{0.0027}{0.004} = \frac{0.00135}{0.002} = 0.67$; this means that airport B has an efficiency score of 0.67. Hence, the input-saving potential for airport B, the percentage by which the airport would have to reduce its inputs to achieve the best practice frontier is 33% (= (1-0.67)*100). The same reasoning can be used to derive the efficiency score and input-saving potential for other inefficient airport E; the efficiency score and the potential for input-saving respectively $\frac{0.0015}{0.002} = \frac{0.001375}{0.005} = 0.75$ and 25%.

There is a further characteristic of DEA in this example that needs to be explained, especially as far as benchmarking is concerned. Consider the inefficient airport B. Intuitively; we see from Figure 1 that it is aiming to produce the same results as A and C who are on the frontier. However, its airport of comparison has been a "virtual or hypothetical" airport B*. The virtual airport B* is a combination of or a weighted average of the operations of airports A and C. If airport B is to be benchmarked against any other airports as role models to improve performance, then it should examine the operations of airports A and C because these are the most efficient airports similar to itself; its 'peers' . Thus, DEA as method of efficiency assessments is able to identify the 'peers' of which the 'learners', the inefficient airports can be compared to in order to learn and improve.

The DEA example above is relatively easy to understand and implement, especially in a two-dimensional diagram as in Figure 3. However, when the there are many inputs and outputs as often is the case for airport services, DEA method is no longer amenable to simple graphical analysis. It is necessary to use linear programming techniques and computer packages to solve for the efficiency scores and potentials for improvements for the individual airports that are being compared. We briefly present here the linear programming (LP) problem due to Charnes, Cooper and Rhodes; hereafter (CCR, 1978), for finding the efficiency score (E_i) for an airport *i*:

$$
\left.
\begin{aligned}
&\text{Min } E_1 \\
&\text{s.t} \sum_{j=1}^{n} \omega_j x_j - E_1 x_o = - s_j^- && \text{(a)} \\
&\sum_{j=1}^{n} \omega_j y_j - y_o = s_j^+ && \text{(b)} \\
&\omega_j \geq 0, j = 1, . . ., n && \text{(c)}
\end{aligned}
\right\}
$$

where x_0 and y_0 respectively, denote the input and output vectors for selected airport. E_1 is the input saving efficiency measure of unit 0 under evaluation. ω_j is the non-negative weight of unit J's outputs and inputs that defines a comparison point on the frontier. Restriction (a) states that the efficiency-corrected use of inputs ($E_1 x_0$) must at least equal the amounts employed by the reference company. Constraint (b) states that the reference company must produce as much output as Company J. Note that the CCR formulation is non-flexible in the sense that it assumes constant returns to scale (CRS) in its production possibility set. It can, however, be modified to include variable returns to scale (VRS); see Banker et al. (1984), hereafter BCC. This modification implies adding a convexity constraint limiting the summation of the multiplier weights (ω) equal to 1, i.e., including $\sum_{j=1}^{n} \omega_j = 1$ in the model above. The linear program above is run sequentially for each of (n) airports. Technically, efficient airports are identified in units that have input and output slack vectors s_j^- = 0 and s_j^+ = 0 in addition to E_1 = 1 at optimality. These best-practice airports display either an optimal composite of inputs (or outputs) or a single exceptional input-output ratio. Less efficient airports will obtain a z-score of less than 1 and might have non-zero input or output slacks. In order to compute the output-oriented measure E2, the reciprocal of model (1) above may be considered. The objective is then to maximize output within the given finite stock of inputs available.

The technical efficiency derived from CCR and BCC formulations can be used to obtain a measure of scale efficiency as:

$$
SE_k = \frac{E_{CCR_k}}{E_{BCCk}}
$$

where SE_k indicates the scale efficiency of k th airport, where E_{CCR_k} and E_{BCC_k} are the technical efficiency measures for airport k, derived from applying CCR and BCC formulations respectively.

$SE_k = 1$ indicates scale efficiency, and $SE_k < 1$ indicates scale inefficiency. Scale inefficiency, however, is due to either increasing or decreasing returns to scale. Which is the case can be determined by inspecting the sum of weights under the CCR formulation:

$$SW = \sum_{j=1}^{n} \omega_j$$

SW= 1 will provide constant returns to scale (optimal scale), SW>1 decreasing returns to scale (superoptimal scale) and SW<1 increasing returns to scale (suboptimal scale).

