Topics in efficiency benchmarking of energy networks: Selecting cost drivers

Report prepared for

The Netherlands Authority for Consumers and Markets

15 December 2017

Denis Lawrence, John Fallon, Michael Cunningham, Valentin Zelenyuk, Joseph Hirschberg

Economic Insights Pty Ltd
Ph +61 2 6496 4005 Mobile +61 438 299 811
WEB www.economicinsights.com.au
ABN 52 060 723 631

CONTENTS

EXECUTIVE SUMMARY

This paper highlights relevant issues for consideration when developing candidate variables to include in benchmarking studies of gas and electricity transmission system operators (TSOs). This study is intended to inform benchmarking analysis, and while data envelopment analysis (DEA) is the main benchmarking method to which it is ultimately directed, the studies examined here are not confined to DEA applications, and include various types of cost and productivity analysis. Although many approaches are discussed in the report, it needs to be recognised that feasibility and resource limitations will influence the most ideal or optimal approach that can be implemented in practice.

The types of variables most commonly used in efficiency benchmarking studies are production inputs and outputs, input prices or costs, and variables that reflect features of the different operating environments of TSOs. As a non-parametric method, DEA does not impose any specific functional form on the technology or production possibility set, however, variables used as inputs and outputs are chosen by the analyst and, as with all benchmarking methods, the choice of inputs and outputs is fundamentally important to the results obtained. DEA itself does not provide guidance on the choice of input and output variables. There is considerable reliance on the judgement and expertise of the analyst. However, it is well known that model misspecification has significant impacts on DEA efficiency estimates, especially in small samples. Furthermore, achieving parsimony in the use of variables is also vitally important in DEA analysis, which loses discriminatory power as the dimensionality of the production space increases (i.e., as the number of outputs and inputs included increases). This is an especially important problem when sample size is small.

Principles for Identifying Relevant Variables

The choice of candidate variables is often developed through consultation processes with industry, which can assist to prioritise the variables for inclusion and filter out those that are unlikely to have any explanatory power. Three general approaches that appear to be taken to identifying the variables are:

- Advice from engineering or business process experts regarding what was logical or plausible from their perspectives.
- Some formal statistical techniques are available to assist to filter and screen variables, and although they have some limitations, especially when data samples are relatively small, they are arguably much better than *ad hoc* or trial and error procedures.
- It is also feasible and desirable to consider the methodologies and variables used in the literature, particularly those that may be considered the 'conventional' or 'best practice' in energy network benchmarking. This study undertakes a literature review of this kind.

It is important in any benchmarking exercise to carefully distinguish between inputs and outputs. There will inevitably be a high correlation between inputs and cost (because of the definition of cost), but if inputs were included as 'cost drivers' this would not yield a true measure of cost efficiency. This is because the economic functional relationship between cost and its determinants (outputs and input prices) would be conflated with the definitional relationship between costs and inputs (as the sum or products of input quantities and prices).

In cost efficiency benchmarking it is common to estimate the DEA cost efficiency model in conjunction with the input-oriented technical efficiency model (which measures the degree to which the use of inputs could be reduced while producing a given levels of outputs) because this enables cost efficiency to be decomposed into technical and allocative efficiency (both input oriented). It is a useful discipline to estimate both models, because the technical efficiency model requires that an economically meaningful distinction be made between the inputs and outputs.

Ideally, the inputs and outputs included in the DEA analysis should be sufficiently comprehensive to capture the relevant features of the production process, including the quality and quantity of outputs. Otherwise the measures of efficiency may be inaccurate. This means that a theory of the production process needs to be formed, which will usually combine engineering and economic perspectives. It is also preferable that each different service provided should be measured by a separate output, and each distinct factor of production should be measured using a separate input. However, unless the data sample is quite large, there will usually be practical limitations that require some compromises. This means there are challenges to undertaking DEA analysis with small samples.

Studies of Gas and Electricity TSOs

This study includes a literature review of thirteen studies of gas TSO cost functions or production technology published between 1987 and 2016 and fourteen studies of electricity TSO cost functions or production technology published between 1978 and 2014. Some of the conclusions that can be drawn from these reviews are as follows.

- A wide variety of different inputs and outputs have been used in these studies. This suggests a lack of consensus on the main features of the technologies of gas and electricity TSOs.
- There have been some distinct changes in the types of variables used as inputs in the studies conducted in the period up to 2005 and in the post-2005 period, especially in relation to gas TSOs. The earlier studies tended to rely on separate measures of non-capital inputs (usually proxied by employee numbers) and capital inputs (measured either using physical capital measures or deflated monetary measures such as real fixed assets), whereas the later studies relied almost exclusively on total cost, or total variable cost, as a single input. There is not such a clear pattern in the input variables used in studies of electricity TSOs, because the variety of variables has been so diverse over all the studies examined.
- There has been a corresponding change in relation to the use of input prices. In regard to gas TSO studies, although the use of input prices was quite common in the studies up to 2005, none of the studies in the post-2005 period used any input prices. Similarly, in regard to electricity TSO studies, whereas a small number of studies included input prices in the period up to 2005, none of the studies in the post-2005 period did so.

- There doesn't appear to be such a clear pattern of difference in the use of output measures in the periods before and after 2005. Although a wide range of output measures have been used, for gas TSOs the most common were: (a) gas throughput and transport distance measures either included separately or combined into a single volume-distance measure; and (b) some measure(s) of maximum delivery capacity, either using a peak day demand measure or a physical supply capacity measure. For electricity TSOs, the most common types of output variables were: (a) electricity throughput, or separate measures of inflows and outflows; (b) transport distance measures (often proxies by the length of the network, for which there are different measures) or capacity × distance measures; and (c) peak day demand measures of maximum delivery capability were used in some studies, whereas physical supply capacity measures were not often used as electricity TSO outputs. Overall, surprisingly few studies of electricity TSOs used some measure of peak supply capability as an output, given that it is widely viewed as a key driver of costs.
- There is a lack of consistency in the categorisation of variables as inputs or outputs. A number of gas TSO studies used as outputs, variables that appear as inputs in other studies, and the same applies to electricity TSO studies. There appears to be confusion in the categorisation of some capacity-related variables for gas pipelines as inputs or outputs. Similarly, there is confusion in the categorisation of electricity TSO variables such as transformer capacity and capacity-distance (MVA-km) type measures. In part this reflects difficulties in deriving proxy measures for capital input services and for aspects of customer services that are capacity-related, such as supply security and ability to meet peak day demand.
- There was an increased reliance on DEA as an analytical method in the period after 2005, whereas in the earlier period there was a relatively greater use of econometric methods (which were more frequent than DEA in the period up to 2005) and multilateral TFP analysis. This observation applies to studies of both gas and electricity TSOs.
- There was a tendency in most studies to disregard differences in operating environments between TSOs. Very few of the studies took account of differences in the operating environments either by including such variables within the analysis or by conducting a second-stage analysis of efficiency scores against operating environment characteristics. This observation applies to studies of both gas and electricity TSOs.

Candidate Outputs and Inputs

Several conceptual issues needed to be addressed in relation to the types of outputs to be considered. These included:

(1) Outputs can be identified with the quantities used for billing customers under the transmission tariffs, or they can be identified as variables that best reflect the services provided to users and that drive the costs of supply (i.e. 'functional' outputs). This divergence can occur because utility tariffs need not be well aligned with the services they provide. On balance, the functional approach has advantages, but it is advisable

to have regard to billed outputs also.

(2) There is a range of difficulties in measurement of capacity, whether required to deliver outputs now and in the future. These include whether the capacity-related security of supply service is mainly related to peak day or peak hour supply capability, or whether it is broader. How peak day or peak hour should be measured, including 'ratcheted' measures, or probabilistic estimates with a low probability of exceedance. Whether a long-term perspective is needed in relation to capacity, since efficiently planned capacity may nevertheless not be optimal in a short-term perspective, either because of historical uncertainties affecting network planning, or because their optimality can only be fully assessed from a long-term perspective.

(3) There are also issues relating to the choice and inclusion of service reliability and quality measures. Preferably, benchmarking analysis will take into account the quality and/or reliability of service if they are important to customers. However, there are issues to be addressed when doing so. Service quality may be partly due to decisions of TSOs (and hence endogenous) and partly due to exogenous factors (e.g. severe weather). The interest for benchmarking is primarily in endogenously determined output quality, and its inclusion as an output in DEA analysis may conflict with the chosen orientation depending on the method chosen. Alternatively, exogenous factors that influence service quality can be included in a second-stage analysis to adjust for their effects.

The main two groups of inputs are durable and non-durable inputs. Specific issues relating to measuring the quantities and prices of capital inputs are addressed in our report on Capital Costs. Some general issues in relation to non-capital input quantities and prices (or together, the input cost) are as follows.

(1) Decisions need to be made in relation to the level of aggregation at which inputs are to be measured. In many benchmarking studies, just two inputs are included, capital and non-capital inputs. In some cases either capital or non-capital inputs (or both) may be disaggregated into their main components. For example, for gas TSOs, non-capital inputs may be separated into compressor fuel and other non-capital inputs. Gas TSO capital inputs are sometimes separated into pipelines and compressors, which can in some circumstances be substitutes.

(2) It may be difficult to directly measure the quantities of some inputs because they are heterogeneous. Instead the quantity may be estimated by dividing the cost of that input by an appropriate price index for that input. A price index for a group of inputs should be a weighted average of the prices of the key components of that group of inputs (where the weights are cost shares). Ideally, this weighted price index will reflect as closely as possible the prices faced by each TSO. As an example, because of the diverse composition of operating and maintenance costs, and differences between businesses in relation to contracting out or in-house provision of services, direct measurement of non-capital input quantities is often difficult and the method of deflating relevant costs is often used.

(3) Consistency of the operating and maintenance cost ('opex') data collected from TSOs is important. Some areas where particular attention is needed to ensure consistency

are:

- capitalisation practices, for example in relation to isolated asset refurbishment;
- cost-allocation methods, such as corporate overhead allocation in businesses that have other activities in addition to electricity (or gas) transmission;
- related party services, such as a network operating agreement with a related company, can cause comparability difficulties if the transfer price is not cost reflective;
- energy losses in transmission may be treated differently between jurisdictions. For example, in some cases generators may bear these costs and in others they may be borne by the TSO. Care is needed to ensure consistency.

Operating Environment Factors

Utilities tend to operate in discrete geographical areas, and features of the geographical location, including topography, characteristics of the urban areas supplied (e.g., density) and climate in those locations, may all have an important influence on observed productivity and cost efficiency. These operating environment characteristics essentially act as constraints, and can influence the ability of businesses to convert inputs into outputs. The aim of making like-for-like comparisons in benchmarking studies supports taking operating environment factors into account. However, there is an issue of regulatory judgement around which types of factors to allow for, since excessive allowances for operating environment factors may lead to over-estimating their influence and under-estimating efficiency differences between TSOs, thereby weakening the efficiency incentives within the regulatory framework.

