Benchmarking of Airports - A Critical Assessment

Vanessa Liebert[1] and Hans-Martin Niemeier[2]

Abstract:

Performance measurement of airports serves different purposes. While the management perspective focuses mostly on partial processes the policy perspective of regulators and politics focuses on the overall productivity and efficiency. These different perspectives are reflected also in the methods. Partial measurers are popular with management, but might be misleading and need careful interpretation. Total measures are mainly used to assess overall performances, but demand sophisticated methods such as DEA, SFA or price-based index approaches and the knowledge of the airport production technology. The latter is especially often missing as required inputs and outputs are often not available. Further, for deriving comparable results, the operating environment of airports needs to be carefully considered because inefficiencies are not only caused by management decisions but also due to the operating environment. So far, most studies failed to provide conclusive answers to questions on the effects of privatization on the airports performance or size effects. Other aspects such as the effects of regulation and competition were hardly covered although it is well known that the type of regulation (cost versus incentive regulation) or regional competition is more important than ownership. This paper provides a literature review on the scope, methods and results of studies on the productivity and efficiency assessment of airports. It aims in particular to critically evaluate the limitations for a deeper understanding of the production process of airports.

Keywords: airports, efficiency, productivity, benchmarking, literature review

Topic area: A1 Airports and Aviation, SIG-8 (Air Transport Research Society)

Session Track

[1] Jacobs University, School of Humanities and Social Sciences, Campus Ring 1, 28759 Bremen, Germany, vanessakamp@yahoo.de

[2] University of Applied Sciences, School of International Business, Werderstr. 73, 28199 Bremen, Germany, hniemeier@fbn.hs-bremen.de

1 Introduction

Performance measurement of companies can serve different purposes as outlined by Oum et al. (1992). It can be used to assess the performance of units within or across companies or industries. The availability of panel data allows measuring productivity and efficiency changes over time. Airports offer an equally challenging and interesting objective for applying performance and benchmarking methods. Traditionally, airports were managed and regulated as public utilities, which is still present in many countries. However, in the Eighties with the privatization of the BAA a worldwide process of privatization emerged which was accompanied by regulatory reforms towards incentive and light-handed regulation. A change in the management style and the attitudes are reflected in the increase of commercialization which has lead to substantial investments in non-aeronautical activities. Furthermore, competition between airports began to arise. Some airports like Manchester or London Stansted face effective competition from local airports. With the deregulation of the airline industry some airports such as Pittsburgh or Brussels lost their hub carrier and experienced competition with other hub airports. The liberalization of bilateral air service agreements offered the opportunity to attract international traffic to hubs and secondary hubs. The increase in traffic has lead in many countries to excess demand at major airports which were rationed rather inefficiently by queuing and slot allocation mechanisms. The vertical and horizontal boundaries have changed over time. While some airports outsourced labour-intensive activities such as ground handling and become international airport companies others have remained highly integrated either as a separate airport or within an airport system organized as a civil aviation authority like in Finland. The technology of airports has changed due to innovations; as an example many airports operate with automated baggage handling systems, thereby shifting the production frontier outwards. For these reasons airports should offer a rich field for performance and benchmarking analyses as it offers diverse and heterogeneous characteristics.

Following Hensher and Waters (1993) there are in principle three quantitative methods that will be applied in the productivity and efficiency analysis among government enterprises which are (i) non-parametric index number approaches to measure the total factor productivity (TFP), (ii) parametric (econometric) analyses such as Stochastic Frontier Analysis (SFA) and (iii) non-parametric linear programming approaches as Data Envelopment Analysis (DEA). All methodologies are substantially different in its model specification and data requirements and might consequently lead to different results, even within one technique as proven by Diewert (1989) with six different index number approaches.

Although the instrument of performance measurement has already been applied in other transport sectors or regulated utilities in the Seventies, it became primary important in the airport industry twenty years later (Forsyth 2000). Since the late Nineties, more than 50 research papers on airport benchmarking emerged. Gillen and Lall (1997) and Hooper and Hensher (1997) were the first who assessed the airports efficiency and productivity with DEA and index-based TFP on US and Australian airports respectively. These articles vary in a number of aspects such as the objective of the study, the underlying methodology, the inputs, outputs or environmental variables and the selection of airports or the time frame. However, over the years, significant progress has been made in the studies to improve the performance assessment of airports. The majority of recent studies consider the heterogeneous character of airports and assume heterogeneous production or cost functions in parametric studies or conduct second stage analyses to explain inefficiencies of DEA estimates. Further, a few studies discovered the overstatement of the airport's performance without undesirable outputs and directly include them in the model.

Kincaid and Tretheway (2006) as well as Morrison (2009) challenged the current practice of airport performance analysis for inconsistencies and its limited value to managers. They stated that the clear definition of an airport model is crucial to understand the industry. Therefore the collection of consistent outputs and inputs (incl. capital) is very important. Further, methodological variety, different underlying assumptions and less consideration of the heterogeneous character of airports can lead to different results across studies. Especially the last argument was discussed in a response to Morrison (2009) by Adler et al. (2009) as econometric approaches have substantially been developed to account for observed heterogeneity across decision making units. From our point of view, the debate has not resolved all methodological issues and will continue because airport managers prefer to use simple partial measures while academics make use of sophisticated overall productivity and efficiency approaches.

Besides the review articles by Forsyth (2000) and Graham (2005) and a substantial number of authors who included a brief overview of the current literature in their studies (e.g. Barros 2009; Oum et al. 2008; Tovar and Martín-Cejas 2010), no extensive survey of airport benchmarking studies has been carried out so far which also discusses the results. For other transport sectors, we could find literature reviews by Oum et al. (1999) on the rail sector, De Borger et al. (2002) on public transport and Gonzalez and Trujillo (2009) on seaports, all critically reviewing the studies with respect to the underlying methodology, the data and the results.

The aim of this paper is to give a review of the current empirical literature on airport efficiency and productivity analysis and to examine what we can learn from existing studies so far. This regards the use of various methodologies, inputs, outputs and other variables and the resulting outcomes of the studies. It will be assessed whether a clear definition of the airport technology can be identified. Further it will be reviewed what factors are mostly determined to contribute to inefficiency.

The paper is organized as follows; Section 2 will introduce the methodology of benchmarking and will review their application in the airport studies. Section 3 reviews the inputs and outputs that have been used as well as how they have been connected to an airport model. Section 4 continues with a presentation of the results of the studies. We will pay special attention to results on country comparison, the effects of privatization, commercialization and the hub status. Concluding remarks are given in section 5 on the issues that still remained unresolved and give recommendations for future research.

2 Productivity and Efficiency Measurement Concepts

The studies under review assess the airports productivity and efficiency with quantitative methods. This section will start with a definition and distinction of the terms productivity and efficiency and gives an overview of different techniques that can be applied. We will then continue with a brief introduction to the concepts of index-based approaches, DEA and SFA which have mostly often been used[3].

2.1 Productivity and Efficiency

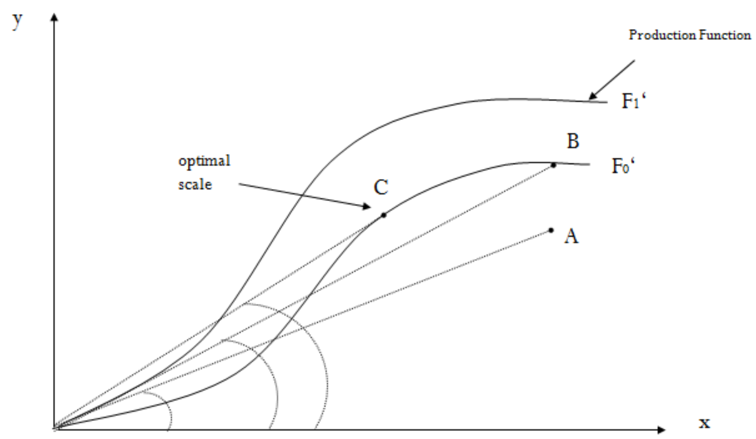
The productivity of an airport can be simply calculated as the ratio of output(s) per input(s). The division of one output by one input (e.g. labour productivity: passengers per employee) gives an indication of the partial productivity and aggregating all factors results in an overall measure given by the TFP. Further, the availability of panel data[4] allows measuring the technical change of airports over time. An upward shift of the production frontier from F_0' to

[3] We do not explain the whole methodological concept the approaches. For further details we suggest the following literature: Coelli et al. (2005) and Fried et al. (2008) for an overview for productivity and efficiency measurement, Kumbhakar and Lovell (2000) for SFA and Cooper et al. (2007) for DEA.

[4] Cross-sectional data is a one-dimensional set of variables including N firms in one year whereas panel data contains information on a set of firms over more than one time period.

F_1' as in Figure 1 can be caused by technical progress such as automated baggage handling systems (Coelli et al. 2005).

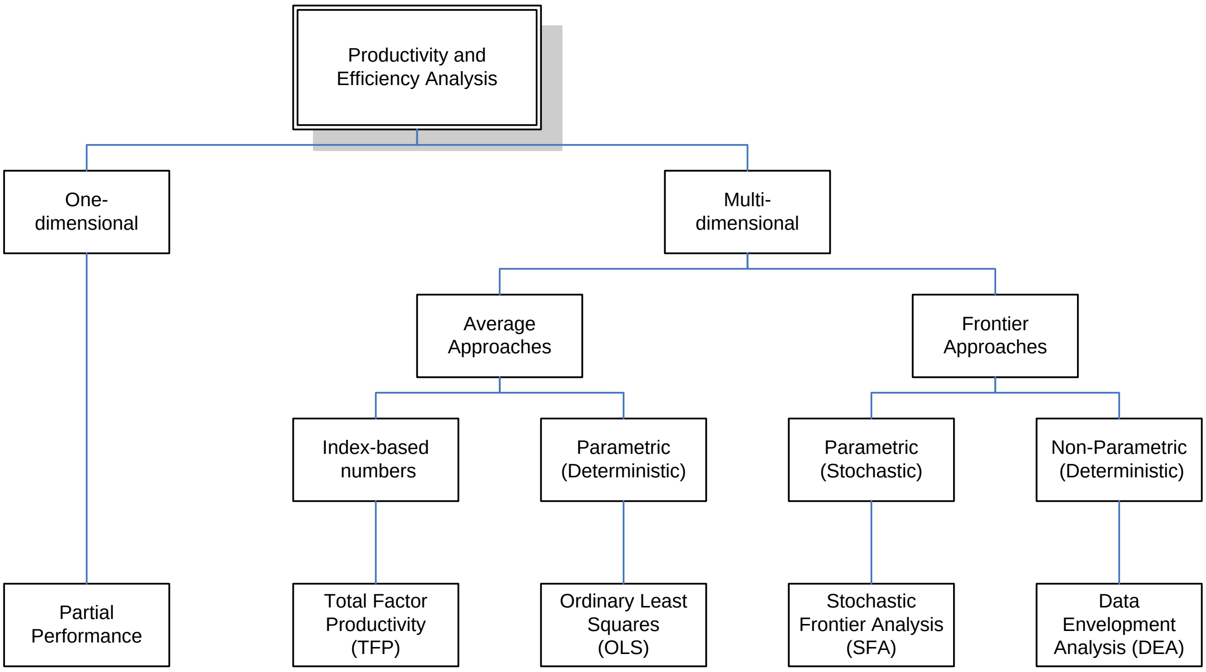
Figure 1: Productivity, Efficiency and Technical Change



Source: adapted from Coelli et al. (2005)

Technical efficiency defines the comparison of the observed outputs to its optimal values holding the inputs constant or vice versa. Farrell (1957) was the first to empirically measure the technical efficiency non-parametrically. His work was then taken up by Charnes, Cooper and Rhodes (1978) who developed DEA but also inspired econometricians such as Aigner and Chu (1968) and Aigner et al. (1977) to assess the technical efficiency with parametric techniques, either as a deterministic function or stochastically like SFA (Kumbhakar and Lovell 2000). The following Figure 2 illustrates an overview of quantitative methods that can be used to measure the productivity or efficiency of airports.

Figure 2: Quantitative Methods in Productivity and Efficiency Analysis



Source adapted from: von Hirschhausen and Cullmann (2006)

The simplest form is the one-dimensional approach computing partial productivities. However, this measure should be treated with caution and as discussed by Forsyth et al. (1986) partial measures should only be used if no data for overall measures is available. The Irish Commission for Aviation Regulation (2005) argued that partial productivity measures might be misleading as it does not capture substitution effects between inputs because differences in productivity might not reflect inefficiencies, but differences in relative factor prices. Furthermore, it cannot sufficiently account for different outsourcing strategies among airports. The second argument is demonstrated in the comparison of the two annual published studies by the Air Transport Research Society (ATRS) and the Transport Research Laboratory (TRL) or Doganis et al. (1995). While ATRS uses original data obtained from annual reports and other public sources in their section[5] on partial productivity assessment, TRL and Doganis et al. (1995) argue that airports differ in terms of the services that are supplied by an airport; a straightforward comparison with raw data might produce misleading results. For a like-with-like comparison, both studies defined core activities of an airport and adjust the original data accordingly by removing figures related of non-core activities such as ground handling and car parking from costs, expenses and staff numbers. As a result, the comparison of the labour productivity of ATRS and TRL studies lead to different results in the ranking as depicted in Table 1 (Kamp et al. 2007).