Now, there may be factors outside the control of the organisation but that may impact efficiency. For instance, weather conditions which are not under the control of airports management may impact the performance of airports. In order to gain the role that external factors such as weather play on the efficiency performance of airports, the so-called second stage DEA may be conducted. It entails regressing the efficiency scores obtained from the first stage on the external factors and interpreting the results. The first stage results may then be corrected up or downwards depending on the regression results.

Truncated regression

The efficiency scores from the DEA where the efficiency of the airport services are compared may have underlying causes that should be addressed.

First, the difference in efficiency may not be statistically significant. The cause may be based on coincidences rather than differences in the contract regime. Second, there may be other explanatory factors. As an example, airports may have different governance structures that may affect efficiency. Another example is differences in the services, like heliports in order to serve offshore activities.

Data for these and other explanatory factors are collected, and their impacts on efficiency can be tested by means of a Tobit regression model as described in Maddala (1983). Tobit regression is specifically useful here because the data on inefficiency is truncated, i.e. the dependent variable is only observed in the interval [0,1]. The model can be written as:

$$\left.\begin{array}{l} Y_i^* = \beta X_i + \mu_i, \\ Y_i = Y_i^*, \ \text{ if } \ Y_i^* > 0, \\ Y_i^* = 0, \ \text{ otherwise} \end{array}\right\}$$

X_i is a vector of explanatory variables and β is a vector of parameters to be estimated. Y_i is the latent variable, which can be viewed as a threshold beyond which the explanatory variables must impact in order for Y_i to “jump” from 0 to some positive value. In this case it is legitimate to view the inefficiency scores as continuous variables limited a minimum value of 0. Thus the threshold has no particular interpretation in this case, but the model specification can be estimated by a maximum likelihood method assuming normally distributed errors μ_i. For the present analysis we formulated a Tobit model with inefficiency, defined as 1 minus the efficiency score as the dependent variable (Y_i). The relevant independent variables in the vector X_i in this study were:

- Form of governance (member of a group or individual airports)
- Serving offshore activities or not
- Degree of outsourcing of operations like firefighting etc.
- Accessibility with public transport.

Concluding remarks

From the expositions above, DEA provides the efficiencies of individual airports or aviation authorities as compared to others and, identifies possible benchmarks towards which performances can be targeted. The actual level of inputs or outputs of efficient airports can serve as target for the inefficient units and managers can improve by identifying 'peers' and learning from them.

- DEA is easy to grasp and understand for managers; the benchmark is other service providers providing the same type of services using the same types of inputs and, these other providers are observable and not derived from some assumed production function.
- DEA readily incorporate multiple inputs and outputs and, it does not require price data to calculate technical efficiency. This makes it especially suitable analysing the efficiency of service production where it is often difficult to assign prices to many of the outputs.
- It determines sources of inefficiency and efficiency levels and provides a means of decomposing economic (cost) efficiency into technical and allocative efficiency. Furthermore, technical efficiency is decomposed scale effects and non-scale effects.
- DEA identifies the 'peers' for units (airports) that are not efficient. It thus provides a set of role models that the inefficient units can look to for way of improving its operations.
- DEA can be extended to study efficiency over time using the Malquist productivity index (MPI). Thus its advantages over other methods are maintained even when efficiency is being studied over time. MPI has not been applied in this study.
- Truncated regression can be used in order to deepen the understanding of the causes behind differences in efficiency scores.

Like any assessment method, DEA too is based on a number of assumptions and hence has some weakness that need to acknowledge. The main ones are follows:

- DEA is a deterministic rather than a statistical approach. Its results would therefore be sensitive to measurement error. However, recently it has been proven that applying DEA together with bootstrapping takes account of noise adequately.
- DEA only measures efficiency relative to best practice within a particular sample. Thus it is not meaningful to compare efficiency scores across samples or across different studies.
- DEA scores are sensitive to the number of inputs and outputs and, the sample size. Increasing the sample size will tend to reduce the average efficiency score because including more observations provides greater scope for DEA to find a comparison partner. Conversely, fewer observations relative to the number of inputs and outputs can inflate the efficiency scores. There are however ways of dealing with this problem. A rule of thumb is that the number of

units in the sample should be at least three times greater than the sum of the number of outputs and inputs included in the analysis.

Despite its few weakness most of which can be corrected for e.g., by applying bootstrapping method, DEA is a useful for investigating the efficiency of service providers such as airports.

Closer examination of the potential for efficiency improvement on a specific airport probably needs an engineering-based approach as described in Avinor (2010).

Northern Periphery and Arctic Programme
2014–2020
EUROPEAN UNION
Investing in your future
European Regional Development Fund