It is important to concentrate on only those operating environment factors that have the most significant effect and which vary the most across TSOs. Where a number of operating environment factors are highly correlated, only the one with the most direct impact on TSOs' costs may be included. Some of the types of operating environment factors that can have an important bearing on energy network costs include:

- Climate and terrain can have an important influence on infrastructure construction and maintenance costs;
- Concentration or dispersion of demand centres and distances between energy sources and demand centres will influence the design of networks including whether their configuration is linear or meshed, etc. These characteristics are sometimes referred to as 'network topography'.
- Regulations and Standards are usually exogenous to the TSO and, if they are binding constraints, they may have a material impact on costs that is difficult to quantify robustly and objectively.
- If peak demand is not included as an output, it is sometimes included as an operating environment factor (e.g. severity of winter cold which influences peak winter energy demand for heating or severity of summer heat, which influences energy demand for air-conditioning).

Various methodologies can be used to control for the influences of non-discretionary or operating environment factors. At a broad level this may be done:

- before the DEA analysis, such as pre-analysis adjustment of data;
- during the DEA analysis, by including operating environment variables in the DEA analysis alongside inputs and outputs, or by using subsamples of like TSOs in the analysis; or
- after the DEA analysis, such as by using 'second stage' approach to analyse and control for the influence of business environment factors on measured efficiency.

The approach to controlling for operating environment characteristics by treating them as additional inputs or outputs in the DEA analysis can be contentious because efficiency measurement in DEA assumes that the inputs produce the outputs, and there is no reason to expect that assumptions derived from production theory, such as monotonicity, convexity, etc., would apply to those variables. Furthermore, since operating environment factors are generally exogenous, they cannot usually be proportionately scaled down in input-oriented DEA (or scaled up in output-oriented DEA) through management discretion, as is typically assumed for regular inputs and outputs.

A more common approach is firstly to carry out the DEA analysis without controlling for the exogenous factors, and then conduct a 'second stage' analysis, in which the estimated efficiencies scores are used as the dependent variable in a regression against the operating environment factors. The model obtained from the second stage regression can be used to calculate adjusted efficiency scores which control for differences in the exogenous factors. More than one approach may be used, for example when some exogenous factors may be readily controlled-for through normalisation of variables before the DEA analysis, others can be addressed through second-stage analysis.

Techniques for Variable Selection or Reduction

This paper also addresses techniques for variable selection and search for parsimony. This includes principles for screening and selecting the most suitable variables to use in a benchmarking study. As previously mentioned, an important limitation of DEA is that, as more input and output variables are added, and its dimensionality increases, it loses some ability to discriminate between efficient and inefficient DMUs, especially with small data samples.

In any parametric or nonparametric modelling task, knowledge of the industry being examined is important to ensure that the variables used and the specifications employed are likely to be sufficiently representative, at least as a starting point for analysis. This is one reason why benchmarking exercises usually involve consultation with industry participants and industry experts to assist in ensuring that the most appropriate variables and definitions are considered. Economic theory also has an important role to play in understanding the way that the industry works, how the variables are likely to be interrelated, and how operating environment factors are likely to influence the production process. In addition to these fundamental 'first principles', a number of techniques for variable selection are briefly reviewed. These include:

- Reliability assessments of trial DEA results obtained when using specific sets of variables. For example, this may include examining whether the marginal rates of transformation or substitution implied by the weights obtained in the DEA solution are consistent with the findings of other studies in the literature or with past regulatory benchmarking analysis of the same TSOs.
- Methods that rely on partial correlations between partitioned sets of candidate variables, such as the method developed by Jenkins and Anderson (2003).
- The 'efficiency contribution measure' method of Pastor *et al* (2002), which involves comparing differently specified DEA models to determine the incremental effect of each variable on the efficiency measures of firms.
- Preliminary regressions may be used to identify variables that are potential explanatory variables of a single input (variable or total cost) as used by Jamasb *et al* (2007), Jamasb *et al* (2008), and Frontier, Consentec & Sumicsid (2013), among others.
- The regression-based approach of Ruggiero (2005), an iterative process beginning with a minimally specified DEA model, regressing the resulting efficiency score estimates on remaining candidate variables, and identifying any significant variables that might be added to the model.
- The Simar and Wilson (2001) bootstrapping method to statistically test, within a DEA setting, whether some outputs or inputs in the model are irrelevant, or whether some inputs, or outputs, can be aggregated.

Even after narrowing them down, the set of variables that describe the technology may remain too large given the size of the sample. Two approaches to achieving greater parsimony in these circumstances are reviewed.

The first of these approaches is aggregation, which can involve: (a) simply adding together variables if they are in the same units (e.g. monetary units); (b) constructing indexes for some group of inputs, or grouping of outputs, for example using the Törnqvist index method; or (c) combining variables in a meaningful way, such as by using engineering formulas or formulas derived from commercial practices (such as the volume × distance measures sometimes used).

The second approach is to use principal components analysis (PCA) to transform the set of original variables into a smaller group of derived variables that contain much of the information in the original variables, thereby reducing dimensionality with minimal loss of information, and hence minimal bias to the efficiency estimates obtained. PCA-DEA has been used in a number of DEA benchmarking studies, including one of the TSO benchmarking studies surveyed. It involves using the leading components (or principal components) as the variables in the DEA analysis rather than the original variables. It has the particular advantages of:

- allowing a richer set of input and output variables to be used in the overall analysis (thereby improving the ability to identify 'true' efficiency); while also
- enabling a reduced number of variables used in the DEA analysis (thereby mitigating

the dimensionality and discrimination problems).

Objectives within the regulatory framework will also be relevant considerations in the selection of the set of candidate variables. This is because the candidate variables used for a benchmarking study will influence the dimensions in which services and inputs are viewed, which will in turn affect the way in which efficiency targets are formulated for regulated businesses. Furthermore, the choice of variables included in the final analysis may have an influence on the incentives of regulated businesses.

1 INTRODUCTION

This study discusses methodologies for selecting the variables to be tested and used as 'cost drivers' in benchmarking analysis, or when examining reasons for differences in the estimated efficiencies, of gas and electricity transmission system operators (TSOs). Generally speaking, the types of variables most commonly used in efficiency benchmarking studies are production inputs and outputs, input prices or costs, and variables that reflect features of the different operating environments of TSOs which go toward explaining differences in costs or estimated efficiencies.

Although the discussion of inputs and outputs of businesses is relevant to a range of benchmarking methods, the principal method of interest in this study is data envelopment analysis (DEA). However, the studies examined here are not confined to DEA applications, and include various types of cost and productivity analysis.

This paper addresses the following topics:

- The economic and benchmarking principles relevant to determining candidate variables for use as 'cost drivers' when benchmarking costs (chapter 2);
- A detailed review of benchmarking studies of energy transmission networks in academic literature, and studies commissioned by economic regulators having regard especially to the variables used in the studies (chapter 3);
- A more general discussion of candidate variables for benchmarking TSOs and discussion of issues in the measurement of variables of interest, service quality and operating environment characteristics. Methods of dealing with differences between TSOs operating environments are discussed, including the pros and cons of alternative approaches such as: pre-analysis adjustment of data; using subsamples of similar TSOs; whether to include operating environment variables in the DEA analysis or use them in a separate 'second-stage' analysis to analyse and control for their effects (chapter 4);
- Methods and principles for screening and selecting among the variables to use in a benchmarking study, and techniques for achieving parsimony in the use of variables in DEA analysis, including potential use of principal components analysis (PCA) (chapter 5); and
- The implications of the study for data collection relating to European gas and electricity TSOs (Appendix A).

The aim of this paper is to highlight relevant issues for consideration when developing candidate variables for inclusion in gas and electricity TSO benchmarking.

2 PRINCIPLES FOR IDENTIFYING RELEVANT VARIABLES

European national regulatory authorities (NRAs) are responsible for setting revenue or tariff caps for network businesses. The Netherlands has one gas transmission system operator (TSO), GTS, and one electricity TSO, TenneT NL. Like several other NRAs, ACM has used benchmarking to ascertain efficient costs for the purpose of setting revenue caps. The measured efficiency scores are used to determine the trajectory of the assumed level of 'efficient cost' over the regulatory period. In both electricity and gas transmission, DEA analysis has been used to benchmark European TSOs.

DEA models the set of input-output combinations that are feasible for businesses, based on observed input-output combinations and on principles of economic theory. The DEA technique uses linear programming to estimate a production or cost efficiency relative to the observed best-practice frontier around a set of data. Different assumptions may be made in regard to returns-to-scale (e.g., constant, non-increasing, variable) as well as inputs and outputs. DEA is applied to data of comparable businesses that produce multiple outputs from multiple inputs and solves for the tightest fitting piecewise-linear convex efficiency frontier that contains all of the included observations. This is referred to as 'the technology', or in economic theory as the 'transformation set' (the set of all combinations of inputs and outputs that are 'feasible' because the inputs can produce the outputs). The boundaries of the transformation set represent the 'efficiency frontier'.

DEA can be used to measure technical efficiency by comparing a firm's use of inputs relative to the outputs it produces relative to the best observed practice in the sample (as given by the efficiency frontier). If our interest is the degree to which the use of inputs could be reduced while producing a given levels of outputs, the DEA input-oriented technical efficiency model is used. If we want to know the degree to which outputs can be increased while using the same quantities of inputs, the DEA output-oriented technical efficiency model is used. DEA can also be used to measure cost efficiency, which is the degree to which a firm is minimising its cost. Unlike the methods to estimate technical efficiency, which require data for input and output quantities, the cost efficiency approach requires information on input prices, because cost is a monetary measure.

In order to be relevant to the different available DEA approaches, this paper addresses the choices of output variables, input variables, input prices and operating environment factors.

2.1 Importance of Variable Choice

As a non-parametric method, DEA does not impose any specific functional form on the technology or production possibility set, however, variables used as inputs and outputs in the analysis are chosen by the analyst and, as with all benchmarking methods, the choice of inputs and outputs is fundamentally important to the results obtained.

DEA itself does not provide guidance on the choice of input and output variables. There is considerable reliance on the judgement and expertise of the analyst. However, it is well known that model misspecification has significant impacts on DEA efficiency estimates, especially in small samples. This can arise either by including irrelevant inputs or outputs or by omitting important variables. Even the omission of a variable that is highly correlated with another variable can have a significant impact on some of the efficiency estimates. A further difficulty is that DEA loses discriminatory power as the dimensionality of the production space increases (i.e., as the number of outputs and inputs included increases), which is an especially important problem when sample size is small. These considerations imply that the variable selection methods are important, and methods for achieving parsimony in a DEA model are necessary in the absence of a large data sample.