[5] The ATRS benchmarking report includes besides partial measures various other aspects of performance assessment such as variable factor productivity, unit and cost competitiveness and a review of airport fees and charges.

Table 1: Airport Labour Productivity in Passenger per Employee (2000) by rankings

		ATRS 2000	TRL 2000		
1	ARN	26.352	26.241	ARN	1
2	OSL	22.955	23.531	AMS	2
3	ZRH	22.249	22.627	ZRH	3
4	AMS	20.270	22.447	OSL	4
5	LGW	17.814	19.066	LHR	5
6	LHR	17.002	18.092	LGW	6
7	GVA	16.008	**18.032**	**MUC**	7
8	CPH	12.617	17.979	GVA	8
9	MAN	7.067	14.632	VIE	9
10	**MUC**	**5.714**	13.174	CPH	10
11	VIE	4.879	10.692	MAN	11
12	**FRA**	**3.459**	**8.050**	**FRA**	12

Source: ATRS and TRL, 2002 edition, taken from Kamp et al. (2007)

For this reason academic studies such as Graham and Holvad (2000) and Oum et al. (2003, 2004) use partial productivity measures only in addition to overall approaches such as DEA and TFP respectively.

To gain an overall measure of the airports performance, multi-dimensional approaches provide more accurate results. These can be distinguished between frontier and average approaches. A classical average approach is the ordinary least square (OLS) method where the sum of squared residuals has its least value and fits the model at best. Another average approach is the index-based number measurement of TFP. However in contrast to frontier approaches they assume that all decision making units (DMU) operate efficiently. This however appears to be unrealistic for airports as various factors beyond managerial control can have an effect on the efficiency. In frontier approaches an efficient production or cost frontier will be estimated, either parametrically with SFA or using linear programming in DEA. However, both require large datasets which are often not available (Coelli 2005).

2.2 Price-Based Index Number Approaches

The price-index number approach (PIN) is often used by statistical agencies to form the consumer price index (CPI) of a country. To assess the TFP information on market prices are required as weights to aggregate multiple inputs and outputs to an index. In contrast to frontier approaches it can already provide information on the productivity with only two observations. The Törnqvist index has widely been used in studies on the TFP. Different to the classical Laspeyres or Paasche indices it assumes the more flexible second-order translog technology. As the initial Törnqvist index only allows to measure productivity changes over time Caves, Christensen and Diewert (1982) proposed the multilateral translog index, also known as the CCD index to compare the TFP of different airports with cross-sectional or panel data:

$$
\begin{aligned}
\ln TFP_{kj} &= (\ln Y_k - \ln Y_j) - (\ln X_k - \ln X_j) \\
&= \frac{1}{2}\sum_i (R_{ik} + \overline{R}_i)(\ln Y_{ik} - \ln \overline{Y}_i) - \frac{1}{2}\sum_i (R_{ij} + \overline{R}_i)(\ln Y_{ij} - \ln \overline{Y}_i) \qquad (1) \\
&- \frac{1}{2}\sum_i (W_{ik} + \overline{W}_i)(\ln X_{ik} - \ln \overline{X}_i) - \frac{1}{2}\sum_i (W_{ij} + \overline{W}_i)(\ln X_{ij} - \ln \overline{X}_i)
\end{aligned}
$$

where Y_{ik} and R_{ik} are the output and the revenue share for output i of DMU k; $\overline{R}_i$ is the arithmetic mean of the revenue share and $\overline{Y}_i$ is the geometric mean of output i over the entire sample. X_{ik} are the input quantities and W_{ik} are the input cost shares for input i of DMU k; $\overline{W}_i$ is the arithmetic mean of cost shares and $\overline{X}_i$ the geometric mean of input i over the entire sample.

In assessing the productivity of airports only a few studies applied a price-based number approach. Nyshadham and Rao (2000), Hooper and Hensher (1997) and Vasigh and Gorjidooz (2006) applied the CCD index to measure the airport's TFP, whereas Oum and Yu (2004) and Oum et al. (2006) assessed the variable factor productivity (VFP) with the same index but excluded the capital input. As information of market prices is hardly available all studies used cost and revenues shares as weights to aggregate the inputs and outputs to indices.

However, price index number approaches do no estimate the true but unknown production technology and cannot account for inefficiency differences. Hence, the TFP cannot be decomposed into changes of pure technical efficiency, scale efficiency and the technology. To overcome these shortcomings, non-parametric linear programming and parametric approaches can be applied instead to measure the TFP. The most popular DEA based approach is the Malmquist index with distance functions (see section 2.3). A parametric approach to assess the TFP is the so-called Endogenous-Weight TFP. This has been applied by Yoshida (2004), Oum et al. (2003) and Yoshida and Fujimoto (2004) and estimates the production transformation function and inherently needs a specification of the functional relationship between inputs and outputs.

2.3 Data Envelopment Analysis

The majority of studies applied DEA. An advantage of this approach is that it can incorporate multiple inputs and outputs without input prices and does not need to specify a functional form of the production or cost frontier. Instead it is uses linear programming to construct a piece-wise linear frontier which is determined by the efficient airports of the sample. All relative inefficient units are enveloped within the efficient frontier. The mathematical framework of DEA has first been proposed by Charnes, Cooper and Rhodes (1978) to assume constant returns to scale (CRS) and has been extended by Banker, Charnes and Cooper (1984) to include variable returns to scale (Coelli 2005). Both models are also known as the CCR and BCC model respectively.

The basic envelopment form of the input-oriented BCC model evaluates the relative efficiency of DMU_o as in the following:

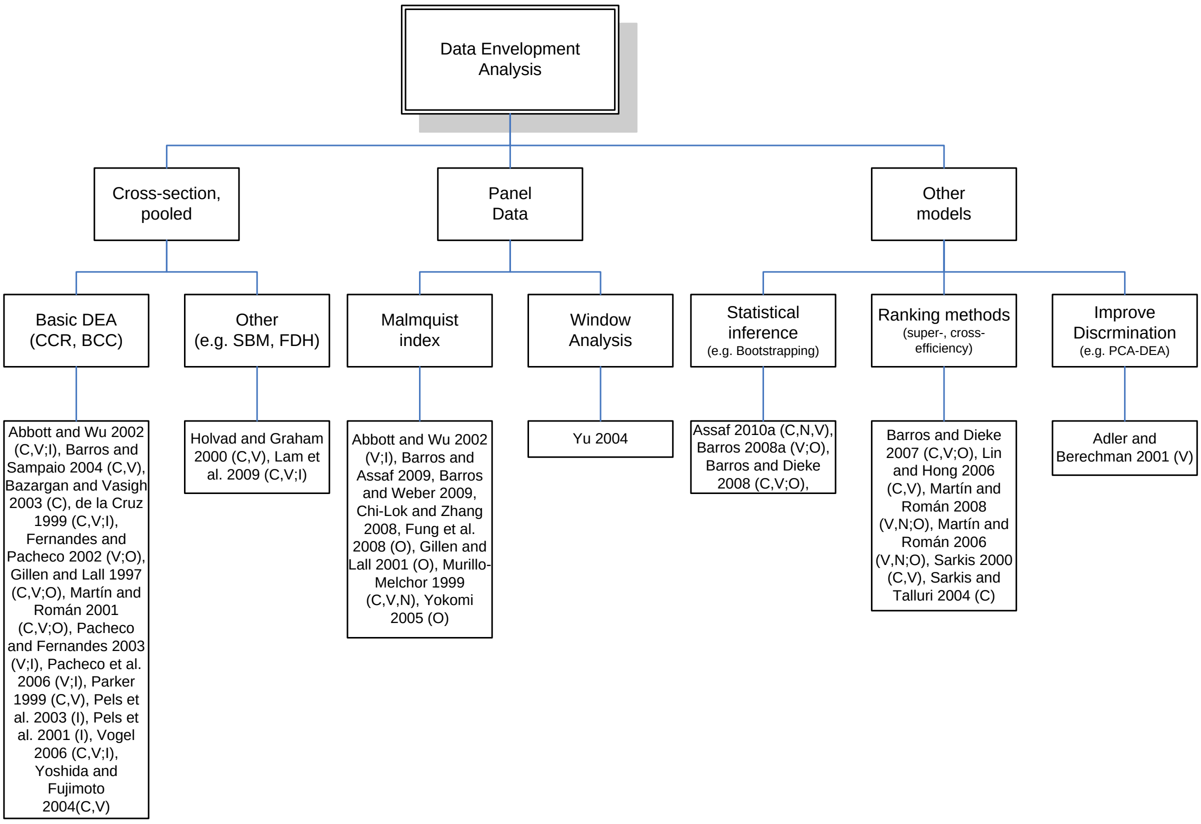
$$
\begin{aligned}
&\min_{\theta,\lambda} \theta \\
s.t.\ &\theta x_o - X\lambda \geq 0 \\
&Y\lambda \geq y_o \\
&e\lambda = 1 \\
&\lambda \geq 0
\end{aligned}
\qquad (2)
$$

where θ is a scalar that indicates the radial contraction of all inputs, hence the efficiency score. λ is a non-negative vector of weights that are determined by the optimization process and x_o and y_o are the input and output value of DMU_o the airport under investigation. X and Y represent the output matrices.

Over the years, the basic model has been further developed. It includes non-radial models such as the additive form, the consideration of non-discretionary variables, the estimation of efficiency changes over time, the improvement of discriminating efficiency estimates, the identification of outliers or introducing statistical inference into DEA (Cooper et al. 2007). The following Figure 3 presents the various DEA applications that have been used in airport efficiency studies.

Figure 3: Models in Data Envelopment Analysis[6]

[6] C= constant returns to scale, V= variable returns to scale, N= non-increasing returns to scale, I= input minimization, O= output maximization



Source: own illustration

The majority of studies assumed variable returns to scale which is mostly explained with a heterogeneous dataset on the airport size. Some studies also applied both scale options to assess the scale efficiency (e.g. Assaf 2010a; de la Cruz 1999). Mostly, early studies focus on the application of basic DEA with cross-sectional or pooled data as a sufficient panel structure was often not available. However, the relation between a low number of observations and a high number of inputs and outputs may lead to a large amount of efficient airports. As an example Parker (1999) assessed the technical efficiency of the British Airports Authority (BAA) as a single unit over 17 years using three inputs and two outputs. He received 9 of 17 years to be efficient with an average efficiency score of 96%. The opposite case could be observed in the study by Vogel (2006), to our concern he only included total costs and total revenues as input and output to assess the overall financial efficiency which would actually drop to a ratio.

To further rank efficient airports, Andersen and Petersen (1993) developed the super-efficiency model where "specialized" DMUs receive excessively high ranking. This model can also be used to identify outliers and remove them from the dataset. An alternative is the cross-efficiency model developed by Sexton et al. (1986) and extended by Doyle and Green (1994). Several airport studies have applied one of the models to rank efficient airports such as Adler and Berechman (2001), Sarkis and Talluri (2004) and Barros and Dieke (2007).

To overcome the shortcoming of receiving too many efficient airports Bazargan and Vasigh (2003), Lam et al. (2009) and Martín and Román (2006) included a virtual efficient airport which possesses the lowest input and highest output values of the sample. Another approach is principal component analysis (PCA) combined with DEA. It is applied to replace the original inputs and/or outputs with a smaller group of principle components (PCs), which explain the variance structure of a matrix of data through linear combinations of variables with minimal information loss (Adler and Golany 2001, 2002). Adler and Berechman (2001) applied PCA-DEA to reduce five outputs to three PCs which explain more than 80% of the variance in the original data over a cross-sectional sample of 26 airports.

The availability of a sufficient panel structure allows the assessment of productivity and efficiency changes over time. Malmquist DEA has already early been applied by e.g. Murillo-Melchor (1999) and Gillen and Lall (2001) on Spanish and US airports respectively, where the first assessed a rather short panel between 1992 and 1994. Most studies found positive productivity and efficiency changes over time (e.g. Abbott and Wu 2002; Chi-Lok and Zhang 2008; Fung et al. 2008; Yokomi 2005). This is not surprising as they used the traffic volume as outputs and physical data as inputs; the latter having remained fairly constant over time if no capacity expansion took place. Barros and Weber (2009) and Murillo-Melchor (1999) found decreases in TFP over the review period for UK and Spanish airports. Different to the other studies they selected cost information as input which might have increased disproportionately high to the passengers (pax), cargo and air transport movements (ATM).