2.2 Approaches to Identifying Variables

The choice of candidate variables is often developed through consultation processes with industry, which can assist to prioritise the variables for inclusion and filter out those that are unlikely to have any explanatory power. Three general approaches that appear to be taken to identifying the variables are:

(a) Advice from engineering or business process experts regarding what was logical or plausible from their perspectives. This is in some sense an ideal approach, because in any economic analysis of industry technology or costs it is vital to make use of industry knowledge. This can be termed a 'first principles' approach and can also include views on which output dimensions should be included from an economic perspective.

(b) Some formal techniques are available to assist the variable selection process, and while not 'fool proof', they are arguably much better than *ad hoc* and time-consuming trial and error processes. Several DEA studies of TSOs previously undertaken for European regulators have used regression analysis prior to the DEA analysis in order to ascertain the variables having a statistically significant explanatory power in relation to costs. Several alternative statistical methods of filtering and screening variables are available, and some of them appear to have advantages over the preliminary regression approach. These approaches are surveyed in chapter 5.

(c) Since a considerable number of benchmarking and cost studies of energy TSOs have been undertaken to date, it is also feasible and desirable to consider the methodologies and variables used in the literature, particularly those that may be considered the 'conventional' or 'best practice' in DEA analysis of energy networks. Chapter 3 examines the relevant academic consultant studies in terms of the variables used, and the methods or arguments used to select them.

In any case, it remains important to have regard to the economic theory of the producer in order to ensure that the results of the benchmarking study have a meaningful economic interpretation. Otherwise, inferences that can be drawn from the benchmarking analysis will be more limited.

2.3 Principles for Variable Choice

2.3.1 Distinguishing between types of variables

The process of production is one of transforming inputs into outputs, and in modelling this it is important to clearly distinguish between the outputs and the inputs. Outputs are goods and services provided to customers, whereas inputs are goods and services used up in the process of producing the outputs. (When a durable input such as capital equipment is only partly used up in a given period, then input quantity in that period is a measure of the services provided by that durable input.) Even where a group of benchmarked firms have the same technology available to them, differences in their operating environments may affect their ability to transform inputs into outputs and hence represent an external factor that can affect efficiency. External factors of this kind can be particularly relevant to utilities because they tend to operate in discrete areas, and those areas may have different characteristics which are relevant. It is important to distinguish between operating environment factors, which are beyond management control, and inputs, which are chosen by the producer.[1] The purpose of efficiency measurement is to enable or incentivise firms to achieve best practice, but this is limited to the variables that they can control. Whereas an inefficient firm can reduce the use of inputs to produce a given set of outputs, it cannot influence the operating environment factors, which therefore have little or no bearing on true efficiency, and their effects on measured efficiency need to be removed.

In some applications it can be difficult to distinguish inputs from outputs. For example, an important feature of the services of energy networks is the distances over which the energy is transferred, which may be indicated by network length. However, this can be related to (or used as a proxy for) physical measures of capital inputs based on network capacity. Care is needed to ensure the model is grounded in a good representation of the production technology of the industry and the services it provides to customers.

In principle, improvements in the quality of outputs can be treated in the same way as output quantities, although in some cases quality variables may be measured as 'bads' (i.e. undesirable outputs, e.g. the number of outages). There may be difficulties with measuring service quality, including the related issue of measuring services such as security of supply. In order to limit the number of variables, some benchmarking studies adjust output quantity measures using indexes of service quality to obtain quality-adjusted output measures. These are among the issues considered in chapter 6.

2.3.2 *'Cost drivers'*

The analysis of cost efficiency raises special issues of its own, because input price data is required, and because care is needed to ensure that the model is consistent with economic theory. Total cost is defined as the sum of the products of the prices and quantities of each input. This is simply how cost is calculated (whether that cost is efficient or inefficient) and should not be confused with the cost function of economic theory, in which the minimum cost of supply is a function of (i.e. determined by) the quantities of outputs produced and the set of input prices, given the available technology.

[1] By 'beyond management control' we mean that the operating environment factors are exogenous for the firm. Management can still make choices in how to deal with operating environment factors (which may be more or less effective), but these responses generally require resources to implement, so that differences in operating environments can affect the observed comparative productivity and cost efficiency of firms even when action is taken to mitigate their effects. The effects of operating environment factors are an empirical question.

Although the DEA cost efficiency model is often loosely described as a 'cost driver' model, the variables used in the model should be consistent with economic theory. Cost inefficiency refers to the degree to which a firm's actual costs exceed the minimum cost of supply, which refers to the economic cost function. The 'cost driver model' terminology is used in stochastic frontier analysis (SFA) to refer to cost function analysis (e.g. Burns and Weyman-Jones, 1996). In DEA, Nieswand *et al* (2010) use the term 'cost driver model' to mean:

> … that costs are explained by output variables that are relevant to costs of the pipelines under consideration. This approach deviates from the purely technical representation of the production process by physical data but [sic] is often applied in regulatory practice (2010) (p 7).

This conception is consistent with an economic cost function (although no mention is made by Nieswand *et al* of input prices, perhaps not being relevant in that study) and the distinction is correctly made between measures of technical and cost efficiency (the latter taking into account input allocative efficiency as well as technical efficiency). This reiterates the need to carefully distinguish between inputs and outputs. There will inevitably be a high correlation between inputs and cost (because of the definition of cost), but if inputs were included as 'cost drivers' this would not yield a true measure of cost efficiency. This is because the economic functional relationship between cost and its determinants would be conflated with the definitional relationship between costs and inputs.

It is useful to note that in cost efficiency benchmarking it is common to estimate the DEA cost efficiency model in conjunction with the input-oriented technical efficiency model (which measures the degree to which the use of inputs could be reduced while producing a given levels of outputs) because this enables cost efficiency to be decomposed into technical and allocative efficiency (both input oriented). It is a useful discipline to estimate both models, because the technical efficiency model requires that an economically meaningful distinction be made between the inputs and outputs.

2.3.3 Trade-offs in the use of variables

Ideally, the inputs and outputs included in the DEA analysis should be sufficiently comprehensive to capture the relevant features of the production process, including the quality and quantity of outputs. Otherwise the measures of efficiency may be inaccurate. This means that a theory of the production process needs to be formed, which will usually combine engineering and economic perspectives. It is also preferable that each different service provided should be measured by a separate output, and each distinct factor of production should be measured using a separate input. However, unless the data sample is quite large, there will usually be practical limitations that require some compromises. This means there are challenges to undertaking DEA analysis with small samples

When the sample size is small, it may not be realistic to multiply the number of variables included in the model. There are limits to the number of variables that can be meaningfully included in a DEA analysis depending on the data sample size. Cooper *et al* (2006) have suggested, as a rule of thumb, that the product of the number of inputs and the number of outputs should not exceed the sample size and the sum of the number of inputs and outputs should not exceed one-third of the sample size. If the recent developments in statistical theory for DEA are to be used, then a much larger amount of data is needed per variable included in the model. It should also be noted that the accuracy of DEA (and its discriminative power in particular) reduces as the number of outputs becomes substantial relative to the sample size, because more firms will be found to be ‘efficient’ simply because they have unique input-output mixes and are not closely comparable to other firms.

When data sample sizes are small it will be necessary to properly narrow (e.g., by properly aggregating) the candidate variables down to a relatively small number of key outputs and inputs. One way of approaching this is to use aggregated input or output measures with the aim of completeness in reflecting all of the essential inputs and outputs of TSOs. There are risks and pitfalls with combining variables into aggregates or indexes, and suitable methods would need to be considered, Methods used for combining variables are among the issues considered in chapter 5, which also notes the risks associated with, and potential errors introduced by, inappropriate aggregation.

2.3.4 Regulatory context

To the extent that it is feasible within a given application, the choice of variables may also need to have regard to the regulatory framework in which the benchmarking will be applied, and its objectives. For example, if specific objectives such as reliability, safety or security of supply were important regulatory objectives, it may be desirable for them to be taken into account within the benchmarking analysis. More generally, attention may need to be given to the incentive effects of including each variable in the benchmarking process.

3 EXAMPLES OF BENCHMARKING STUDIES

The variables used in previous studies can provide a useful guide to the variables that should be considered when undertaking a benchmarking study.

This section will provide a literature review of:

- the academic literature on the analysis of cost structure and/or benchmarking of energy networks. Although particular emphasis will be given to studies of gas and electricity transmission businesses, studies relating to energy distribution networks will also be included where they provide useful analogies for transmission businesses.
- energy network benchmarking studies carried out by or on behalf of regulators.

The studies will be presented in a table showing the sector studied, the output variables used, the input variables used, input prices where relevant and the operating environment variables included in the study.

3.1 Studies of Gas Transmission

Table 3.1 summarises a number of cost function or productivity studies carried out for gas transmission. The discussion in section 3.1.1 briefly summarises relevant information from those studies. Sections 3.1.2 and 3.1.3 analyse the studies to draw broad conclusions.

3.1.1 Individual studies

Aivazian *et al* (1987) estimated an econometric translog production function together with factor cost-share equations, therefore requiring input prices in addition to output and input quantities. Output was measured in terms of energy throughput × distance. The four inputs were labour, compressor fuel, pipeline capital services, measured by pipeline tonnage, and compressor capital services, measured in horsepower (Hp).

The study by Sickles and Streitwieser (1992) estimated the technical efficiency of 14 US gas transmission firms over 1977-1985 using a stochastic frontier translog production function and DEA with a time-varying frontier. The output measure was volumes of gas delivered (including gas transported for third parties) multiplied by an estimate of the distance gas is transported, based on the average length of the major trunkline pipelines from the gas production sources to the major delivery points for each firm in the sample. Three input measures were included: labour, energy and capital, and in each case both quantity and price measures. The quantity of labour was the estimated number of employees (in gas transmission activities only). Energy input (i.e. the gas used in compressors) was measured in cubic feet. Two measures of capital were used: total horsepower ratings of transmission compressor stations as a proxy for compressor capital services; and tons of steel as a proxy for pipeline capital services.

Prices for labour and energy inputs were derived from labour and energy expenses divided by the quantity measures. A ‘value added methodology’ was used to derive input prices for capital services. The costs of gas purchases (for resale), labour and energy used for gas transportation were subtracted from total revenue from gas resale and from transportation of third-party gas to obtain net revenue. Then net revenue was allocated between compressors and pipelines proportionately to the book values of those assets, the resulting values being divided by the capital quantity measures to obtain prices.

Three related studies in 1999 by Lee, Kim et al (1999; 1999a, 1999b) benchmarked Korean gas utilities against international comparators and included separate benchmarking for gas transmission businesses and integrated gas businesses.[2] The studies differed in terms of the sample of utilities included and the benchmarking techniques used, but were largely consistent in terms of the assumed outputs and inputs. For both gas transmission and integrated gas utilities:

- Output was defined as total gas throughput (measured in energy units).
- Input quantities were:
  - labour, defined as the total number of employees;
  - capital, defined as total tangible fixed assets in constant prices; and
  - administrative inputs: calculated by dividing administration cost by a proxy price measure discussed below.
- Input prices were defined as follows:
  - The price of the labour input is the sum of payroll and other employee-related expenses, divided by the labour input quantity.
  - The price of capital inputs differed between studies. In Lee, Park and Kim (1999), capital cost was calculated as the sum of maintenance, depreciation, taxes (other than income tax), insurance, interest, and other capital-related expenditures, and the capital price was defined as the ratio of capital cost to the capital inputs measure. A similar method appears to have been used in Kim et al (1999). However, in Lee, Oh and Kim (1999) price of capital services was computed using a simplified version of the neoclassical user cost of capital services formula in which tax terms were omitted due to lack of data.
  - A proxy price for administrative inputs was obtained by regressing administration cost against some index of employees and length of pipelines, and the slope of that regression was used as the proxy unit price of administration.

In order to convert to a common currency, US$, a reference year was chosen (1991) and all monetary amounts were converted to US dollars using the purchasing power parity (PPP) for the reference year in the Penn World Table.

Granderson (2000) studied the effects of open access on efficiency with a sample of 20 US interstate natural gas pipeline companies from 1977 to 1989, and using an econometric translog cost function. The inputs for pipeline companies were considered to be: labour, compressor station capital and transmission pipeline capital. The chosen output measure was

[2] When benchmarking against integrated gas utilities, the Korean gas transmission business, Kogas, was combined with 26 distributors to form a hypothetical integrated business.

the volume of compressor station fuel. Prices for fuel and labour were obtained by dividing expenditures on them by their physical quantities. The quantity of compressor station capital was measured as the sum of the horsepower ratings of all compressor stations on the pipeline. The quantity of pipeline capital was estimated by the formula: $a \times d^2 \times l$; where d is the average diameter and l is the length of the pipeline, and a is a constant.[3] The user price of capital was based on the neoclassical user cost of capital formula. Total cost was a function of the output quantity and the input prices.

Although output was defined in the study as the amount of compressor station fuel used, it was noted that the "ideal output measure for natural gas transmission is the sum across all shipments of the volume times the distance transported" (Granderson, 2000, p. 259). However, data for this measure was unavailable.

Hawdon (2003) compared countries rather than businesses, and examined the efficiency of the gas industries as a whole in those countries, rather than being confined to just transmission or distribution. The sample included 33 countries. The analysis was at a high level, with two outputs and two inputs. The outputs were total customer numbers and total gas supplied/consumed. The inputs were total gas industry employees and capital services of the pipeline system.

Hawdon observed that the capital services of a pipeline is a complex matter that depends on a number of factors:

> The services of a pipeline system depend on a wide variety of factors including pipeline diameter and length, inlet and outlet pressures and the availability of compressor equipment to regulate operating pressures. Other factors affecting supply include the availability of storage capacity for seasonal and other top-up to regular supplies. (Hawdon, 2003, p. 1169)

Since data was not available for all of these features, the length of pipelines was used as the measures of capital services.

A number of operating environment variables were included in the study: (a) share of gas in total energy; (b) growth in demand; (c) reform in terms of privatisation or deregulation; and (d) responsiveness to the EU gas directive. There was little discussion of the specific rationales for these variables. Their significance was tested within a bootstrap framework and variable (a) was excluded in the final analysis. Variables (c) and (d) had negligible effect. The market growth variable was a significant one for efficiency, and this was interpreted to

[3] This formula may be interpreted in terms of physical input or service potential. In terms of physical input, the formula for the required wall thickness of a steel pipe is: $T = b \times p \times d$, where p is pressure, d is diameter and b is a constant that depends on technical standards in relation to minimum yield stress etc. Combining this with the formula for the circumference of a circle (πd), the amount of steel in a pipeline of length l is proportionate to: $\pi b \times p \times d^2 \times l$ (so that $a = \pi b \times p$). In terms of service potential, the volume flow rate of gas is: $V = \pi d^2 (v/4)$, where d is the diameter and v is the velocity of the gas (so that $a = \pi v/4$). Velocity depends on pressure, temperature, and internal friction. These derivations show that pipeline pressure could be an explicit argument in this type of formulation.

be because “efficiency improving investments occur in an expanding [market] which would be difficult to justify during periods of contracting sales” (Hawdon, 2003, p. 1172).

Two studies were undertaken by Jamasb *et al* (2007, 2008). The 2007 study benchmarked a sample of US and European gas transmission businesses. The US data was more abundant with 43 US gas TSOs included covering the period 1996 to 2004 (317 observations), compared to only 4 European gas TSOs over 2000 to 2004 (a total of 11 observations).

A preliminary econometric ‘cost-driver’ analysis was used to test the significance of the chosen variables prior to their use in DEA. It was noted that although “statistical significance does not have to be the ultimate arbitrator for the inclusion of a variable it gives important guidance especially for DEA, which cannot discriminate between relevant and irrelevant variables itself” (Jamasb et al., 2007, p. 23). The additional use of corrected ordinary least squares (COLS) and stochastic frontier analysis and consistency checks between the methods, were aimed at making the results more robust (Jamasb et al., 2007, p. 24). A literature review was also used to identify appropriate cost drivers.

Cost was used as the single input, with four alternative measures of cost tested:

- O&M (i.e., variable cost)
- O&M plus depreciation
- O&M plus depreciation and cost of capital (i.e., total cost)
- Revenue less gas sales (i.e. transportation revenue).

The outputs, or cost-drivers, included:

- Technical indicators of capacity, including: pipeline length, number of compressor stations, number of compressor units and total compressor horsepower.
- Measures of gas deliveries, including: annual gas throughput and peak day delivery (record to date × days in year), both in m^3/year. Load factor, which is the ratio of these two measures, was also used.

The study did not include input prices, which it recognized was a potential limitation, but costs were inflation adjusted (using consumer price indexes), and converted to a common currency using purchasing power parities (PPPs).

The 2008 study of Jamasb, Pollitt and Triebs benchmarked a sample of 39 US gas transmission businesses for the period 1996 to 2004 (351 observations). A preliminary econometric analysis was used to test the significance of the chosen variables prior to estimation, but not used to select among alternative variables. The authors noted: “Admittedly, our choice of variables is rather ‘ad hoc’ in the sense that we do not test alternative cost-drivers but rather verify the econometric significance of the variables at hand” (Jamasb et al., 2008, p. 3400).

The input variable was total cost, defined as O&M expenses (excluding fuel), depreciation and the cost of capital (defined as asset written-down valued × 6%). The alternate measure of total cost was total revenue from gas transported for others. The output variables were the total network length, the total horsepower of compressor stations, and the total annual throughput of gas transported for shippers (thus not including any gas sold in the end-markets by the pipeline company). Among the other variables mentioned in the study, but not used in the analysis, was network age, defined as: accumulated depreciation ÷ annual depreciation.

The use of a single monetary measure of input (total cost or revenue) was considered by the authors to have several advantages.[4] Firstly, the "trade-offs between the various inputs are accounted for", and secondly, the resulting efficiency measures "have incentive properties different from standard technical efficiency measures" (Jamasb et al., 2008, pp. 3407–8).

Nieswand *et al* (2010) use principal components analysis (PCA) in conjunction with DEA (section 6.2 discusses the PCA-DEA methodology) to estimate the cost efficiency of a group of 37 US gas transmission companies. The cost driver model uses cost (opex) as the only input and chooses as output the variables most relevant to costs. The authors noted that the 'cost driver' approach "deviates from the purely technical representation of the production process by physical data but is often applied in regulatory practice" (p 7). Two models were tested. The first used as outputs: total natural gas delivered; transmission system length; peak deliveries; total compressor station capacity (Hp). The second model also included transmission system losses as a 'bad' output.

Related studies of gas transmission efficiency benchmarking by Sumicsid (Agrell et al., 2014) and Sumicsid and Swiss Economics (Agrell et al., 2016) were carried out on behalf of the Council of European Economic Regulators (CEER). The first of these, the PE2GAS study, was a feasibility analysis. It proposes a 'size of grid' or 'Net Volume' measure: $\sum_k N_k v_k$, where N_k is the number of assets of type k, and v_k is "the relative costs of these assets" (Agrell et al., 2014, p. 38). This measure can be applied separately to groups of asset types, to obtain a Net Volume measure for each group, or applied to all asset types to obtain a single Net Volume measure. The second study (E2GAS) carried out benchmarking study of gas TSOs using data for 13 businesses in 2010 and 9 businesses in 2014. The input was Totex, although adjusted for gas purchases and other factors. The output variables were:

- 'Normalized Grid' (which corresponds to variable described as 'Net Volume' above),
- the number of connection points to the transmission grid; and
- peak capacity: the maximum of total injection and total delivery capacities. Total injection capacity was measured as the highest concurrent hourly total of injections at all injection points (nm^3/hour). Total delivery capacity was measured as the highest concurrent hourly total of deliveries at all delivery points (nm^3/hour).

Lastly, ACM commissioned Frontier and Consentec (2016) to produce a benchmarking study of 14 European gas TSOs using 2012 data. This study used DEA and derived three final models, each with three outputs (or cost drivers) and one input, but with slightly different output specifications. The variables used as cost drivers in the final analysis were:

- the number of connection points to the transmission grid (used in Models A, B & C)
- pipeline volume, defined as the total physical volume of the pipelines taking into account their lengths and diameters: (used in Models A & B)

[4] The use of revenue as a measure of input is not intuitive, but was intended to be a proxy for total cost in a context where revenue is subject to regulatory constraints.

- supply area, defined as "defined as the area of the convex hull of the entry and exit points" (p. 38) (used in Models B & C)
- Transport Momentum, a logistics concept which in the "simplest case of a direct point-to-point pipeline, the transport momentum is the product of the throughput (maximum of feed-in and withdrawal in [m³/h]) and the distance between entry and exit point (transport distance in [m])" (p 38): (used in Model C)
- $\sqrt{(\text{Transport Momentum} \times \text{supply area})}$ (used in Model B).