Compared with parametric approaches DEA has its limitations of not allowing for hypothesis tests by itself. Bootstrapping, a re-sampling technique developed by Efron (1979) and firstly applied to DEA by Simar and Wilson (1998, 2000) can be used for statistical inference and correct the efficiency estimates from biasness. Just recently, bootstrapping has been applied to airport benchmarking studies by Assaf (2010a) and Barros and Assaf (2009). However, this approach needs to be treated with caution. As stated by Simar and Wilson (2000) the higher the number of variables to the number of observations the lower the ratios of convergence the bootstrapping provides. In their experiment with one input and one output the inclusion of 100 DMUs leads to almost exact confidence intervals, ten observations in contrast were far too less (Simar and Wilson 2000).

2.4 Stochastic Frontier Analysis

An advantage of SFA over DEA is that it does not only explain inefficiency due to mismanagement but also incorporates a stochastic random error which accounts for noise, i.e.

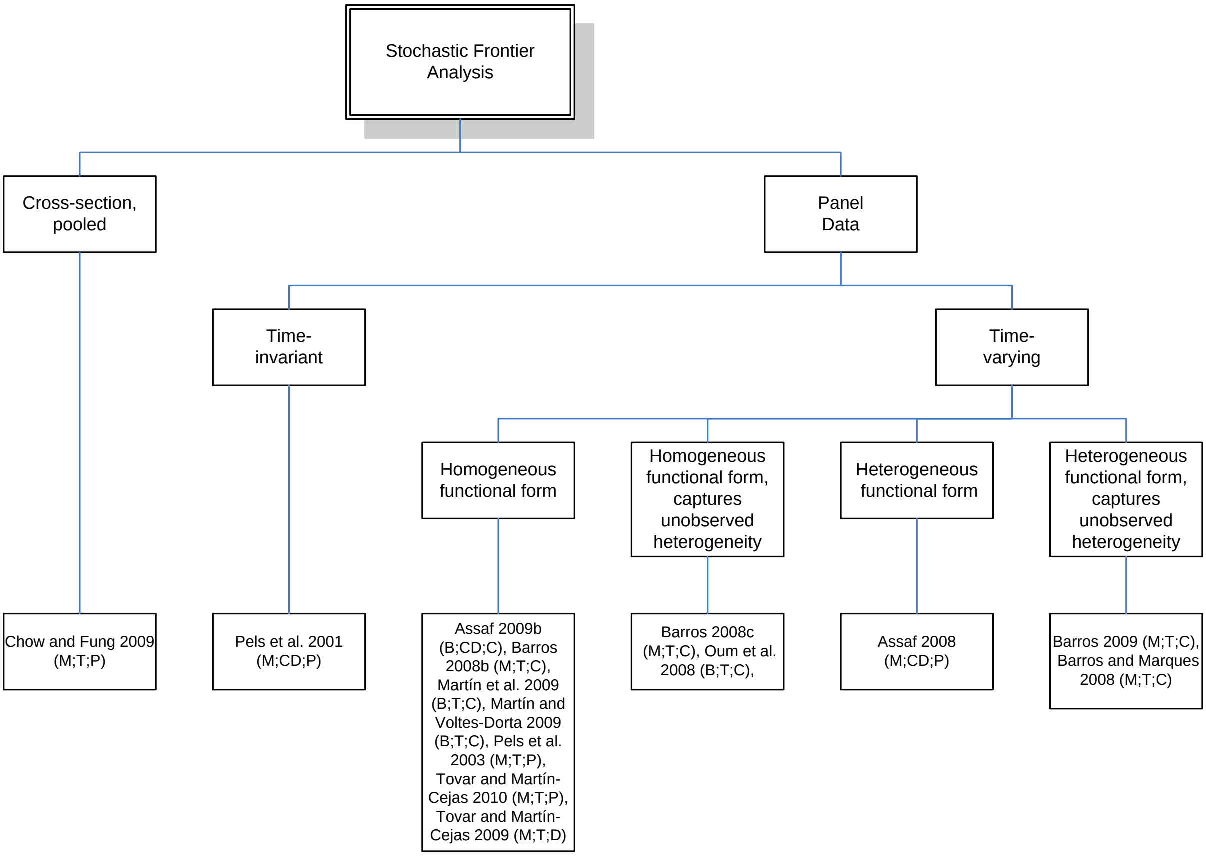
a not observable relation between the inputs and the output. This parametric frontier approach was first independently proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977):

$$\ln(y_i) = x_i'\beta + v_i - u_i \quad (3)$$

where the scalar $ln(y_i)$ is the observed output, x_i a vector of inputs and β a vector of technology parameters to be estimated. v_i is the stochastic random error and u_i the term for technical inefficiency indicating mismanagement.

Over the years the initial model has been further extended by considering panel data or by capturing unobserved and observed heterogeneity to improve the significance. The latter can enter either the production technology or the inefficiency term and will be discussed in section 4.3 which deals with the explanation of efficiency. The following Figure 4 presents the various models that have been applied in the airport sector.

Figure 4: Models in Stochastic Frontier Analysis[7]



Source: own illustration

[7] M= maximum likelihood estimation, B= Bayesian estimation, CD= Cobb-Douglas, T= Translog, C= cost function, P= production function, D= distance function

Although SFA has rarely been used in early years, it received increasing importance since 2007. Whereas early studies primarily estimated a production function with physical data (Pels et al. 2001, 2003) most recent studies estimated a translog cost function (e.g. Martín and Voltes-Dorta 2007; Barros 2008a/b; Oum et al. 2008) where the price for labour and capital has mostly been constructed from the wage rate and accounting ratios. Although the Cobb-Douglas function is simpler in its estimation it has its drawbacks that it assumes fixed returns to scale and an elasticity of substitution that equals unity. It has nevertheless been applied by Assaf (2008, 2010b) who tested the adequateness of a Cobb-Douglas and translog function with hypothesis testing and found the Cobb-Douglas to be preferable.

In the studies under review only Chow et al. (2009) used cross-sectional data with only 46 observations on Chinese airports. This small sample size might certainly weaken the model as SFA needs a large number of observations to obtain significant results. Indeed, in their study not many parameters were found to be significant. Panel data models were firstly developed by Sickles and Schmidt (1984) and Pitt and Lee (1981) as the fixed effects and random effects model respectively. Both assume the inefficiency to be time-invariant and were applied by Pels et al. (2001). To relax this rather restrictive, Cornwell et al. (1990) and Battese and Coelli (1992, 1995) have introduced time-varying inefficiency models for the fixed and random effects model respectively where the models by Battese and Coelli have been applied by the majority of studies (e.g. Assaf 2010b; Pels et al. 2003). To also capture cross-firm heterogeneity which is not related to technical inefficiency Greene (2005) introduced the true fixed and true random effects models. Both models shift time-invariant effects to unobserved heterogeneity whereas the inefficiency term varies over time. These rather recent developments have been applied by Barros (2008c, 2009) and Oum et al. (2008).

Further developments regard the assumption of a heterogeneous production or cost frontier due to observed heterogeneity and was applied by Assaf (2008), Barros (2009) and Barros and Marques (2008). They argued that airports operate under different conditions as with respect to the airport size, hence different technologies should be assumed for the airports instead. These approaches are helpful to explain inefficiencies across airports and will be explained in section 4.3.

2.5 Choice of Technique

Having reviewed the different methodologies it appears that the decision for a certain technique still remains difficult as they significantly differ from each other (see Table 2). DEA and SFA as frontier approaches are accurate as they both do not assume that all airports operate efficiently. Further, both approaches do not necessarily need input prices to aggregate multiple inputs and outputs. They only become important for the estimation of cost functions or assessing the allocative efficiency. However, both methods require large datasets in order to obtain robust efficiency estimates. In addition, especially DEA is very sensitive to outliers as the efficient frontier will be constructed by the data and using SFA requires prior information on the functional relationship between inputs and outputs. However, different to DEA, SFA does not assume that all deviations from the efficient frontier are attributed to managerial inefficiency but instead allows for a random noise that captures measurement errors. Using benchmarking as a managerial tool, DEA provides additional information on composite benchmarks to reach the target inputs or outputs but it cannot offer a complete ranking of the airports. With price index number approaches in contrast, a full ranking can be provided and only two observations, either in cross-sectional or time-series dimension, are necessary to estimate unbiased productivities. However, this method requires information on market prices to weight the inputs and outputs.

Table 2: Comparison of DEA, SFA and PIN Properties

Category	DEA	SFA	PIN
Method and assumptions			
Parametric approach	No	Yes	No
Assumes that all airports are efficient	No	No	Yes
Accounts for noise (e.g. weather, strike)	No	Yes	No
Types of measurement:			
- Technical efficiency	Yes	Yes	No
- Allocative efficiency	Yes	Yes	No
- Technical change	Yes	Yes	No
- Scale effects	Yes	Yes	No
- TFP change	Yes	Yes	Yes
Data Collection			
Data that can be used:			
- Cross sectional	Yes	Yes	Yes
- Time series	No	No	Yes
- Panel	Yes	Yes	Yes
Robust estimates with small number of observations	Yes	Yes	No
Basic method requires data on:			
- Input and output quantities	Yes	Yes	Yes
- Input and output prices	No[8]	No[9]	Yes
Testing hypotheses			
Conventional hypothesis testing	No[10]	Yes	No

Source: adapted from Coelli et al. (1998; 2003; 2005)

As a result the German regulator for the electricity sector, Bundesnetzagentur, applies both DEA and SFA when assessing the efficiency in order to ensure the reliability and verifiability of results (Reinhold et al. 2008). However, based on empirical studies that have compared the efficiencies estimates from DEA and SFA models were far from being conclusive (see e.g.

[8] Only if the allocative efficiency is estimated.
[9] Only if a cost function will be estimated.
[10] Unless applying the bootstrapping techniques proposed by Simar and Wilson (2000).

Ferrier and Lovell 1990 or Bauer et al. 1998 for the banking industry). Although the overall picture seems to be similar it is still not the same. In summary, the correct choice of the methodology depends on various criteria which include the objective of the study and the underlying assumptions or prior knowledge of the airport technology, the sample size and form (cross-section, time-series or panel) but also the available data (Coelli et al. 2003). The following section will continue with a review on the choice of inputs and outputs in airport benchmarking studies and how they will be connected to an airport production technology.

3 Productivity and Efficiency of Airports: Technology Specifications

Today, airports are very complex enterprises. Traditionally they were publicly owned and primary provided and operated the infrastructure for passengers and cargo handling, i.e. runways, terminals and apron. Hence the objective of an airport was to offer a good level of service irrespective of commercial and financial purposes. However as a result of the airline liberalization in the late Seventies and Eighties some airports might have become more competitive and underwent structural changes. Some airports became privatized, others who remained public increased their degree of commercialization and benchmarking gained more importance.

As stated by Coelli et al. (2003) "*Irrespective of which methodology […] -PIN, SFA or DEA- it cannot avoid the first rule for empirical economics: garbage in = garbage out*" (p.83). The data included in the model should clearly explain the underlying airport production technology. There is a consensus in the majority of studies that capital, employees, material and other external services are necessary to produce traffic volume. Depending on the data availability and the objective of the study physical or financial data will be collected.

On the output side, typical physical data are the number of passengers, cargo volume and the number of air transport movements. Various studies (e.g. Martín et al. 2009; Martín-Cejas 1999) combined passengers and cargo to work load units[11] (WLU) but we would argue that passengers and cargo generate different costs and revenues and should therefore not be seen as an equal output to the airport. Non-aeronautical activities are mostly given in revenues as a physical measure might be difficult unless an appropriate quantity index can be created. So far, especially early studies (e.g. Fernandes and Pacheco 2002; Gillen and Lall 1997; Pels et al. 2001) but also some recent publications (e.g. Barros 2009; Chi-Lok and Zhang 2008; Fung et al. 2008) did not consider non-aeronautical activities at an airport but only assessed the

[11] One WLU equals one passenger or 100kg cargo.

technical efficiency of the airside with physical data. However, studies ignoring the non-aeronautical side might cause biased results as the inclusion of staff and capital input would then require the split between airside and non-aeronautical activities which is to our concern not the case in any study. Today, in an increasingly competitive environment where airports are still subject to economic regulation an airport emphasizes to generate additional revenues from non-aeronautical activities to attract airlines with lower charges. Therefore, especially recent studies (e.g. Barros and Dieke 2007; Oum et al. 2006, 2008; Pacheco et al. 2006) view the airport as a production technology with both operating activities, including the non-aeronautical side as revenues.

In the recent past it was not only of interest to include desirable outputs in the studies but also undesirable outcome as they also contribute to the performance of an airport. Pathomsiri et al (2008) analysed US airports included delays as a negative output. Their results clearly indicated that ignoring the quality of airport services would otherwise overestimate the technical efficiency gains of airports from higher utilization. Unfortunately, so far this method has not been applied to analyse airports in Europe and other countries where slot allocation rations excess demand with less but still substantial levels of delays.

Yu (2004) included aircraft noise as an undesirable output having studies Taiwanese airports. He found that undesirable outputs severely affect the technical efficiency of airports. As with congestion one would expect that including noise and other externalities will change the efficiency scores of those airports located in highly populated areas in the US and Europe.