Table 3.1: **Studies of Gas Transmission Costs, Productivity or Efficiency**

Study	***Method***	***Data***	***Inputs/cost***	***Outputs***	***Input Prices***	***Operating environment variables***
Aivazian et al (1987)	Econometric translog production function	14 US gas transmission companies, 1953-1979	- No. employees - Fuel - Line-pipe capital services (tonnage) - Compressor capital services (Hp)	Cubic feet-miles (volume of gas delivered × distance transported)	- Labour expenses / employees - Fuel expenses / fuel quantity - (Transmission Revenue less labour and fuel expenses) × (allocation between pipeline and compressors) ÷ compressor Hp or pipeline tonnage	
Sickles and Streitwieser (1992)	SFA (translog production function) & DEA (with time-varying frontier)	14 US gas transmission firms, 1977-1985	No. employees Energy used in transportation (cu.ft) Compressor capital – Total compressor horsepower Pipeline capital – tons of steel pipe.	Cubic feet-miles (volume of gas delivered × distance transported)		

Study	***Method***	***Data***	***Inputs/cost***	***Outputs***	***Input Prices***	***Operating environment variables***
Lee, Park & Kim (1999b)	Multilateral productivity and profitability indexes	Two samples: (i) 9 gas transmission businesses from USA (6), Belgium Germany & Korea; (ii) 19 integrated gas utilities from USA (3) Canada (4), Japan (9) France, Italy & Korea. Both samples for 1987-1995.	- No. employees - Administrative inputs (derived from regressing O&M cost against no. employees and pipeline length to obtain administrative input proxy price and dividing administrative costs by this proxy) - Gross capital stock (tangible fixed assets) at constant prices	Total gas throughput delivered to end users (10^9 kcal)	Labour: employment-related expenses ÷ employees Admin inputs: admin. input proxy price (described above) Capital: Interest, depreciation and maintenance costs divided by capital stock.	
Lee, Oh and Kim (1999a)	FGLS regression of translog variable cost function	Integrated gas utilities from USA (2) Canada (4), Japan (3) France (1), Italy (1) & Korea (1), 1987-1995 (104 obs)	As above	Total gas throughput delivered to end users (10^9 kcal)	As above, but input price for capital calculated differently. Here using $P(r+d)$, where P is an index of capital goods, r and d are the interest and depreciation rates.	
Kim et al (1999)	DEA, Multilateral TFP index, Multilateral Edgeworth managerial index	9 gas transmission and 19 integrated gas utilities in 8 countries	As above	Total gas throughput delivered to end users (10^9 kcal)		

Study	***Method***	***Data***	***Inputs/cost***	***Outputs***	***Input Prices***	***Operating environment variables***
Granderson (2000)	Econometric translog cost function	20 US interstate natural gas pipeline companies 1977 to 1989	Total cost	Compressor station fuel	Price of labour Price of fuel User cost transmission pipeline capital User cost compressor station capital inputs.	
Hawdon (2003)	DEA	Integrated gas industries in 33 countries	No. employees Network length	Gas supplied No. customers		Share of gas in total energy Growth in demand Reform (privatisation or deregulation) Responsiveness to EU gas directive
Meyrick and Associates (2004)	Multilateral TFP	Gas transmission: 1 New Zealand and 7 Australian, 2003 Gas distributors: 4 New Zealand and 10 Australian (not detailed here)	Real opex Capital quantity (km of pipeline)	Throughput System capacity (proxied by asset value)		

Study	***Method***	***Data***	***Inputs/cost***	***Outputs***	***Input Prices***	***Operating environment variables***
Jamasb et al. (2007)	DEA, SFA, COLS	43 US gas TSOs, 1996 to 2004, and 4 European gas TSOs, 2000 to 2004 (328 observations).	Cost, alternatively measured as: - O&M (i.e., variable cost) - O&M plus depreciation - O&M, depreciation and cost of capital (i.e., total cost) - Revenue less gas sales (i.e. transportation revenue)	Pipeline length No. compressor stations* No. compressor units* Compressor capacity (Hp)* Annual gas throughput (m3)* Peak day delivery (record to date × days in year) (m3/year) Load factor* * excluded from final model		
Jamasb, Pollitt and Triebs (2008)	DEA (Constant Returns to Scale)	39 US gas transmission businesses, 1996-2004	Total cost (excl. fuel) Total revenue (as an alternative)	Network length Compressor capacity (Hp) Gas deliveries		
Nieswand, Cullmann, and Neumann (2010)	PCA-DEA cost driver model	37 US gas transmission businesses in 2007	Opex	Total deliveries (Dth)* Peak deliveries (Dth)* Network length (km)* Total compressor horsepower (Hp)* Transmission line losses (Dth)* *Combined into a single input using PCA		

Study	***Method***	***Data***	***Inputs/cost***	***Outputs***	***Input Prices***	***Operating environment variables***
Frontier & Consentec (2016)	DEA	14 European gas TSOs, 2012	Total costs (measured equivalent to regulatory costs)	No. grid connection points Pipeline physical volume Supply area Transport Momentum Root Transport Momentum area		
Agrell, Bogetoft and Trinkner (2016)	DEA & Unit cost analysis	13 European TSOs in 2010 and 9 European TSOs in 2014	Totex (opex + annuity-standardised capex)	-‘Normalized Grid’ - No. grid connection points - Peak capacity (maximum of total injection and total delivery capacities)		Adjustments for some environmental complexities such as land use, topography, soil structure and humidity

3.1.2 Methods of Analysis

Table 3.2 presents a summary of the methodologies used in the gas transmission studies presented in Table 3.1. Approximately one in five of the studies surveyed used more than one method, in which case they appear twice in the table. The main observations are:

- DEA analysis, or some variant such as PCA-DEA, was used in 47 per cent of all the gas TSO studies. Econometric production or cost function analysis was used in 30 per cent of the studies.[5]
- DEA analysis was used in 30 per cent of the studies up to 2005 and 71 per cent of the studies after 2005 (including PCA-DEA). Econometric production or cost functions were used in 40 per cent of the studies up to 2005, and 14 per cent of the studies after 2005. Multilateral TFP indexes were used in 30 per cent of the studies up to 2005.

This indicates that the studies carried out after 2005 relied mostly on DEA analysis whereas in the period up to 2005, econometric analysis and multilateral TFP indexes were each used as frequently as DEA.

Table 3.2: **Methods of Analysis used in Gas TSO Studies**

	Up to 2005		*After 2005*		*Total*	
Method	*No.*	*%*	*No.*	*%*	*No.*	*%*
DEA (static)	3	30	4	57	7	41
DEA (dynamic)	0	0	0	0	0	0
PCA-DEA	0	0	1	14	1	6
Econometric cost function[a]	2	20	1	14	3	18
Econometric production function	2	20	0	0	2	12
Multilateral TFP index[b]	3	30	0	0	3	18
Unit cost analysis	0	0	1	14	1	6
Total	10	100	7	100	17	100

Notes:

(a) The term 'econometric' here includes ordinary least squares, stochastic frontier analysis and/or corrected or modified least squares. The majority of these applications are stochastic frontier analysis.

(b) 'Multilateral TFP indexes' refers to the index number method introduced by Caves, Christensen and Diewert (1982).

3.1.3 Variables Used in Gas Transmission Studies

Table 3.3 lists the output variables used in the studies of gas transmission. The most frequently used output variables were:

[5] The term 'econometric' here includes ordinary least squares, stochastic frontier analysis and/or corrected or modified least squares. The majority of these applications are stochastic frontier analysis.

- gas throughput (used in 8 studies);
- the distance over which gas is transported is, in some cases proxied by pipeline length (3); while some studies use combined throughput-distance measures (3 including 'Transport Momentum');
- several physical capacity-related measures, such as: pipeline volume (1); compressor capacity (3); and number of compressor stations (1);
- Peak demand measures (3, one of which was a ratcheted measure);
- the number of grid connection points (2).

The remaining variables were each used in only one study.

Table 3.3: **Output Variables used in Gas Transmission Studies**

Variable	*No.*	*%*
Gas throughput	8	24
Compressor capacity (Hp)	3	9
Pipeline length	3	9
Gas volume x distance	2	6
Grid connection points	2	6
Peak load	2	6
Peak load (record to date)	1	3
'Transport Momentum' (a)	1	3
Asset value (system capacity)	1	3
Compressor station fuel	1	3
Line losses	1	3
Load factor	1	3
No. compressor stations	1	3
No. compressor units	1	3
No. customers	1	3
Pipeline physical volume	1	3
Root Transport Momentum-Area	1	3
Supply area	1	3
'Normalized Grid'	1	3
Total	33	100

(a) Related to volume x distance

This summary suggests the most commonly used output variables are: firstly, gas throughput and transport distance measures either included separately or combined into a single volume-distance measure; and secondly, some measure(s) of maximum delivery capacity, either using a peak day demand measure or a physical supply capacity measure.

Table 3.4 lists the input variables used in the studies of gas transmission. The most frequently used input variables were:

- the number of employees (used in 6 studies);
- several physical capacity measures, including pipeline tonnage, pipeline volume (length x diameter^2) and pipeline length (together used in 4 times); and compressor capacity (used in 2);
- real fixed assets (3)
- administrative inputs (3);
- opex (3);
- Total cost or annuity-based Totex (together used in 5 studies).

There is a distinct difference between the variables commonly used as inputs in the studies from the period up to 2005 and those after 2005. The latter generally used a cost-related measure such as opex or opex plus depreciation (3 studies), total cost (incl. annuity based) (4) or 'value-added' estimate of total cost (i.e. transportation revenue) (2 studies). Only two studies used any of these variables in the period up to 2005. In the earlier period there was a greater use of variables such as employee numbers and either a monetary-based measure of capital inputs such as real fixed assets, or physical measures of capital inputs.

Table 3.4: **Input Variables used in Gas Transmission Studies**

Variable	*No.*	*%*
No. employees	6	19
Total cost	4	13
Total cost (annuity-based)	1	3
Administrative inputs (est.)	3	10
Opex	3	10
Real fixed assets	3	10
Compressor capacity (Hp)	2	6
Pipeline length	2	6
Transportation revenue (VA est.)	2	6
Compressor fuel	1	3
Fuel	1	3
Pipeline capital services (tonnage)	1	3
Opex + capital depreciation	1	3
Pipeline length x diameter^2	1	3
Total	31	100

It is also notable that pipeline capacity-related measures have appeared as outputs in some studies and inputs in others. For example, compressor capacity was used as an output in three studies and as an input in two studies. Pipeline length was used as an output in three studies, and as an input in two studies. A pipeline physical volume measure was also used as an output in one study and an input in another.

This reflects a general challenge of measuring capital inputs distinctly from measures of customer services that are produced by capital facilities. In principle, capital inputs are productive services of the capital goods employed, which are often proxied by some measure of stock of capital employed. Outputs are the services provided to customers, and some of these, such as the distance over which gas is transported, or the ability to deliver gas volumes on the peak day, are closely related to attributes of the pipeline network length and capacity. This measurement issue needs to be given careful consideration in a benchmarking study.

Table 3.5 lists the input price variables used in the studies of gas transmission. Input prices are typically derived from other information, rather than being directly available data. Firm-specific estimates were obtainable for labour and fuel prices. Most of the other input prices related to capital inputs. Only one study used the conventional neoclassical user cost of capital formula, whereas three studies used unconventional methods for deriving a firm-specific capital inputs price.

It is also notable that none of the studies published in the period after 2005 included any input prices. This corresponds to the much greater reliance on cost as a single input in the studies undertaken in this later period.

Table 3.5: **Input Price Variables used in Gas Transmission Studies**

Variable	*No.*	*%*
Labour avg. cost	4	29
Admin inputs proxy price	2	14
Fuel avg. cost	2	14
Interest, cap. depr. & maint. exp. / capital stock	1	7
User cost of capital (est.)	1	7
User cost transmission pipeline capital	1	7
User cost compressor station capital	1	7
User cost of compressors (VA est.)	1	7
User cost of pipelines (VA est.)	1	7
Total	14	100

Only four operating environment variables were explicitly included in these studies, all of them in one study. These variables were: growth in demand; reform (privatisation or deregulation); responsiveness to EU gas directive; and share of gas in total energy. In addition, one study used supply area as an output, which might alternatively be viewed as an operating environment characteristic. Another study (Agrell et al., 2016), when calculating the 'Normalized Grid' variable, made adjustments for certain environmental complexities such as land use, topography, soil structure and humidity.

3.1.4 Conclusions

Thirteen studies of gas transmission cost functions or production technology published between 1987 and 2016 have been reviewed. Some of the main conclusions that can be drawn from this review are as follows. Firstly, a wide variety of different inputs and outputs have been used in these studies, but there appears to be a quite distinct difference between the studies conducted in the period up to 2005 and the post-2005 studies. The latter have relied almost exclusively on total cost, or total variable cost, as a single input, whereas the earlier studies tended to rely on separate measures of non-capital inputs (usually proxied by employee numbers) and capital inputs (measured either using physical capital measures or deflated monetary measures such as real fixed assets). Corresponding to this shift, although the use of input prices was quite common in the earlier studies, none of the studies in the post-2005 period used any input prices.

There doesn't appear to be such a clear pattern of difference in the use of output measures in the periods before and after 2005. Although a wide range of output measures have been used, the most common are:

- gas throughput and transport distance measures either included separately or combined into a single volume-distance measure
- some measure(s) of maximum delivery capacity, either using a peak day demand measure or a physical supply capacity measure.

However, a number of studies used as outputs, variables that appear as inputs in other studies. There appears to be confusion in the categorisation of some capacity-related variables as inputs or outputs. In part this reflects difficulties in deriving proxy measures for capital input services and for aspects of customer services that are capacity-related, such as supply security and ability to meet peak day demand.

Two other general observations are that:

- There was considerably more reliance on DEA as an analytical method in the period after 2005, whereas in the earlier period there was a relatively greater use of econometric methods (which were more frequent that DEA in the period up to 2005) and multilateral TFP analysis.
- Very few of the studies took account of differences in the operating environments of gas TSOs, either by including such variables within the analysis or by conducting a second-stage analysis of efficiency scores against operating environment characteristics.

3.2 Studies of Electricity Transmission

Table 3.6 summarises a number studies of electricity transmission cost functions or productivity. There are comparatively few published studies of this kind. This is in marked contrast to electricity distribution, where we found approximately 100 benchmarking studies published since 1990 (not reviewed here). Data limitations may be a factor, since many jurisdictions have only a small number of electricity TSOs. Even for Europe, Agrell and Bogetoft (2007) found that there was insufficient data to benchmark electricity TSOs at that time. However, data limitations are not so apparent for the USA.

The discussion in section 3.2.1 briefly summarises relevant information from the studies presented in Table 3.6. Analysis of the methods and variables used in those studies is presented in sections 3.2.2 and 3.2.3.

3.2.1 Individual studies

The early studies of electricity transmission costs, for example Huettner and Landon (1978) and Pollitt (1995) were primarily interested in examining economies of scale as part of an inquiry into whether TSOs are natural monopolies. These two studies estimated a variable cost function using ordinary least squares. Huettner and Landon considered utility costs to be a function of peak delivery capacity, the rate of capacity utilisation (i.e. the average quantity of deliveries divided by capacity), input prices and transmission line length. These models explained only a small amount of the variation in the samples used.

Gilsdorf (1994) sought to determine whether there is any substantial cost efficiency from vertical integration between electricity generation and distribution, because if so, this could have implications for any natural monopoly test. Transmission output was seen as "simply the path of the transmission system, which is a function of circuit mileage and voltage capacity" (p.265). The product of average voltage and circuit-miles, called 'circuit voltage miles' (circuit length × kV) was used as the transmission output measure.

Dismukes *et al* (1998) estimated an opex cost function for transmission, again for the purpose of assessing economies of scale, but used a significantly larger sample than the previous studies. These authors felt that several of the earlier cost function specifications were rather *ad hoc*, and they included a more disaggregated set of explanatory variables. Their model explained a higher proportion of the variation in their sample. Dismukes *et al* rejected Pollitt's capacity-distance measure—circuit length × kV—because "distance has nothing to do with capacity", whereas:

> ... transformers serve to connect different circuits and can be viewed as a facilitator, or bottleneck, for moving power. Thus, the transformer capacity in a transmission system is probably a much more reliable measure of overall transmission system capacity than some arbitrary product of voltage and distance. (Dismukes et al., 1998, p. 157)

The key explanatory variables for variable cost used by Dismukes *et al* were:

- Transmission system power flows in MWh (separately for inward, outward; inward wheeling; and outward wheeling flows);[6]
- Wage costs ($)
- Transmission line circuit miles (separately for underground and overhead lines)
- Transformer capacity (MVA).

Two studies by von Geymueller (2007, 2009) estimated the technical efficiency of electricity TSOs using the Dynamic DEA method. The first used a sample of 7 European TSOs for the

[6] The term "wheeling" means flows out of or into an adjacent transmission network.

period 1999 to 2005, and the second used a (larger) sample of 50 US TSOs for the period 2000 to 2006. Dynamic DEA treats certain assets such as transformers and transmission lines as 'quasi-fixed' inputs that cannot be adjusted instantaneously, and switch from being outputs in one period to being inputs in the next period. In the 2007 study the variable input was employee numbers, and the quasi-fixed input was transformer capacity, and electricity demand was used as a proxy for the amount transmitted. In the 2009 study, the variable inputs were materials and supplies expenses, and salaries and wages. The quasi-fixed inputs were line length (miles) and total transformer capacity (MVA). The output was transmission of electricity for others (MWh). The paper did not discuss the reasons for the choice of variables. The output variable would seem inappropriate if any businesses were vertically integrated with generation facilities, and had their own electricity transmission requirements.

Cadena *et al* (2009) is a productivity benchmarking study of Columbian electricity TSOs with the purpose of assisting the economic regulator to determine efficient revenue caps. This study has the considerable merit of explicitly developing an economic model of electricity transmission in mathematical form, which is used to guide the choice of variables. For example, one of the preferred input variables was the summation over all lines of wire length × wire thickness (in MCM, ie, thousands of circular mils), because this is closely related to the power delivery capacity. Other inputs included non-electrical assets and opex. The two output variables were the quantity and quality of the energy service:

- Although energy delivered × distance would have been the preferred measure of quantity, the necessary data was unavailable, and a capacity variable was used instead: the summation over all lines of power capacity (in MVA) × length.
- Quality of the energy service was proxied by average available hours.

The first of these two output variables seems problematic because it is closely related to the network capacity input variable. The operating environment variables were of two kinds: three measures of the extent of exposure of assets to salinity in coastal areas; and a measure of network complexity. Since network complexity and configuration indexes are of particular interest, the three measures mentioned in the paper are reproduced below (although it is not clear whether they were alternative measures or combined in some way):

(1) $$Substation\ Complexity = \frac{No.Bays}{No.Substations}$$

(2) $$Area\ Complexity = \frac{Service\ Area}{Total\ lines\ length}$$

(3) $$Network\ Complexity = \frac{No.Substations}{Total\ lines\ length}$$

Agrell and Bogetoft (2009, 2006, 2014) have carried out a series of benchmarking studies of the cost efficiencies of European TSOs on behalf of European regulators. A common feature of the methodologies used in these studies is that they relate costs to indexes of physical assets, rather than to outputs and input prices, as used in cost functions of economic theory.

In the 2006 study (the ECOM+ model) actual opex and standardised capex[7] were compared to two measures of the 'size' of the grid, one defined in a way most relevant to opex requirements ($Size^{Opex}$), and the other defined in a way most relevant to capex ($Size^{Capex}$). Unit cost measures could then be defined as: $Opex/Size^{Opex}$, and $Capex/Size^{Capex}$, and the overall measure: $UC = Totex/(Size^{Opex} + Size^{Capex})$. The relative efficiency measure for a TSO i was then defined as: $E_i = UC^{min}/UC_i$, where UC^{min} is the smallest unit cost in the sample of TSOs. By implications there was two inputs (opex and standardised capex) and two outputs ($Size^{Opex}$ and $Size^{Capex}$).

An important question about this model is what the 'size' variables actually measure. The $Size^{Opex}$ measure for a given firm in a particular year was defined as: $\sum_a N_a w_a$, where N_a is the number of assets of type a owned by the firm, and w_a is "the minimal cost of operating and maintaining one unit of asset a" (p.26). Therefore the $Size^{Opex}$ measure is actually an estimate of the minimal cost of operating and maintaining all of the firm's assets, and not a measure of network size at all. The same issue applies to the $Size^{Capex}$ variable, which is defined using the same annuity summation formula over the firm's historical investment series as is used for standardised capex. The only difference is that the firm's actual investment in year s, is replaced with: $\sum_a n_{a,s} v_a$, where $n_{a,s}$ is the number of assets of type a purchased by the firm in year s, and v_a is "the minimal costs of installing one unit of asset a" (p.26). This is an estimate of the capital cost if past investments had been made at minimal cost, and is not a measure of network size.

In the development of country-specific weights, several operating environment variables were identified, but do not appear to have been used in the final analysis:

- climate (e.g. mean temperatures, air salinity and air humidity)
- types of landscape (5 categories)
- population density
- abnormal safety regulations
- abnormal environmental restrictions, delays or public demands
- quality of supply
- interconnectedness of the country
- location of sources of generation and load
- pricing and universal service obligations
- country specific and international market structures in generation and consumption.

Agrell and Bogetoft's (2009) study (e[3]GRID) examined the cost efficiency of 22 European electricity TSOs. The primary benchmarking method was DEA, unlike the previous study. The single input was Totex, defined as opex plus standardised capex, using the same annuity method of standardisation as used in their 2006 study. The three outputs were:

[7] Standardization refers to the application of an annuity formula to the historical capex series, which provides an estimate of the economic Cost of Capital employed.

- 'Normalised Grid', a physical network measure "based on a system of techno-economic weights, [and] permits to take into account over 1200 different assets in eight groups, differentiated with respect to voltage, power, current, cross-section, complexity and other relevant dimensions" (p. ii). It is also referred to as a "cost norm".
- Customer density: reflecting a number of additional construction and maintenance costs, as well as greater network design complexity, associated with higher density areas;
- Connected generating capacity from renewable sources: reflecting additional grid costs to provide for stability of power and cater for highly intermittent power flows.