On the input side, physical data are usually information on staff as operating inputs and gates, the terminal size or the runway length as capital inputs. Material and other outsourced services are normally measured in financial terms unless constructing a quantity index (e.g. Barros and Dieke 2008c; Hooper and Hensher 1997; Oum et al. 2003, 2006, 2008). Therefore, studies that only considered physical data did not include information on outsourcing strategies (e.g. Abbott and Wu 2002; Assaf 2010; Chi-Lok and Zhang 2009).

The inclusion of staff appeared to be very controversial among the studies. The airport is a very heterogeneous industry regarding the vertical integration, especially with respect to the labour intensive ground handling as discussed in Kamp et al. (2007). For this reason, Pels et al. (2001, 2003), Yoshida (2004) and Fung et al. (2008) did not use staff in their model to avoid an apple-with-pear comparison and assuming labour and capital to be perfect complements. We would argue the latter argument as for example automated baggage handling systems are substitutes rather than complementary to workers. The majority of studies included the number of employees which is mostly reported in annual reports or other open sources (e.g. Gillen and Lall 1997, 2001; Parker 1999; Murillo-Melchor 1999; Abbott and Wu 2002). This figure however cannot distinguish between full-time and part-time employees. The studies of Oum et al. or Assaf (2010a/b) collected information on full-time equivalents but this figure is often publicly not available. Staff costs were normally used in combination with staff quantities for the estimation of a parametric cost function or to assess the allocative efficiency in DEA (e.g. Assaf 2010a; Barros and Sampaio 2004; Oum et al. 2008) or to differentiate between unskilled, skilled and management staff (e.g. Hooper and Hensher 1997; Martín and Román 2001, 2006, 2008). Nevertheless, it should be satisfied that staff costs are adequately adjusted for different staff costs level across countries or regions.

Further, the consideration of capital input varies among the studies and seems to be a difficult task in airport benchmarking. Due to data and comparability problems as discussed by Oum et al. (2004) they have ignored the capital input and assessed the variable factor productivity instead. Most studies that consider capital included physical data (e.g. Barros 2008a; Gillen and Lall 1997; Pels et al. 2003; Oum et al. 2003) which seems to be for cross-border studies more appropriate to avoid the comparison of different national accounting procedures. Thus, studies that included monetary values of the physical capital were basically national studies. A classical approach is to take the book value of physical assets (e.g. Assaf 2008; Barros and Weber 2009; Barros and Sampaio 2004). However, the book value only reflects the value of the assets which has not been depreciated so far, therefore a comparison between new and mature airports will cause some bias. To consider the monetary value of physical capital Abbott and Wu (2002) and Hooper and Hensher (1997) applied the perpetual inventory method (PIM) for Australian airports. Collecting the data of past capital investments as well as information on the expected useful life is however very time-consuming, especially for cross-border studies.

As discussed by Kamp et al. (2007) the degree of outsourcing especially with respect to the labour-intensive activities such as ground handling severely affects the efficiency of an airport. Whereas in the US, UK or Australia ground handling has traditionally been operated by airlines or independent third-parties the airports in Germany and Austria provide this activity in-house. As a result, German airports automatically receive lower efficiency scores in input-oriented models if the airport staff will be compared against traffic volume as tested by Adler et al. (2009). To prevent an imbalance in international studies the ground handling activity should be considered on the input and output side. To our knowledge, all studies by Oum et al. were aware of this heterogeneous character of airports and included revenues from ground handling activities in their commercial revenues.

In DEA studies, researcher need to decide whether to maximize the outputs given the inputs or vice versa. This question can only be answered by analyzing the actual behaviour incentives of the airports under review. For example, an airport in a competitive environment with no capacity constraints will probably seek to maximize the traffic volume. This might also be true for public airports trying to maximize the connectivity for regional development. In contrast, monopolistic and unregulated airports are likely to maximize their revenues whereas monopolistic airports that are regulated by a revenue-cap might try to minimize their costs. In airport benchmarking the majority of studies assume an airport to maximize their outputs with given inputs (e.g. Barros 2008a; Fernandes and Pacheco 2002; Gillen and Lall 1997; Martín and Román 2001) as the capital input cannot be easily adjusted in the short-term. To overcome this problem, Pels et al. (2003) treated the number of runways as a fixed (non-discretionary) input in his input-oriented DEA model.

Having now reviewed the variables that have been used to define the airport technology it is also important how these have been connected. This regards particularly DEA studies which do not require prior assumptions on the functional relationship between inputs and outputs. As discussed, many studies considered estimated an aggregate model of the airside including pax, cargo and ATM but ignoring non-aeronautical activities on the output side. To differentiate between different aeronautical activities, Gillen and Lall (1997, 2001) followed by Pels et al. (2001, 2003) and Barros and Assaf (2009) applied a DEA model separating them with respect to passengers and movements due to different assumptions on the returns to scale. Whereas constant returns to scale were assumed for the airside producing movements, it is expected to exhibit variable returns to scale with respect to passengers on the terminal side. Indeed, Pels et al (2003) found different scale effects for the two operational sides and Gillen and Lall (1997) as well as Barros and Assaf (2009) further identified productivity growth on the terminal side and a decline on the airside.

So far, the overview of the measuring techniques and the choice of inputs and outputs in past studies indicate a great variability and no concluding answer on a commonly defined airport model. In the following section we will highlight the findings of the studies and will compare the results.

4 Productivity and Efficiency of Airports: Results

Studies that assess the airport's performance try to find reasons for inefficiency or productivity changes. Inefficiency can for example be explained with diseconomies of scale, technical or allocative inefficiency or externalities such as delay, noise or other factors that are beyond managerial control whereas productivity changes might be caused by innovations as the production frontier will be shifted outwards (Oum et al. 1992). The following chapter will present the findings of airport benchmarking studies with respect to country comparisons, economies of scale and how and by what factors the inefficiency can be explained.

4.1 Productivity and Efficiency Measurement

When reviewing studies with respect to productivity and efficiency differences within and across countries we found concurrent results but also some open questions. Only the studies by Chi-Lok and Zhang (2008) and Fung et al. (2008) reached similar results on the Chinese market proving a positive TFP growth. This is however not surprising; both studies reviewed a similar time period between 1994 and 2006 using basic DEA and Malmquist DEA as well as the same inputs and outputs.

The studies by Tovar and Martín-Cejas (2010) and Murillo-Melchor (1999) reached the same conclusions on Spanish airports on the years 1993/94. The first estimated a distance function with SFA and decomposed the TFP change into its components for the years 1993-1999. They found especially for hub airports positive TFP changes, which were mainly contributed by technological changes but a decrease in TFP 1993 and 1995. The study by Murillo-Melchor (1999) applied Malmquist DEA for a three-year period between 1992 and 1994 and also proved a decrease in TFP over time. Martín et al. (2009) found slight technological progress at Spanish airports with SFA for a long-run cost function from 1991-1997 however, it is not clear if the progress was constant over time.

Several studies were conducted comparing the airports performance across countries. The studies by Abbott and Wu (2002) and Graham and Holvad (2000) used basic DEA for the years 1998/99 and 1992/93 respectively and found that Australian airports perform better than European or US airports. Lin and Hong (2006) in contrast found a better operational performance of European and US airports rather than Asia and Australia with DEA due to higher GDP rates but no significant difference between Europe and the US and Asia and Australia. Their time period included the years 2001/02. This again contradicts with the findings by Oum et al (2006) who concluded with a CCD index that Asian and European airports between 2001 and 2003 have a negative influence on the operating performance rather than airports in the US and Australia. Surprisingly, it appears that the often criticised US airports are more efficient than airports on other continents (Morrison and Winston 2008).

In contradiction to the studies above Pels et al. (2003) argued that when assessing the efficiency of European airports there are no region-specific effects as he found different efficiency scores for airports in London or Berlin with DEA. Hence, it can be argued that different techniques, time periods or different inputs and output combinations might not provide a clear picture on efficiency differences within and across countries.

4.2 Estimation of Returns to Scale

Benchmarking airports of different size can also raise the question of how to eliminate the effects of size for a multi- product ‘airport’ firm[12]. According to Graham (2004) after the airport reaches the size of about 3 to 5 million passengers, economies of scale effects flatten out, so that for benchmarking of medium and large sized airports the size does not matter. However, various benchmarking studies lead to different conclusions on the scale effects at airports.

Martín and Voltes-Dorta (2007) and Oum et al. (2006) assessed the performance of worldwide airports with SFA and the CCD index respectively. Surprisingly the first concluded that even Atlanta and Chicago as the two largest airports in the world are operating under increasing returns to scale. In other words, economies of scale will never flatten out at airports. Oum et al. (2006) in contrast found positive size effects which were not significant in all cases but still support the study by Jeong (2005) who concluded that economies of scales exhaust at 3 million passengers. Pels et al. (2003) applied DEA and SFA to the European market and found decreasing returns to scale from 12.5 million passengers for the airside but increasing returns on the terminal side for an average airport. Similar conclusions were reached by de la Cruz (1999) with DEA on Spanish airports in an aggregated model considering aeronautical and non-aeronautical activities together. Murillo-Melchor (1999) found increasing returns to scale for smaller Spanish airports but constant or decreasing returns for larger airports which in turn contradict the findings of Martín et al. (2009) who found that all Spanish airports operate at increasing returns to scale. Further Barros and Sampaio (2004) argued that for Portuguese airports scale economies are less important than location and agglomeration.

4.3 Explaining inefficiency with exogenous factors

An important issue in efficiency assessment is to find answers on performance differences between airports. On the one hand, these can occur due to managerial decisions but very often

[12] The studies often receive information on the existence of economies of scale as a by-product but do not primarily focus on economies of scale. There are various studies exclusively focussing on economies of scale which are not benchmarking studies. These will not be considered here as this is not the primer focus of this paper. For more information we refer to Kamp et al. (2005).

factors beyond managerial control can have an effect on the airports efficiency. Frontier methods offer various solutions to accommodate these effects in an assessment.

The majority of DEA studies applied a second stage regression. Two-stage approaches use econometrics to regress the environmental variables against the DEA efficiency estimates, mostly using the censored Tobit regression (e.g. Abbott and Wu 2002; Barros and Sampaio 2004; Gillen and Lall 1997) or truncated regression that includes bootstrapping to account for statistical inference as applied by Barros (2008a) and Barros and Dieke (2008). As the environmental variables are not included in the model, their consideration does not affect the discriminatory power of the DEA model, however as with all parametric estimates it requires a specification of the functional form.

Non-parametric possibilities such as Mann-Whitney and Kruskal-Wallis tests assess the significance of efficiency differences of various groups. Bazarghan and Vasigh (2003) used both to assess efficiency differences between public and private airports and Graham and Holvad (2000) applied a Mann-Whitney test on Australian and European airports.

Still in its early stages is the consideration of observed heterogeneity in SFA. In contrast to unobservable heterogeneity, this form of heterogeneity is reflected in measured variables. An important question in parametric research studies is to define if the factors affect the production technology or the inefficiency term? Early studies that included environmental factors in the model assume a homogenous production technology for all airports such as by Battese and Coelli (1995) where observed heterogeneity has an effect on the efficiency of firms (e.g. Tovar and Martín-Cejas 2009; Chow et al. 2009). The heterogeneous models that account for different production technologies across airports have recently been developed. Assaf (2008) used a meta-frontier approach which accounts for technological differences between heterogeneous groups, in his case small and large airports in the UK. Similarly a study by Barros (2009) applied the latent class model on British airports as introduced by Orea and Kumbhakar (2004) which first clusters the data into certain groups and then run the models separately. Barros (2009) defined the classes according to the market share indicating the airport size.

Reviewing past studies many different factors were defined that contribute to efficiency differences across airports (see Table 3).

Table 3: Factors contributing to inefficiency

Airport Characteristics	Management Strategies	Governance Structure	Other Exogenous Effects
Hub Status *(e.g. Barros 2008a, Gillen and Lall 1997, Oum*	Commercialization *(e.g. Oum et al. 2003, 2006, Tovar and*	Economic Regulation *(Oum et al. 2004)*	Agglomeration *(Barros and Dieke 2007, Yu2004)*

et al. 2004)	*Martín-Cejas 2009)*		
Slot Coordination *(Pels et al. 2003)*	Noise Strategy *(Gillen and Lall 1997)*	Ownership *(e.g Barros and Dieke 2008, Chi-Lok and Zhang 2008, Fung et al. 2008, Oum et al. 2006, 2008, Vasigh and Gorjidooz)*	Economic Development *(Lin and Hong 2006, Parker 1999)*
Traffic Structure *[Chow et al. 2009 (int. traffic), Oum et al. 2006, 20008(cargo and int. traffic), Tovar and Martín-Cejas 2009 (cargo)]*	Outsourcing *(Tovar and Martín-Cejas 2009, Oum et al. 2003)*	Congestion *(Oum et al. 2004)*	Location *(Fung et al. 2008)*
	Service Quality *(Oum et al. 2003)*	Regional Competition and Market Share *(Barros and Sampaio 2004, Chi-Lok and Zhang 2008, Pathomsiri et al. 2008)*	

Source: own illustration

From this list of factors we will present the findings that have mostly been explained to affect the efficiency of airport, which are the ownership form, commercialization and the hub status.

Ownership

Like in other studies on the empirical evidence of ownership[13] the results are rather inconclusive if privatization increases the efficiency of airports. Additionally the emergence of mixed ownership forms complicates the debate on the effects of ownership as discussed by Boardman and Vining (1989).

The first study in this field was an analysis of the effects of privatization of the BAA airports. Parker (1999) used DEA to estimate the technical efficiency prior and after privatization (1979-1996). In his first model he assessed the technical efficiency of BAA as a single unit from 1979 to 1995 and in a second model he included 22 UK airports for the years 1988-1996 and separated the BAA into its single units. He found no evidence that privatization has improved the airport's performance and concluded that the golden share which is kept by the government does not induce enough capital market pressures. Further he argued that BAA is still subject to economic regulation and it is of question if incentives to operate more efficiently can be distorted by government regulation. In contrast, Yokomi (2005) reviewed the technical and efficiency change of 6 BAA airports from 1975 to 2001 with Malmquist DEA. Different to Parker she found that the BAA airports have improved after their privatization exhibiting positive changes in technical efficiency and technology. In particular on the non-aeronautical side, the growth after privatization is much higher. Positive effects of privatization were also found by Barros and Marques (2009) and Barros and Dieke (2007), the first applying SFA to estimate a homogenous and heterogenous translog cost function using a

[13] Shirley and Walsh (2001) assessed the effects of privatization in empirical studies and the theoretical literature. They found stronger support for privatization in empirical studies but nevertheless the results are still far from being conclusive.

worldwide sample from 2003-2004, the latter solving DEA on 31 Italian airports between 2001 and 2003 and applying Mann-Whitney hypothesis tests.

Vasigh and Gorjidooz (2006) measured the effects of ownership on the airport's TFP with the multilateral index on a sample of 22 airports from the UK (7; private), other European countries (7; public-private) and the US (8; public) between 2000 and 2004. Having regressed the relationship between the TFP and the ownership, their study came to the result that there is no significant relationship for financial and operational efficiencies. The same was concluded by Lin and Hong (2006) on worldwide airports in 2001/2002 using DEA to assess the operating efficiency. Their hypothesis testing leaded to no correlation between the ownership form and the performance.

Different studies by Oum lead to different results on the effects of privatization. Whereas Oum et al. (2003) concluded that the ownership structure has no statistically significant effect on the airport's productivity with data of 50 major airports worldwide from 1999 and conducting a second-stage regression, Oum et al. (2006) and Oum et al. (2008) were in favour of privatization. Both latter studies included more than 100 airports worldwide and in contrast to all other studies do not only distinguish between privatized and public but also consider the governance structure. Oum et al. (2006) assessed the impact of different governance structures using VFP and regression to obtain the residual VFP and reached the conclusion that a public corporation is not statistically different from major private airports. However, airports that are major publicly owned or have multiple government involvement seem to operate significantly less efficient from the other airports. Oum et al. (2008) estimated a heterogeneous translog cost function with SFA to measure the cost efficiency between 2001 and 2004. In this study the authors concluded that airports with a major private shareholder are more efficient than public airports or airports with major public influence.

Again, the DEA results by Vogel (2006) on the financial performance of European airports conflict the outcome of Oum et al. Having reviewed the period between 1990 and 1999 he indicated that partially and fully privatized airports outperform its public counterparts, thereof fully privatized are more efficient than PPPs. Further, in a second stage regression proved significance of a positive correlation of efficiency estimates and private ownership. This has again been reconfirmed in a second regression mode which assessed the effects of changes in from public toward private ownership.

<u>Hub effect</u>

On the effect of being a hub airport the majority of studies found positive effects (Sarkis

2000; Gillen and Lall 1997; Barros 2008a; Chow et al. 2009; Fung et al. 2008; Martín and Román 2008). Sarkis (2000) applied various DEA models and argued on US airports between 1990 and 1994 that especially major hubs are more efficient as since the liberalization flag carriers tend to chose more efficient airports as hubs or consolidate operations at their major hubs in more recent years. Also Gillen and Lall (1997) assessed the US market with DEA between 1989 and 1993. As they separated the terminal and airside they found general higher efficiency on the terminal side for hub airports but especially large gateway hubs such as Atlanta perform better on the airside. This has been supported by the studies of Fung et al. (2008) on Chinese airports and by Martín and Román (2008) on Spanish airports. Both also found especially large hubs to outperform the others with DEA. Different to the studies above, Pathomsiri et al. (2008) and Bazargan and Vasigh (2003) found that small airports operate more efficiently.

In contrast to these results Oum et al (2004) concluded that a hub status reduces the TFP of an airport. They argue that airlines create peak problems artificially by running highly intensified flight departure and arrival flight banks which make airports over-utilized during peaks and under-utilized during off-peaks. However, with respect to measuring the capital input productivity this factor was found to be insignificant therefore its effects less clear.

Except for Pathomsiri et al. (2008) all other studies do not consider delay as an undesirable output or account for capacity constraints which could be rationed with queuing or an effective slot allocation. We would argue that the consideration of these factors will severely affect the results and should certainly be included in the model.

Commercialization

Since the liberalization of the airline industry also the airports faced structural changes as the industry became more competitive. Some airports privatized whereas other stayed public but shifted their focus on non-aeronautical activities. The intention was often to lower charges and to attract airlines (Oum et al 2006). In airport benchmarking several studies tested for the effects of commercialization on the airports productivity and efficiency. These include studies by Oum et al. (2003, 2004, 2006) and Oum and Yu (2004) on worldwide airports and Tovar and Martín-Cejas (2009) on Spanish airports. They all found demand complementarities between aeronautical and non-aeronautical activities which positively affect the airport's performance. In this case, although various techniques (CCD index, EW-TFP and SFA), airport samples and variables have been selected, they all reached the same conclusion which might indicate a certain level of robustness on the result. The studies point out to strong economies of scope between different activities. However, unlike for example the airline industry the strengths of economies of scope have not been assessed so far.

5 Conclusion and Directions for Future Research

Since the late Nineties several studies on the productivity and efficiency of airports emerged that applied quantitative methods such as DEA, SFA or PIN. PIN has only been used in a small number of studies which could be explained with the non-availability of price information or the general assumption of inefficiency across airports. Instead, efficiency techniques have been preferred. The majority of studies applied DEA to assess the technical efficiency but there have been an increasing number of studies who recently applied SFA. DEA as being a non-parametric approach requires fewer assumptions than SFA and can easily handle multiple inputs and output. However as with SFA, DEA can cause biased results if not carefully assessed. Over the time, both DEA and SFA have become more sophisticated approaches. Whereas DEA provides for example possibilities to open up the black box with network DEA or introducing statistical inference via bootstrapping, researchers focussed on the consideration of unobserved and observed heterogeneity regarding SFA. To consider the often discussed heterogeneous character of airports especially Assaf and Barros made use of recently developed SFA models that account for heterogeneous airport production technologies. However, what is still not fully developed is to account for time lags when assessing the dynamic efficiency. The airport industry is typical for being capital intensive with lumpy investments and technological progress over time. For future studies we suggest to also account for time-lags between capital investments and the optimal use when assessing the dynamic efficiency of airports.

Having reviewed the majority of airport benchmarking studies we could not find a consensus airport production technology which defines the relevant inputs and outputs nor their connection with each other. The debate regards the measurement of labour, capital and the inclusion of material and outsourcing on the input side. On the output side various studies only consider aeronautical activities although the commercial side is gaining increasing importance. Others combined passengers and cargo to WLU as a single measure which we doubt should be treated equally. For future studies we suggest to include the relevant inputs and outputs that explain the airport production technology if the data is available. This especially regards a meaningful measure of capital and the consideration of undesirable outputs which clearly affect the performance of an industry that is influenced by externalities.

The results of various studies still lead to open questions with respect to the airport's optimal governance structure or the most productive scale size. Many studies aimed to explain inefficiency across airports, mostly due to different ownership structures, being a hub airport or increasing commercialization. Whereas we found positive effects of commercialization by all studies, the results on ownership seems to be rather inconclusive with studies in favour of privatization and others that concluded no effects of ownership. However, we argue that the effects of ownership should not be separated from economic regulation and competition as stated by Vickers and Yarrow (1988, 1991). Instead privatization might not have any positive effects in the airports performance if they are cost-based regulated but could increase efficiency if airports are effectively price-cap regulated or facing effective competition. Thus, future studies should account for all three factors as a combined effect i.e. to assess the most efficient regulatory and ownership form with and without regional and hub competition.

To improve the use of airport benchmarking and performance measurement we argue that airport managers should contribute with their industry knowledge to improve the models. This can also increase the use of results obtained with overall measures for managers as they often argue to prefer partial measures. For policy makers benchmarking results can be the first step in decision processes to assess efficiency differences and find answers on the optimal governance structure. In summary, as also stated by Morrison (2009, p.157): "*[...]there should be more dialogue between airport managers, government agencies and research organizations towards developing a more balanced perspective that better aligns desired end use of benchmarking results to the data collected and methodology employed.*"

6 References

Abbott, M. & Wu, S., 2002. Total factor productivity and efficiency of Australian airports. *Australian Economic Review*, 35, 244–260.

Adler, N. & Berechman, J., 2001. Measuring airport quality from the airlines' viewpoint: an application of data envelopment analysis. *Transport Policy*, 8(3), 171–181.

Adler, N. & Golany, B., 2001. Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe. *European Journal of Operational Research*, 132(2), 260–273.

Adler, N. & Golany, B., 2002. Including principal component weights to improve discrimination in data envelopment analysis. *The Journal of the Operational Research Society*, 53(9), 985–991.

Adler, N., Liebert, V. & Yazhemsky, E. 2009. Can we really benchmark airports? Paper presented at the German aviation Research Society Workshop on Airport Benchmarking 2009, Berlin.

Adler, N., Oum, T.H. & Yu, C., 2009. A response to 'Understanding the complexities and challenges of airport performance benchmarking'. Journal of Airport Management, 3(2), 159–163.

Aigner, D., Lovell, C.A.K. & Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6, 21-37.

Aigner, D.J. & Chu, S.F., 1968. On estimating the industry production function. *The American Economic Review*, 58(4), 826–839.

Air Transport Research Society 2002. Airport Benchmarking Report – 2003: Global Standards for Airport Excellence. ATRS, Vancouver.

Andersen, P. & Petersen, N.C., 1993. A procedure for ranking efficient units in data envelopment analysis. *Management Science*, 39(10), 1261-1264.

Assaf, A., 2008. Accounting for size in efficiency comparisons of airports. *Journal of Air Transport Management*, 15(5), 256-258.

Assaf, A., 2010a. Bootstrapped scale efficiency measures of UK airports. *Journal of Air Transport Management,* 16(1), 42-44.

Assaf, A., 2010b. The cost efficiency of Australian airports post privatisation: A Bayesian methodology. *Tourism Management*, 31(2), 267-273.

Banker, R.D., Charnes, A. & Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 30(9), 1078–1092.

Barros, C.P., 2008a. Airports in Argentina: Technical efficiency in the context of an economic crisis. *Journal of Air Transport Management*, 14(6), 315–319.

Barros, C.P., 2008b. Technical change and productivity growth in airports: A case study. *Transportation Research Part A*, 42(5), 818–832.

Barros, C.P., 2008c. Technical efficiency of UK airports. *Journal of Air Transport Management*, 14(4), 175–178.

Barros, C.P., 2009. The Measurement of Efficiency of UK Airports, Using a Stochastic Latent Class Frontier Model. *Transport Reviews*, 29(4), 479-498.

Barros, C.P. & Assaf, A., 2009. Productivity Change in USA Airports: The Gillen and Lall Approach Revisited. Working Paper 22/2009/DE/UECE, School of Economics and Management, Technical University of Lisbon.

Barros, C.P. & Dieke, P.U., 2008. Measuring the economic efficiency of airports: A Simar–Wilson methodology analysis. *Transportation Research Part E*, 44(6), 1039–1051.

Barros, C.P. & Dieke, P.U., 2007. Performance evaluation of Italian airports: a data envelopment analysis. *Journal of Air Transport Management*, 13(4), 184–191.

Barros, C.P. & Marques, R.C., 2008. Performance of European Airports: Regulation, Ownership and Managerial Efficiency. Working Paper 25/2008/DE/UECE, School of Economics and Management, Technical University of Lisbon.

Barros, C.P. & Weber, W.L., 2009. Productivity growth and biased technological change in UK airports. *Transportation Research Part E*, 45(4), 642–653.

Barros, C.P. & Sampaio, A., 2004. Technical and allocative efficiency in airports. *International Journal of Transport Economics*, 31(3), 355–377.

Battese, G.E. & Coelli, T.J., 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical economics*, 20(2), 325-332.

Battese, 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(1), 153–169.

Bauer, P.W., Berger, A.N., Ferrier, G.D. & Humphrey, D.B., 1998. Consistency conditions for regulatory analysis of financial institutions: a comparison of frontier efficiency methods. *Journal of Economics and Business*, 50(2), 85–114.