The Normalised Grid measure was the most important output, and was equal to: $Size^{Opex} + Size^{Capex}$. The weights used in these measures appear to represent average European costs (e.g. of maintaining and operating one unit of the asset) rather than the minimal cost previously used.

A similar study was carried out again in 2013 by a consortium of Frontier Economics, Consentec and Sumicsid (2013). The study used weighted DEA, simple log-linear OLS and robust-OLS regression. Again, there was one input, Totex, and the three cost drivers in this instance were:

- Normalised Grid;
- the total area (in km2) of highest population density;
- the totex-weighted line length of angular towers.

The latter two variables were described as operating environment variables. As with the previous two studies, benchmarking was limited to pure 'wire-company' functions.

Lastly, there were two studies by Llorca *et al* in 2013 and 2014. Both used a sample of US TSOs for 2001 to 2009. The first study estimated the cost function for electricity transmission using a translog specification and stochastic frontier analysis. The specification was based on "the basic economic theory of production and the literature on electricity networks" (Llorca et al., 2013). The authors noted Ofgem's view that the four key cost drivers for electricity transmission were: network length, peak demand, energy delivered, and asset age (Ofgem, 2011). The dependent variable was total economic costs (or 'totex') defined as "the sum of Opex, which includes operation and maintenance expenses incurred by the company over one year, and Capex, which is the sum of annual depreciation on capital assets and the annual return on the balance of capital" (Llorca et al., 2013). The explanatory variables were:

- Two outputs: peak demand (maximum peak load of the year during 60 minutes) and annual energy delivered;
- Network length (in pole miles);
- Transmission plant 'additions', to take account of the fact that network capital cannot be adjusted instantaneously;
- Input prices for labour and capital inputs, based on published price indexes average rates and producer prices for electricity networks by US state.

A number of environmental variables were used, including:

- three weather variables; and the weather variables were also interacted with the average capex/opex ratio of each business;
- measures of the average growth of demand for each firm.

The second study by Llorca *et al* investigates the benefits of the latent class model (LCM), used to divide the sample into a number of groupings, where each group is associated with a somewhat different technology due to the influence of unobserved variables. The LCM involves estimating the model in two stages: the first stage is used to search for the best number of latent classes to use and to identify which class each firm belongs to. In this study the first-stage analysis involves linear regression. Once the firms are assigned to classes, the benchmarking analysis is used in the second stage. This method can be applied with both SFA and DEA, and the latter is used in this study.

The variables used in the study do not include input prices because Totex is used as the only input in a DEA analysis. The outputs are slightly different to the first study, including: peak demand; annual energy delivered; and network length; and a new output: total energy of the system, which includes "total net own generation, total purchases from others, net exchanges in the system (received–delivered), net transmission for others and transmission by others" (p 12). A variant to the standard model incorporates four operating environmental variables: the three weather variables and the growth in demand. However, the study found that "a simple latent class model is able to control for heterogeneity in firms' operating environment without explicitly including environmental variables" (p 15).

The Economic Insights (2014) study is discussed in section 4.2.

Table 3.6: **Studies of Electricity Transmission Costs, Productivity or Efficiency**

Study	***Method***	***Data***	***Inputs/Cost***	***Outputs***	***Input Prices***	***Operating environment variables***
Huettner and Landon (1978)	Econometric cost functions. Separate models for generation, transmission, distribution and administration. Only transmission considered here.	74 US electricity utilities, 1971	Transmission expense ($ per kWh delivered)	- Total capacity (MW capacity of company's generators) - Utilization of capacity (%) - Underground line length (circuit miles) - Overhead line length (structural miles)	-Company wage cost ($/hour)	- Regional indicator variables - Indicators for holding company - Commercial (% of total) - Industrial (% of total) - Utility/municipal (% of total)
Gilsdorf (1994)	Econometric multiproduct translog cost function, for both generation and transmission combined.	72 US electricity utilities, 1985	Total cost (Variable cost plus capital user cost × capital stock), generation and transmission	Transmission output: - Circuit miles × Average voltage Generation output: - MWh produced	- Labour (employee expenses ÷ estimated FTE employees) - Fuel (fuel expenses / MMBTU) - Capital services (index of capital user costs for generation, transmission, distribution and general plant)	Distribution density
Pollitt (1995)	DEA	129 US utilities in 1999	No. employees Circuit length × kV Energy losses	Energy delivered Maximum demand Route km		Public / Private ownership

Study	***Method***	***Data***	***Inputs/Cost***	***Outputs***	***Input Prices***	***Operating environment variables***
Dismukes et al (1998)	Translog cost model	US investor-owned utilities, 1986-1991 (805 obs)	Opex	Transmission system power flows: inward; outward; inward wheeling; and outward wheeling (MWh) Wage costs ($) Transmission line circuit miles: underground; and overhead. Transformer capacity (MVA)		Regional indicator variables Indicators for years
Agrell and Bogetoft (2006)	Average unit cost	6 European TSOs, 2000-2003	Totex (opex + annuity-standardised capex)	Grid size measure – opex Grid size measure - capex		
von Geymueller (2007)	Dynamic DEA	7 EU utilities 1999-2005	Variable input: - No. employees Quasi-fixed input: - Transformer capacity (MVA)	Domestic demand (TWh)		
von Geymueller (2009)	Dynamic DEA	50 US electricity transmission businesses, 2000 - 2006	Variable inputs: - materials & supplies ($) - salaries and wages ($). Quasi-fixed inputs: - transmission line length (miles) - transformer capacity (MVA).	Transmission of electricity for others (MWh).		

Study	*Method*	*Data*	*Inputs/Cost*	*Outputs*	*Input Prices*	*Operating environment variables*
Cadena et al (2009)	DEA and SFA	7 Columbian transmission utilities, 2001 - 2004	- Quantity of electric conducting material proxied by either: ▪ summation over all lines of cross-section (MCM) × length ▪ regulatory asset value - Non-electrical assets proxied by non-electrical assets depreciation + rental fees - Opex.	- Nominal transport capacity: summation over all lines of MVA × length - Quality: weighed average of the number of available hours of each line		- Length of lines exposed to salinity in coastal regions - No. substation bays exposed to salinity in coastal regions - Value of electrical assets exposed to salinity in coastal regions - Infrastructure Complexity index
Agrell & Bogetoft (2009)	DEA	22 European transmission system operators from 19 different jurisdictions, 2003-2006	-Totex (opex + annuity-standardised capex)	- Normalized grid measure - Connection density - Capacity of connected power for renewable energy (including hydro).		

Study	*Method*	*Data*	*Inputs/Cost*	*Outputs*	*Input Prices*	*Operating environment variables*
Economic Insights (2014)	Data definition and collection, and variable choices, for Multilateral TFP	5 Australian electricity TSOs	- Opex (total opex deflated by a composite labour, materials and services price index) - Overhead lines (proxied by overhead MVAkms) - Underground cables (proxied by underground MVAkms) - Transformers and other capital (proxied by transformer MVA).	- Energy throughput - Ratcheted maximum demand - Voltage–weighted entry and exit connections - Circuit length - Energy not supplied (minus)		

Study	*Method*	*Data*	*Inputs/Cost*	*Outputs*	*Input Prices*	*Operating environment variables*
Llorca, Orea & Pollitt (2013)	SFA	59 US electricity transmission businesses, 2001-2009 (405 obs)	Totex (= opex + capital cost)	Outputs: o Peak demand (maximum hourly load) o Energy delivered System size variable: o Network length (pole miles) 'Investment proxy': o total transmission plant 'additions'	- Labour price (average wage for electricity transmission and distribution by state) - Capital price (producer price index for power transmission by state).	Included in SFA efficiency term: o Weather: • minimum temp • rainfall • avg wind speed o Capex/Opex (interacted with weather) o Growth of demand: • positive growth • negative growth
Llorca, Orea & Pollitt (2014)	Latent Class Model & DEA	59 US electricity transmission businesses, 2001-2009 (405 obs)	Totex (= opex + capital cost)	o Peak demand (maximum hourly load) o Energy delivered o Total energy of the system o Network length (pole miles)		
Frontier et al (2013); Agrell & Bogetoft (2014)	DEA	21 TSOs from 19 EU countries, 2007-2011 (102 obs)	-Totex (opex + annuity-standardised capex)	Normalised Grid measure Densely populated area Value of weighted angular towers		

3.2.2 Methods of Analysis

Table 3.7 presents a summary of the methodologies used in the studies surveyed in the previous section. Approximately one in five of the studies surveyed used more than one of the methods listed, in which case they appear twice in the table. Some form of DEA analysis was used in 50 per cent of the electricity TSO studies surveyed. Econometric cost functions were not as frequently used as DEA in the electricity TSO studies (36 per cent overall). Multilateral TFP indexes and unit cost analysis were used much less frequently.

Although the number of studies is comparatively small, Table 3.7 shows that DEA analysis was used in 25 per cent of the studies up to 2005 and 60 per cent of the studies after 2005. Econometric production or cost functions were used in 75 per cent of the studies up to 2005, and 20 per cent of the studies after 2005. Multilateral TFP indexes and unit cost analysis were each used in 10 per cent of the studies after 2005. These observations are consistent with the general pattern observed for benchmarking studies of gas TSOs, where there has been a shift toward relatively greater reliance on DEA analysis after 2005 compared to the period up to 2005.

Table 3.7: **Methods of Analysis used in Electricity TSO Studies**

	Up to 2005		*Post-2005*		*Total*	
Method	*No.*	*%*	*No.*	*%*	*No.*	*%*
DEA (static)	1	25	4	40	5	36
DEA (dynamic)	0	0	2	20	2	14
PCA-DEA	0	0	0	0	0	0
Econometric cost function[(a)]	3	75	2	20	5	36
Econometric production function	0	0	0	0	0	0
Multilateral TFP index[(b)]	0	0	1	10	1	7
Unit cost analysis	0	0	1	10	1	7
Total	4	100	10	100	14	100

Notes:

(a) The term ‘econometric’ here includes ordinary least squares, stochastic frontier analysis and/or corrected or modified least squares. The majority of these applications are stochastic frontier analysis.

(b) ‘Multilateral TFP indexes’ refers to the index number method introduced by Caves, Christensen and Diewert (1982).

3.2.3 Variables Used in Studies of Electricity Transmission

Table 3.8 lists the output variables used in the studies of electricity transmission discussed in section 3.2.1. There is a wide range of output variables used in the studies, and apparently, not much consensus on the outputs of the electricity transmission industry. Table 3.8 shows 35 different output variables were used as outputs in the 12 studies of electricity TSOs, and only a few of them were used in more than one study. The wide range of variables used suggests that there is a lack of agreement on how to specify the outputs of electricity TSOs.

Since, the key functions of electricity TSOs are broadly similar to those of gas TSOs (namely transmission of energy over distances and ensuring enough capacity to supply on peak days and hours), Table 3.8 shows sub-totals for variables relating to: (a) energy delivered; (b) transmission distances; (c) the product energy and distance; and (d) measures of peak day demand (which are common types variables in gas TSO studies).