Bazargan, M. & Vasigh, B., 2003. Size versus efficiency: a case study of US commercial airports. *Journal of Air Transport Management*, 9(3), 187–193.

Boardman, A.E. & Vining, A.R., 1989. Ownership and performance in competitive environments: A comparison of the performance of private, mixed, and state-owned enterprises. *Journal of Law and Economics*, 32(1), 1-33.

Caves, D.W., Christensen, L.R. & Diewert, W.E., 1982. Multilateral comparisons of output, input, and productivity using superlative index numbers. *The Economic Journal*, 92, 73–86.

Charnes, A., Cooper, W.W. & Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429-444.

Chi-Lok, A.Y. & Zhang, A., 2009. Effects of competition and policy changes on Chinese airport productivity: An empirical investigation. *Journal of Air Transport Management*, 15(4), 166–174.

Chow, W., Kong, C. & Fung, M.K., 2009. Efficiencies and scope economies of Chinese airports in moving passengers and cargo. *Journal of Air Transport Management*, 15(6), 324-329.

Coelli, T., Estache, A., Perelman, S. & Trujillo, L., 2003. *A primer on efficiency measurement for utilities and transport regulators*, World Bank.

Coelli, T., Rao, D.S.P. & Battese, G.E., 1998. *An introduction to efficiency and productivity analysis*, Kluwer Academic Publishers.

Coelli, T., Rao, D.S.P., O'Donnell, C.J. & Battese, G.E., 2005. *An introduction to efficiency and productivity analysis.* Second Edition, Springer, Berlin.

Commission for Aviation Regulation, 2005. *The Performance of Dublin Airport: The Findings of the comparative reports of the TRL and the ATRS, Commission for Aviation Regulation*, Dublin.

Cooper, W.W., Seiford, L.M. & Tone, K., 2007. *Data envelopment analysis: a comprehensive text with models, applications, references and DEA-solver software*, Second Edition, Springer Verlag, Berlin.

de Borger, B., Kerstens, K. & Costa, A., 2002. Public transit performance: What does one learn from frontier studies? *Transport Reviews*, 22(1), 1–38.

de la Cruz, S.F., 1999. A DEA approach to the airport production function. *International Journal of Transport Economics*, 26, 255–270.

Diewert, W.E., 1989. The measurement of productivity. *Working Paper,* Department of Economics, University of British Columbia, Vancouver.

Doganis, R., Graham, A. & Lobbenberg, A., 1995. The economic performance of European airports. *Research Report* 3. Dept of Air Transport, Colleague of Aeronautics, February 1995, University of Westminster, London.

Doyle, J. & Green, R., 1994. Efficiency and cross-efficiency in DEA: derivations, meanings and uses. *The Journal of the Operational Research Society*, 45(5), 567–578.

Efron, B., 1979. Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, 7(1), 1-26.

Farrell, M.J., 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A*, 120(3), 253-290.

Fernandes, E. & Pacheco, R.R., 2002. Efficient use of airport capacity. *Transportation Research Part A*, 36(3), 225–238.

Ferrier, G.D. & Lovell, C.A., 1990. Measuring cost efficiency in banking: econometric and linear programming evidence. *Journal of Econometrics*, 46(1-2), 229–245.

Forsyth, P.J., 2000. Models of airport performance. In: Hensher, D.A. & Button, K.J (eds.) *Handbook of Transport Modelling*, 597-608.

Forsyth, P.J., Hill, R.D. & Trengove, C.D., 1986. Measuring airline efficiency. Fiscal Studies, 7(1), 61–81.

Fried, H.O., Lovell, C.A.K. & Schmidt, S.S., 2008. *The Measurement of Productive Efficiency and Productivity Change*, Oxford University Press, Oxford.

Fung, M.K., Wan, K.K.H., Hui, Y.V. & Law, J.S., 2008. Productivity changes in Chinese airports 1995–2004. *Transportation Research Part E*, 44(3), 521–542.

Gillen, D. & Lall, A., 1997. Developing measures of airport productivity and performance: an application of data envelope analysis. *Transportation Research Part E*, 33(4), 261–273.

Gillen, D. & Lall, A., 2001. Non-parametric Measures of efficiency of US airports. *International Journal of Transport Economics*, 28(3), 283–306.

González, M.M. & Trujillo, L., 2009. Efficiency measurement in the port industry: A survey of the empirical evidence. *Journal of Transport Economics and Policy*, 43(2), 157–192.

Graham, A., 2005. Airport benchmarking: a review of the current situation. *Benchmarking: An International Journal*, 12(2), 99–111.

Graham, A., 2004. Managing airports an international perspective, 2nd edition, Amsterdam, Elsevier.

Graham, A. & Holvad, T., 2000. Efficiency Measurement for Airports. Paper presented at the Trafik Dags PAA Aalborg Universitet 2000 Conference, Aalborg University.

Greene, W., 2005. Fixed and random effects in stochastic frontier models. *Journal of Productivity Analysis*, 23(1), 7-32.

Hensher, D.A. & Waters, W.G., 1993. Using Total Factor Productivity and Data Envelopment Analysis for Performance Comparisons Among Government Enterprises: Concepts and Issues. *Institute of Transport Studies Working Paper No. ITS-WP-93-10,* The University of Sydney.

Hooper, P.G. & Hensher, D.A., 1997. Measuring total factor productivity of airports—an index number approach. *Transportation Research Part E*, 33(4), 249–259.

Jeong, J., 2005. An Investigation of Operating Costs of Airports: Focus on the Effects of Output Scale, *MSc-Thesis, Sauder School of Business,* University of British Columbia.

Kamp, V., Niemeier, H.M. & Müller, J., 2005. Can we Learn From Benchmarking Studies of Airports and Where do we Want to go From Here. Paper presented at the German Aviation Research Society Conference on Airport Benchmarking, November 2005, Vienna.

Kamp, V., Niemeier, H.M. & Müller, J., 2007. What can be learned from benchmarking studies? Examining the apparent poor performance of German airports. *Journal of Airport Management*, 1(3), 294–308.

Kincaid, I. & Tretheway, M., 2006. Guidelines for benchmarking airports. Paper presented at the German Aviation Research Society Conference, February 2006. Hamburg.

Kumbhakar, S.C. & Lovell, C.A.K., 2000. *Stochastic frontier analysis*, Cambridge University Press, Cambridge.

Lam, S.W., Low, J.M. & Tang, L.C., 2009. Operational efficiencies across Asia Pacific airports. *Transportation Research Part E*, 45(4), 654–665.

Lin, L.C. & Hong, C.H., 2006. Operational performance evaluation of international major airports: an application of data envelopment analysis. *Journal of Air Transport Management*, 12(6), 342–351.

Martín, J.C. & Román, C., 2006. A benchmarking analysis of Spanish commercial airports. A comparison between SMOP and DEA ranking methods. *Networks and Spatial Economics*, 6(2), 111–134.

Martín, J.C. & Román, C., 2001. An application of DEA to measure the efficiency of Spanish airports prior to privatization. *Journal of Air Transport Management*, 7(3), 149-157.

Martín, J. & Román, C., 2008. The relationship between size and efficiency: A benchmarking analysis of Spanish commercial airports. *Journal of Airport Management*, 2(2), 183-197.

Martín, J.C., Román, C. & Voltes-Dorta, A., 2009. A stochastic frontier analysis to estimate the relative efficiency of Spanish airports. *Journal of Productivity Analysis*, 31(3), 163–176.

Martín, J. & Voltes-Dorta, A., 2007. Stochastic Frontier Efficiency Estimation in Airports: a Bayesian Approach. Paper presented at the 11th ATRS Conference 2007, Berkeley.

Martín-Cejas, R.R., 2002. An approximation to the productive efficiency of the Spanish airports network through a deterministic cost frontier. *Journal of Air Transport Management*, 8(4), 233–238.

Meeusen, W. & van den Broeck, J., 1977. Efficiency estimation from Cobb-Douglas production functions with composed error. *International Economic Review*, 18(2), 435-444.

Morrison, W.G., 2009. Understanding the complexities and challenges of airport performance benchmarking. *Journal of Airport Management*, 3(2), 145–158.

Morrison, S.A. & Winston, C., 2008. Delayed! US Aviation Infrastructure Policy at a Crossroads. Aviation Infrastructure Performance: A Study of the Comparative Political Economy.

Murillo-Melchor, C., 1999. An analysis of Technical Efficiency and Productivity changes in spanish airports by using the Malmquist index. *International Journal of Transport Economics*, 26, 271–292.

Nyshadham, E.A. & Rao, V.K., 2000. Assessing efficiency of European airports: a total factor productivity approach. *Public Works Management & Policy*, 5(2), 106-114.

Orea, L. & Kumbhakar, S.C., 2004. Efficiency measurement using a latent class stochastic frontier model. *Empirical Economics*, 29(1), 169-183.

Oum, T.H., Adler, N. & Yu, C., 2006. Privatization, corporatization, ownership forms and their effects on the performance of the world's major airports. *Journal of air Transport management*, 12(3), 109–121.

Oum, T.H. et al., 1992. Concepts, methods and purposes of productivity measurement in transportation. *Transportation Research Part A*, 26(6), 493–505.

Oum, T.H. & Yu, C., 2004. Measuring airports' operating efficiency: a summary of the 2003 ATRS global airport benchmarking report. *Transportation Research Part E*, 40(6), 515–532.

Oum, T.H., Yu, C. & Fu, X., 2003. A comparative analysis of productivity performance of the world's major airports: summary report of the ATRS global airport benchmarking research report—2002. *Journal of Air Transport Management*, 9(5), 285–297.

Oum, T.H., Zhang, A. & Zhang, Y., 2004. Alternative forms of economic regulation and their efficiency implications for airports. *Journal of Transport Economics and Policy*, 38(2), 217–246.

Oum, T.H., Waters, W.G. & Yu, C., 1999. A survey of productivity and efficiency measurement in rail transport. *Journal of Transport Economics and Policy*, 33(1), 9-42.

Oum, T.H., Yan, J. & Yu, C., 2008. Ownership forms matter for airport efficiency: A stochastic frontier investigation of worldwide airports. *Journal of Urban Economics*, 64(2), 422-435.

Pacheco, R.R., Fernandes, E. & de Sequeira Santos, M.P., 2006. Management style and airport performance in Brazil. *Journal of Air Transport Management*, 12(6), 324–330.

Pacheco, R.R. & Fernandes, E., 2003. Managerial efficiency of Brazilian airports. *Transportation Research Part A*, 37(8), 667–680.

Parker, D., 1999. The performance of BAA before and after privatisation: A DEA study. *Journal of Transport Economics and Policy*, 33(2), 133-145.

Pathomsiri, S., Haghani, A., Dresner, M. & Windle, R.J., 2008. Impact of undesirable outputs on the productivity of US airports. *Transportation Research Part E*, 44(2), 235–259.

Pels, E., Nijkamp, P. & Rietveld, P., 2003. Inefficiencies and scale economies of European airport operations. *Transportation Research Part E*, 39(5), 341–361.

Pels, E., Nijkamp, P. & Rietveld, P., 2001. Relative efficiency of European airports. *Transport Policy*, 8(3), 183–192.

Pitt, M.M. & Lee, L.F., 1981. The measurement and sources of technical inefficiency in the Indonesian weaving industry. *Journal of Development Economics*, 9(1), 43-64.

Reinhold, A., Niemeier, H.-M., Kamp, V. & Müller, J., 2008. The Pros and Cons of Benchmark Regulation for Airports. Paper presented at the 12th ATRS Conference 2008, Athens

Sarkis, J., 2000. An analysis of the operational efficiency of major airports in the United States. *Journal of Operations Management*, 18(3), 335–351.

Sarkis, J. & Talluri, S., 2004. Performance based clustering for benchmarking of US airports. *Transportation Research Part A*, 38(5), 329–346.

Schmidt, P. & Sickles, R.C., 1984. Production frontiers and panel data. *Journal of Business & Economic Statistics*, 2(4), 367-374.

Sexton, T.R., Silkman, R.H. & Hogan, A.J., 1986. Data envelopment analysis: Critique and extensions. *New Directions for Program Evaluation*, 1986(32), 73–105.

Shirley, M.M. & Walsh, P., 2001. Public vs. private ownership: the current state of the debate. *Policy Research Working Paper Series No. 2420*, World Bank, Washington.

Simar, L. & Wilson, P.W., 2000. A general methodology for bootstrapping in non-parametric frontier models. *Journal of Applied Statistics*, 27(6), 779–802.

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

Tovar, B. & Martín-Cejas, R.R., 2009. Are outsourcing and non-aeronautical revenues important drivers in the efficiency of Spanish airports? *Journal of Air Transport Management*, 15(5), 217-220.

Tovar, B. & Martín-Cejas, R.R., 2010. Technical efficiency and productivity changes in Spanish airports: A parametric distance functions approach. *Transportation Research Part E*, 46(2), 249-260.

Transport Research Laboratory, 2002. Airport Performance Indicators. TRL, Wokingham.

Vasigh, B. & Gorjidooz, J., 2006. Productivity Analysis of Public and Private Airports: A Causal Investigation. *Journal of Air Transportation*, 11(3), 144-63.

Vickers, J. & Yarrow, G., 1991. Economic perspectives on privatization. *The Journal of Economic Perspectives*, 5(2), 111-132.

Vickers, J. & Yarrow, G., 1988. *Privatization: an economic analysis*, MIT Press, London.

Vogel, H.A., 2006. Impact of privatisation on the financial and economic performance of European airports. *Aeronautical Journal*, 110(1106), 197-213.

von Hirschhausen, C. & Cullmann, A., 2005. Questions to Airport Benchmarkers - Some Theoretical and Practical Aspects Learned from Benchmarking Other Sectors. Paper presented at the German Aviation Research Society Conference on Airport Benchmarking, November 2005, Vienna.

Yokomi, M., 2005. Measurement of Malmquist Index of Privatized BAA plc. Paper presented at the 9th ATRS Conference 2005, Rio de Janeiro.

Yoshida, Y., 2004. Endogenous-weight TFP measurement: methodology and its application to Japanese-airport benchmarking. *Transportation Research Part E*, 40(2), 151–182.

Yoshida, Y. & Fujimoto, H., 2004. Japanese-airport benchmarking with the DEA and endogenous-weight TFP methods: testing the criticism of overinvestment in Japanese regional airports. *Transportation Research Part E*, 40(6), 533–546.

Yu, M.M., 2004. Measuring physical efficiency of domestic airports in Taiwan with undesirable outputs and environmental factors. *Journal of Air Transport Management*, 10(5), 295–303.

Cornwell, C., Schmidt, P. & Sickles, R., 1990. Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels. *Journal of Econometrics*, 46, 185–200.

7 Appendix

Appendix 1: Studies using Non-Parametric Approaches

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	Account for observed heterogeneity	other remarks	Inputs	Outputs	Non-discretionary Variables
Abbott, M. and Wu, S. (2002)	12 main airports (all private except SYD)	1990 - 2000	132	Malmquist-DEA (BCC; inputmin.)	Tobit regression		=> # employees => capital stock => runway length	=> # pax => cargo	rate of return, capital labour ratio, aircraft standing area, total asset growth rate, ownership dummy, year dummy
	24 airports from Australia (12), NZ (3), UK (2), Canada (2) and US (5)	1998-1999	48	DEA (BCC, CCR; inputmin.)			=> # employees => total runway length => airport area => # aircraft standing positions	=> # pax => cargo	
Adler, N. and Berechman, J. (2001)	26 airports in Western Europe, North America and Far East (mainly Europe)	1998	26	PCA-DEA (BCC; inputmin.; super efficiency for ranking)	Banker and Morey (1986)		objective data: => airport charges => # passenger terminals => # runways => minimum connection times in minutes (1998 data) => distance to city centre (ND) => average delay (incl. in sensitivity analysis)	subjective data: => level of satisfaction grouped into five types	
Assaf, A. (2010a)	27 small and large UK airports	2007	27	DEA (CCR, BCC and NIRS; Bootstrapping)			=> # employees (FTE) => airport area => # runways	=> # pax => cargo => # ATM	
Barros, C.P. (2008a)	33 Argentine airports that are operated by Aeropuertos Argentina 2000	2003-2007	165	DEA (BCC; outputmax.)	Truncated bootstrap regression (Simar and Wilson 2007)		=> # employees => runway area => apron area => pax terminal area	=> # pax => cargo => # ATM	=> year dummy => hub dummy => WLU

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	Account for observed heterogeneity	other remarks	Inputs	Outputs	Non-discretionary Variables
Barros, C.P. and Assaf, A. (2009)	35 major US airports	2002-2007	210	Malmquist-DEA (Bootstrapping)			Terminal Services => # runways => # gates => airport area => staff costs Movement Model => airport area => # runways => staff costs	Terminal Services => # pax => cargo Movement Model => ATM => commuter ATM	
Barros, C.P. and Dieke, P.U.C. (2008)	31 Italian airports	2001-2003	93	DEA (CCR and BCC; outputmax.)	Truncated bootstrap regression (Simar and Wilson 2007) on CCR efficiency		=> staff costs => capital invested => other op. costs	=> ATM => # pax => cargo => handling revenues => aeronautical revenues => non-aeronautical revenues	=> hub dummy => WLU (airport size) => privatization dummy => north dummy
Barros, C.P. and Dieke, P.U.C. (2007)	31 Italian airports	2001-2003	93	DEA (CCR and BCC; outputmax.; Cross-Efficiency, Super-Efficiency for ranking)	Mann-Whitney-Hypothesis-Test on super-efficiency scores		=> staff costs => capital invested => other op. costs	=> ATM => # pax => cargo => handling revenues => aeronautical revenues => non-aeronautical revenues	=> airport size => privatization dummy => WLU
Barros, C.P. and Sampaio, A. (2004)	10 Portuguese airports	1990 - 2000	110	DEA (CCR and BCC)	Tobit regression		=> # employees => capital stock => price of labour => price of capital	=> # planes => # pax => cargo => mail cargo => landing revenues => pax revenues	=> market-share => share held by regional governments => location dummy => agglomeration (population) => cost structure (ratio of op. costs to sales)
Barros, C.P. and Weber, W.L. (2009)	27 UK airports	2000-2004	135	Malmquist-DEA			=> # employees => capital stock => other op. Costs	=> # pax => cargo => # ATM	

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	**Account for observed heterogeneity**	**other remarks**	**Inputs**	**Outputs**	**Non-discretionary Variables**
Bazargan, M. and Vasigh, B. (2003)	45 US airports (15 small, medium, and large hub airports)	1996 - 2000	225	DEA (CCR)	Kruskal-Wallis-Test and pairwise Mann-Whitney-Test among the three hub sizes	Inclusion of a virtual super efficient DMU	=> op. costs => non-operating costs => # runways => # gates	=> # pax => # ATM => # other ATM (commuter, GA, military) => aeronautical revenue => non-aeronautical revenue => % of on-time operations	
Chi-Lok, A. and Zhang, A. (2008)	25 Chinese (major) airports	1995-2006	300	DEA and Malmquist-DEA	OLS and Tobit regression		=> runway length => terminal size	=> # pax => cargo => # ATM	=> airport localization program => regional competition intensity => public listing Further (airport characteristics: Hub, local economy, coastal city, tourist city, population, demand and supply shocks, event: airline mergers, opensky agreements, new airports)
de la Cruz, F. (1999)	16 Spanish airports (largest airports)	1993-1995	48	DEA (CCR and BCC; inputmin.)		Several models with different outputs representing the airport as => infrastructure supplier => airport company => diversified airport company	=> total economic costs (operational, current costs and internal interests on net assets)	=> # pax => aeronautical revenues => handling revenues => non-aeronautical revenuea	
Fernandes, E. and Pacheco, R.R. (2002)	35 Domestic Brazilian airports	1998	35	DEA (BCC; outputmax.)			=> airport surface area => departure lounge => # check-in-desks => curb frontage in metres => # vehicle parking places => baggage-claim area	# pax (domestic)	

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	Account for observed heterogeneity	other remarks	Inputs	Outputs	Non-discretionary Variables
Fung, M.; Hui, Y., Law, J., Wan, K. and Ng, L. (2008)	25 Chinese regional airports	1995-2004	250	DEA (CCR; outputmax.) and Malmquist-DEA (dynamic)	OLS		=> runway length => terminal size both inputs are fixed in each period but variable over time	=> # pax => cargo => # ATM	=> location => hub dummy (int. , reg., non-hub) => stock ownership dummy (listed, non-listed)
Gillen, D. and Lall, A. (2001)	22 major US airports	1989 - 1993	110	Malmquist-DEA (Terminal: BCC, ATM: CCR; outputmax.)			Terminal Services => # runways => # gates => terminal area => # employees => # baggage collection belts => # public parking spots Movement Model => airport area => # runways => runways area => # employees	Terminal Services => # pax => cargo Movement Model => ATM => commuter ATM	
Gillen, D. and Lall, A. (1997)	21 major US airports	1989 - 1993	105	DEA (Terminal: BBC; ATM: CCR; outputmax.)	Tobit regression		Terminal Services => # runways => # gates => terminal area => # employees => # baggage collection belts => # public parking spots Movement Model => airport area => # runways => runways area => # employees	Terminal Services => # pax => cargo Movement Model => ATM => commuter ATM	Different sets of variables clustered according to => year dummy => hub dummy => noise strategy variables (only airside) => management operational and investment variables

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	Account for observed heterogeneity	other remarks	Inputs	Outputs	Non-discretionary Variables
Graham, A. and Holvad, T. (2000)	25 European and 12 Australian airports	1993 (Europe), 1992/93 (Australia)	37	DEA (CCR and BCC; FDH) and Partial Performance	Mann-Whitney-Test, Pearson Correlation Coefficient	Two Models: => all in one analysis => Europe and Australia separately Comparison of DEA-scores with partial productivity measures	=> # employees => capital costs => other op. costs	=> # pax (terminal) => cargo)	
Lam, S., Low, J.M.W. and Tang, L.C. (2009)	11 major Asia-Pacific airports	2001-2005	55	DEA (CCR and BCC; inputmin; SBM)		=> introduction of a virtual efficient airport	=> totalcosts => # employees => price of labour => other op. costs => price of capital => terminal area => trade value	=> # pax => cargo => # ATM	
Lin, L. and Hong, C. (2006)	20 major airports worldwide	2001 or 2002	20	DEA (CCR and BCC; superefficiency for ranking, cross-efficiency and FDH)	Hypothesis Testing		=> # employees => # check-in counters => # runway => # parking spaces => # baggage collection belts => # aprons => # boarding gages => terminal area	=> # pax => cargo => # ATM	=> ownerhip dummy => size, hub dummy => location dummy => economic growth of the country
Martín, J.C. and Román, C. (2008)	34 Spanish airports of different sizes	1997	34	DEA [cross-efficiency, super-efficiency (VRS and NIRS) and virtual rank-efficiency models; outputmax.]			=> staff costs => capital expenditures => other op. costs	=> # pax => cargo => # ATM	
Martín, J.C. and Román, C. (2006)	34 Spanish airports	1997	34	SMOP and DEA [cross-efficiency, super-efficiency for ranking (NIRS and BCC) and virtual-efficiency]			=> staff costs => capital expenditures => other op. costs	=> # pax => cargo => # ATM	
Martín, J.C. and Román, C. (2001)	37 Spanish airports (AENA)	1997	37	DEA (CCR and BCC; outputmax.)			=> staff costs => capital expenditures => other op. costs	=> # pax => cargo => # ATM	

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	**Account for observed heterogeneity**	**other remarks**	**Inputs**	**Outputs**	**Non-discretionary Variables**
Murillo-Melchor, C. (1999)	33 Spanish civil airports run by AENA	1992 - 1994	99	Malmquist-DEA (CCR, BCC and NIRS)			=> # employees => capital stock => other op. Costs	=> # pax	
Pacheco, R.R. and Fernandes, E. (2003)	35 Brazilian Domestic airports (Infraero)	1998	35	DEA (BCC; inputmin.)			=> # employees => staff costs => other op. costs	=> # pax (domestic) => cargo => operating revenue => non-aeronautical revenue => other revenues	
Pacheco, R.R., Fernandes, E. and Peixoto de Sequeira-Santos, M. (2006)	58 Brazilian airports administered by Infraero	1998-2001	232	DEA (BCC; inputmin.)			=> # employees => staff costs => other op. costs	=> operating revenue => non-aeronautical revenue => other revenue => # pax => cargo	
Parker, D. (1999)	(1) BAA as a whole and (2) 22 UK airports	1979 - 1995 and 1988- 1996 second model	17 (198)	DEA (CCR and BCC)	Banker and Morey (1986)		=> # employees => capital input => other op. Costs	=> # pax => cargo	=> changes in GDP
Pathomsiri, S., Haghani, A., Dresner, M. and Windle, R.J. (2008)	56 US airports	2000-2003	56 (224)	Directional distance function with Luenburger productivity indicator (instead of Malmquist	correlation to analyse factors affecting airprot productivity	To analyse impact of undesirable outputs, there are three cases (see paper)	=> land area => # runways => runway area	Desirable Outputs: => non-delayed ATM => # pax => cargo Undesirable Outputs: => delayed ATM => time delays	=> load factor (pax/ATM) => cargo load factor => % GA => % int. pax => market share
Pels, E., Nijkamp, P. and Rietveld, P. (2003)	34 European airports	1995 - 1997	102	DEA (inputmin.)	see SFA model		Movement Model => airport surface area => # aircraft parking places at terminal => # remote aircraft parking places => # runways (ND) Terminal Model => # check-in-desks => # baggage claim units	Movement Model => ATM Terminal Model => # pax	Movement Model => dummy for slot-coordinated airport => dummy for time restrictions Terminal Model => time dummy => airlines' load factor

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	**Account for observed heterogeneity**	**other remarks**	**Inputs**	**Outputs**	**Non-discretionary Variables**
Pels, E., Nijkamp, P. and Rietveld, P. (2001)	34 European airports	1995 - 1997	102	DEA (inputmin.)			Terminal Model: => terminal size (in sqm) => # aircraft parking places at the terminal => # remote aircraft parking places => # check-in-desks => # baggage-claim units Movement Model: => total airport area => total length of runway => # aircraft parking positions at the terminal => # remote aircraft parking positions	Movement Model => ATM Terminal Model => # pax	
Sarkis, J. (2000)	44 Major US airports	1990 - 1994	220	DEA (CCR and BCC; Agressive Cross-Efficiency, Ranked Efficiency , Radii of Classification Rankings)		=> Multi Airport Systems vs. Single Airport Systems => Hubs vs. Non-Hubs => Airports in Snowbelts vs. Airports outside Snowbelts	=> op. costs => # employees => # gates => # runways	=> operational revenue => # pax => # ATM => # ATM (GA) => cargo	
Sarkis, J. and Talluri, S. (2004)	44 Major US airports	1990 - 1994	220	DEA (CCR; Basic and Cross-Efficiency models) and clustering method to receive appropriate benchmarks		=> cluster analysis after DEA estimation from cross-efficiency results: based on correlation coefficient b/w pairs of columns which indicate that the corresponding airports are similar	=> op. costs => # employees => # gates => # runways	=> operational revenue => # pax => # ATM => # ATM (GA) => cargo	
Vogel, A. (2006)	31 European airports and 4 airport systems	1990-2000	341	DEA (CCR and BCC), partial performance, financial ration analysis			=> total costs	=> total revenues	

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	Account for observed heterogeneity	other remarks	Inputs	Outputs	Non-discretionary Variables
Yokomi, M. (2005)	6 BAA airports (SOU excluded)	1975-2001	162	Malmquist-DEA (outputmax.)			For Both Sides => # employees => other op. costs	Aeronautical Analysis => # ATM Non-Aeronautical Analysis => non-aeronautical revenue	
Yoshida, Y. and Fujimoto, H. (2004)	67 Japanese airports	2000	67	DEA (CCR and BCC)	DEA: Tobit Regression	airport categories serving (1) int. airlines (2) domestic airline (3) regional airlines (4) other	=> runway length => terminal area => # employees in terminal => average access cost (monetary and time costs)	=> # pax => cargo => # ATM	Both: => thirdmainland dummy => start of operation
Yu, M.-M. (2004)	14 domestic Taiwanese airports	1994-2000	98	DEA (directional output distance function with window analysis with 2 years window)	Banker and Morey (1986)		=> runway area => apron area => terminal area => # connections with other domestic airports	Desirable: => # ATM => # pax Undesirable: => aircraft noise	=> population in the county

Appendix 2: Studies using Parametric Approaches

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	Account for observed heterogeneity	other remarks	Inputs	Outputs	Non-discretionary Variables
Assaf, A. (2010b)	13 major Australian airports (post-privatization)	2002-2007	78	SFA (homogenous Cobb-Douglas cost function with Bayesian)			=> total costs => price of labour => price of capital premises	=> # pax => cargo => # ATM	
Assaf, A. (2008)	27 UK airports (16 large and 11 small)	2002-2006	135	SFA (heterogenous Cobb-Douglas production-Metafrontier with ML)	see SFA model: two group frontiers for small and large airports and the metafrontier)		=> # employees => fixed assets => operational costs => other costs	=> operational income	
Barros, C.P. (2009)	27 UK airports	2000-2006	189	SFA (heterogenous Translog costfunction latent class model with ML)	see SFA model: data set is clustered in three homogenous groups in terms of market share		=> op. costs => price of labour => price of capital => price of capital investment	=> # pax => ATM	
Barros, C.P. (2008b)	13 Portuguese airports	1990 - 2000	143	SFA (homogenous translog cost function, calculating technical change with Malmquist indices with ML)			=> op. costs => price of capital => price of labour	=> landing revenues => pax revenues => non-aeronautical revenues	
Barros, C.P. (2008c)	27 UK airports	2000-2005	125	SFA (comparison of homogenous and heterogenous long-run translog cost function based on true random effects with ML)			=> op.costs => price of workers => price of capital premises => price of capital investment	=> # pax => # ATM	
Barros, C.P. and Marques, R.C. (2008)	117 worldwide airports	2003-2004	234	SFA (Translog Cost frontier), (1) Homogenous non-random and (2) Heterogenous random Model with ML	see SFA model: investigation of effects of unobserved heterogeneity, unobserved managerial ability, regulation and ownership		=> op. costs => price of labour => terminal size	=> # ATM => # pax	=> hub dummy => rate of return dummy => incentive regulation dummy => ownership dummy

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	Account for observed heterogeneity	other remarks	Inputs	Outputs	Non-discretionary Variables
Chow, C.K.W., Kong, C. and Fung, M.K.Y. (2009)	46 Chinese airports (three int hubs, six reg hubs and 37 regional airports)	2000	46	SFA (1) estimated homogenous single output translog production function with ML (2) partial translog input distance function	see SFA model		(1) single output First stage: => airport area => runway length => # of terminal parking positions Second-stage: Both: => predicted ATM APM: => terminal area => car-park area ACM: => cargo handling facility area (2) multi-output => ATM (intermediate output) => terminal area => cargo facilities area => # of aircraft parling positions => airport area => runway length	(1) single output First-stage: => # ATM Scond-stage: APM: => # pax ACM: => cargo (2) multi-output => # pax => cargo	(1) single output Both: => regional effects dummy => major airlines dummy ACM: => % international cargo (2) multi-output => regional effect dummy => major airlines dummy => int. or reg. hub dummy => % int. traffic
Martín, J.C. and Román, C. and Voltes-Dorta, A. (2009)	37 Spanish airports (AENA)	1991-1997	259	SFA (homogenous translog multiproduct long-run cost function with Bayesian)			=> staff costs => capital expenditures => other op. costs => price of labour => price of other (totex per ATM, totex per WLU)	=> # WLU => # ATM	
Martín, J.C. and Voltes-Dorta, A. (2007)	41 worldwide airports	1991-2005	284	SFA (homogenous translog multiproduct long-run cost function with Bayesian approach)			=> staff costs (and prices) => other op. costs (and prices) => capital costs (and prices) => # employees (FTE) => fixed assets	=> # pax => cargo => # ATM	
Martín-Cejas, R.R. (2002)	40 worldwide airports	1996-1997	80	translog cost function estimation with OLS			=> total costs => price of labour => price of capital	=> # WLU	

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	Account for observed heterogeneity	other remarks	Inputs	Outputs	Non-discretionary Variables
Oum, T.H. and Yu, Ch. (2004)	60 major worldwide airports (only 50 of 1999 in TFP)	1999-2000	110 (50 in TFP)	partial productivity measures and gross EW-TFP			=> # employees (FTE) => # runways => # gates => total terminal area => other op. Costs	=> # pax => cargo => commercial revenues	=> airport size => congestion delays => hub dummy => ownership dummy => continent dummy => % non-aeronautical revenues => regulation dummy
Oum, T.H., Yu, Ch. and Fu, X. (2003)	109 worldwide airports	2001-2004	776	SFA (homogeneous translog cost function with Bayesian approach, random effects, ownership enters u and other NDs enter the technology)	see SFA model	2 models: - base model - controlling for multiple market effects inclusion of labour share equation	=> # employees (FTE) (and price) => other op. costs (and price) => # runways (fixed) => terminal size (fixed)	=> # pax => # ATM => non-aeronautical revenues	=> % int. traffic => % cargo => regional dummy variables => ownership dummy (see paper) => multi-airport dummy (in model 2)
Oum T.H, Zhang, A. and Zhang, Y. (2004)	50 major worldwide airports	1999	50	partial productivity measures and EW-TFP	Regression analysis to obtain residual TFP	Three models: (1) all explanatory variables (2) all except hub (3) excludes hub and ownership (treats ownership as public or other)	=> # employees (FTE) => # runways => # gates => total terminal area => other op. Costs	=> # pax => cargo => non-aeronautical revenues	Factors beyond managerial control: '=> ownership dummy => airport size => average aircraft size => composition of air traffic Factors within managerial control: => business diversification => extent of outsourcing => service quality
Pels, E., Nijkamp, P. and Rietveld, P. (2003)	34 European airports	1995 - 1997	102	SFA (homogenous translog production function, ML, z-Var in u)	see SFA model		Movement Model => airport surface area => # aircraft parking places at terminal => # remote aircraft parking places => # runways (ND) Terminal Model => # check-in-desks => # baggage claim units	Movement Model => ATM Terminal Model => # pax	Movement Model => dummy for slot-coordinated airport => dummy for time restrictions Terminal Model => time dummy => airlines' load factor

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	**Account for observed heterogeneity**	**other remarks**	**Inputs**	**Outputs**	**Non-discretionary Variables**
Pels, E., Nijkamp, P. and Rietveld, P. (2001)	34 European airports	1995 - 1997	102	SFA (homogenous CD production function with ML; z-Var in u)	see SFA model		Terminal Model => # baggage-claim units => # aircraft parking places at the terminal => # remote aircraft parking places Movement Model => # runways => # aircraft parking places at the terminal => # remote aircraft parking places	Movement Model => ATM Terminal Model => # pax	=> year dummy
Tovar, B., Martín-Cejas, R.R. (2010)	26 Spanish airports	1993-1999	182	SFA (homogenous with input-oriented translog distance function, z-Var in u, ML)	see SFA model		=> # employees => surface area => # gates	=> # ATM => average size of aircraft => share of non-aeronautical revenues	=> outsourcing degree => non-aeronautical revenues => cargo
Tovar, B., Martín-Cejas, R.R. (2009)	26 Spanish airports	1993-1999	182	SFA (homogenous translog production input distance function and decomposing TFP into its components)			=> # employees => surface area => # gates	=> # ATM => average size of aircraft => share of non-aeronautical revenues	
Yoshida, Y. (2004)	30 Japanese airports	2000	30	EW-TFP		airport were categorized into airports serving mainly (1) international routes (2) domestic trunk routes (3) regional air-traffic	=> terminal size => runway length	=> # pax => cargo => # ATM	
Yoshida, Y. and Fujimoto, H. (2004)	67 Japanese airports	2000	67	EW-TFP	EW-TFP: OLS	airport categories serving (1) int. airlines (2) domestic airline (3) regional airlines (4) other	=> runway length => terminal area => # employees in terminal => average access cost (monetary and time costs)	=> # pax => cargo => # ATM	Both: => thirdmainland dummy => start of operation

Appendix 3: Studies using Price-based index Approaches

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	Account for observed heterogeneity	other remarks	Inputs	Outputs	Non-discretionary Variables
Hooper, P.G. and Hensher, D.A. (1997)	6 Australian airports	1988- 1992	30	TFP (CCD)		2 regression analysis in second step with output index and airport-specific dummy variables as explanatory variables respectively to assess effects on TFP and to obtain an output-adjusted TFP index	=> capital stock => staff costs => other op. Costs	=> aeronautical revenue => non-aeronautical revenue	
Nyshadham, E.A. and Rao, V.K. (2000)	25 European airports (Doganis et al. data)	1995	25	Partial Performance and TFP (CCD)			=> op. cost per WLU => capital expenditures per WLU => other costs per WLU	=> aeronautical revenue per WLU => non-aeronautical revenue per WLU	
Oum, T.H., Adler, N. and Yu, Ch. (2006)	116 worldwide airports	2001 - 2003	254	VFP (CCD)	Regression analysis		=> # employees (FTE) => other op. costs	=> # pax => # ATM => non-aeronautical revenues	Airport Characteristics: => airport size (aggr output) => average aircraft size => % int. traffic => % cargo => runway utilization (ATM/runway) Management Strategies: => % non-aviation revenue Ownership: => ownership form (see paper) Other: => continental dummy variable => year dummy

Authors	Sample	Time Period	# of obs.	Methodology			Data		
				Model	Account for observed heterogeneity	other remarks	Inputs	Outputs	Non-discretionary Variables
Oum, T.H., Yan, J. and Yu, Ch. (2008)	76 worldwide airports	2000-2001	152	VFP (CCD)	Regression analysis to obtain residual VFP	2 models: - base model - controlling for multiple market effects	=> # employees (FTE) => other op. costs	=> # pax => cargo => # ATM => non-aeronautical revenues	Factors beyond managerial control: => airport size => average aircraft size => % int. traffic => % air cargo in total traffic => capacity constrained airports Factors within managerial control: => pax service levels => non-aeronautical business =>airline or independent company operated terminals
Vasigh, B. and Gorjidooz, J. (2006)	22 major airports in the US and Europe (8 US public, 7 BAA, 7 busiest EU private and public)	2000-2004	110	TFP (CCD)	stepwise-regression between TFP-results and ownership structure		=> op. cost => net total assets => runway area	=> operational revenues => non-operational revenues => # pax => # ATM => landing revenues	=> ownership dummy other factors see paper p.157 clustered acc. to financial condition and management policy