The most frequently used output variables for electricity TSOs were:

- Energy throughput (used in 5 studies). In addition, one study splits energy throughput into four variables relating to inward and outward energy flows and corresponding wheeling measures. Another used total domestic electricity demand as a proxy for power flows;
- Peak demand measures (3, one of which was a ratcheted measure) or maximum demand (1 study);
- Network length measures, defined variously including pole miles or circuit miles for the network as a whole (3 studies), or separate measures for overhead and underground transmission lines (2 studies), or 'route km' (1 study);
- Energy or capacity (MVA) times distance was used in two studies; and
- Grid 'size' measures (including 'Normalized Grid') were used in three studies.
- The remaining variables were each used in only one study.

Table 3.9 lists the input variables used in the studies of electricity transmission. The two most commonly used types of approaches appear to be:

- Using opex or number of employees (5 studies overall) as measures of non-capital inputs, and using a physical measure of capital inputs such as: transformer capacity (2) line cross section × length (1) or line length (1); or a monetary measure of capital inputs such as regulatory asset value (1);
- Using a single input such as total cost or totex measure (4 studies), or in some cases opex, and using a combination of network capacity and service delivery measures as outputs.

Again, some of the variables used as inputs in some studies were used as outputs in others. For example, opex and salaries and wages or average labour costs, which are an important component of opex, were used as inputs in several, studies, but wage costs was also used as an output in one study. Transformer capacity was used as an output in one study and as an input in three studies. Transmission line length was used as an input in one study and as outputs in several studies. Combined measures of capacity and distance (in MVA-km) were used as inputs in some studies and outputs in others.

Once again, this shows that there appears to be some confusion about what variables should be regarded as inputs and as outputs. Furthermore, there is a lack of consensus in the studies undertaken to date about the variables that best describe the inputs and outputs of electricity TSOs. It is not possible to discern any trends in the use of variables as inputs or outputs.

Table 3.8: **Output Variables used in Electricity Transmission Studies**

Variable	*No.*	*%*
Energy throughput	5	12
Inward power flows (MWh)	1	2
Inward wheeling power flows (MWh)	1	2
Outward power flows (MWh)	1	2
Outward wheeling power flows (MWh)	1	2
Domestic demand (TWh)	1	2
sub-total	*10*	*24*
Network length (pole miles)	2	5
Underground line circuit miles	2	5
Overhead line circuit miles	1	2
Overhead line length (structural)	1	2
Network circuit length	1	2
Route km	1	2
sub-total	*8*	*19*
Energy x distance (Circuit miles)	1	2
Lines MVA x length	1	2
sub-total	*2*	*5*
Peak or Maximum demand	3	7
Peak demand (ratcheted)	1	2
sub-total	*4*	*10*
Capacity of connected power for renewables	1	2
Connection density	1	2
Energy not supplied (minus)	1	2
Generation (MWh)	1	2
Generation capacity	1	2
Grid size measure -capex	1	2
Grid size measure –opex	1	2
‘Normalized Grid’	2	5
Lines avg availability	1	2
Total transmission plant ‘additions’	1	2
Transformer capacity (MVA)	1	2
Utilization of capacity (%)	1	2
Voltage–weighted entry and exit	1	2
Wage costs ($[(a)])	1	2
Value of weighted an angular towers1	1	2
Densely populated area	1	2
Total energy of the system	1	2
Total	42	100

Note: (a) $ here refers to relevant currency unit.

Table 3.9: **Input Variables used in Electricity Transmission Studies**

Variable	*No.*	*%*
Opex	3	11
Transformer capacity (MVA)	3	11
No. employees	2	7
Total cost	3	11
Total cost (annuity-based)	3	11
Capital price (producer price index)	1	4
Energy losses	1	4
Energy x distance (Circuit miles)	1	4
Labour avg. cost	1	4
Lines cross-section (MCM) x length	1	4
Materials & supplies ($)	1	4
Non-electrical assets depreciation	1	4
Overhead line MVA-km	1	4
Regulatory asset value	1	4
Salaries and wages ($)	1	4
Transmission line length	1	4
Underground line MVA-km	1	4
Unit opex ($/kWh)	1	4
Total	27	100

Table 3.10 lists the small number of input price variables used in the studies of electricity transmission. The variables used include labour average cost, fuel average cost and the user cost of network assets. These studies were in the period up to 2005, as none of the studies after 2005 included any input prices. This observation is consistent with studies of gas TSOs.

Table 3.10: **Input Price Variables used in Electricity Transmission Studies**

Variable	*No.*	*%*
Labour avg. cost	2	50
Fuel avg. cost	1	25
User cost transmission & generation	1	25
Total	4	100

Table 3.11 shows operating environment variables used in electricity TSO studies. Six studies used operating environment variables, four of which were prior to 2005 and two since 2005. Only two of the five DEA studies surveyed used operating environment variables, whereas four out of the five econometric cost function studies used them. The most common types of variables used appear to be regional indicator variables and weather related variables.

Table 3.11: **Operating Environment Variables used in Electricity Transmission Studies**

Variable	*No.*	*%*
Regional indicator variables	2	11
Avg wind speed	1	5
Capex/Opex (interacted with weather)	1	5
Minimum temp	1	5
Rainfall	1	5
Electrical assets exposed to salinity	1	5
Lines exposed to salinity in coastal	1	5
No. substation bays exposed to salinity	1	5
Commercial (% of total)	1	5
Industrial (% of total)	1	5
Distribution density	1	5
Growth of demand - positive	1	5
Growth of demand - negative	1	5
Indicators for years	1	5
Infrastructure Complexity index	1	5
Indicators for holding company	1	5
Public / Private ownership	1	5
Utility/municipal (% of total)	1	5
Total	19	100

3.2.4 Conclusions

Fourteen studies of electricity TSO cost functions or production technology published between 1978 and 2014 have been reviewed. As with the survey of gas TSOs, a wide range of different inputs and outputs have been used in the electricity TSO studies, but the variety of variables has been so diverse it is difficult to discern differences between the types of variables used in earlier and later studies. However, one common pattern is that whereas a small number of studies in the period up to 2005 included input prices, none of the studies in the post-2005 period did so.

Although a wide range of output measures have been used, the most common types of variables are:

- electricity throughput, or separate measures of inflows and outflows.
- transport distance measures (often proxies by the length of the network, for which there are different measures) or capacity × distance measures. Throughput-distance measures don’t appear to be as common for electricity TSOs as for gas TSOs.
- peak day demand measures of maximum delivery capability were used in some studies, whereas physical supply capacity measures were not often used as electricity TSO outputs. Overall, surprisingly few studies of electricity TSOs used some measure

of peak supply capability as an output, given that it is widely viewed as a key driver of costs.

Once again, a number of studies used as outputs variables that were used in other studies as inputs. There appears to be some confusion in the categorisation of variables such as transformer capacity and capacity-distance (MVA-km) type measures. Again, this may reflect difficulty in clearly differentiating between capital input services and capacity-related services such as energy delivery, supply security or ability to meet peak period demand.

Two other general observations are broadly consistent with observations made in regard to gas TSO studies:

- There appears a relatively greater use of DEA compared to other benchmarking methods after 2005, which was not the case in the earlier period.
- Although not as pronounced as for gas TSOs, the majority of the electricity TSO studies did not account for differences in operating environment by using operating environment variables.

3.3 General Conclusions

The conclusions in regard to gas TSO studies in section 3.1.4 and those in regard to studies of electricity TSOs in section 3.2.4 highlight some interesting commonalities. These include:

- lack of consistency in the categorisation of variables as inputs or outputs
- a tendency in most studies to disregard differences in operating environments
- increased use of the DEA method, relative to econometric methods, perhaps associated with the increased importance of regulatory benchmarking studies in the sample after 2005
- at least in regard to gas TSO studies, an increasing tendency to use a single input (deflated total cost).

4 CANDIDATE VARIABLES & MEASUREMENT ISSUES

This chapter draws from the literature review in chapter 3 and gives further consideration to the candidate variables for outputs and inputs of gas and electricity transmission businesses. The alternative methods used in the literature for measuring the variables of interest are also examined.

4.1 Candidate Outputs and Inputs: Gas TSOs

Section 3.1 discussed the alternative methods used in the literature for measuring the outputs and inputs for gas TSOs. This section gives further consideration to the measurement of outputs and inputs for gas TSOs.

4.1.1 Gas Transmission Outputs

The way in which services are charged for is often a useful starting point for investigating the outputs of a business, although as we emphasise in section 4.2 below, there is in general a distinction to be made between functional versus billed output measures. The formats of gas pipeline tariffs vary between pipelines but most commonly depend on factors such as the duration of contracts, the volumes of gas, reservation of firm maximum hourly capacity, load factor, and/or the distance of transportation (ERGEG, 2007). Some of these are related (e.g. capacity, volume and load factor) and some appear to be less important when aggregating over all customers (e.g. contract duration). This suggests that three important candidate outputs, or services of particular interest, are the overall gas throughput, the distances over which the gas is transported and the aggregate capacity used or peak day/peak hour demand.

With regard to gas throughput there is a distinction between the volume of gas measured in terms of cubic metres of gas and the quantity of gas measured in terms of the quantity of energy (e.g. in terajoules). The significance of this distinction is that gas transmission pipelines do not all operate with the same gas specification, and in some jurisdictions such as the Netherlands, distinct pipelines deliver two distinct types of gas in terms of calorific values. It may be argued that the volume of gas in cubic metres is the more useful measure of output because it most closely represents the physical quantity of gas transported, and the calorific value of each cubic metre of gas may be outside the control of the TSO.

Some gas pipeline tariffs may not include an explicit distance component but different rates apply to different geographical areas, or may have different tariffs that apply to different entry and exit points. Some jurisdictions have postage stamp tariffs (e.g. Belgium, Denmark and Hungary). However, there will usually be an implicit distance component in the way the tariff differentials for geographical areas or entry and exit points are formulated (ERGEG, 2007, p. 7). Several benchmarking studies have used the length of a pipeline network as a proxy measure of the distance over which gas is transported, but unless the network comprises a single pipeline extending between one entry point and one exit point, this measure will be crude and inaccurate. More generally, the average distance over which gas is transported will depend on the distances between multiple exit points and multiple entry points weighted by the volumes of gas entering at the various entry points and the volumes of gas delivered at the various exit points.

Peak demand is an important cost driver, but peak demand in a particular year is only useful for benchmarking purposes if there is little volatility from year to year. Since energy demand is generally sensitive to weather, and peak demand especially so, there is usually volatility in the peak demand from year to year. The weather-adjusted measure of maximum demand with a relatively small probability of exceedance (POE) is the most relevant measure for network planning and for benchmarking purposes. For example, for network planning purposes a very small POE, between 2 and 4 per cent, is usually used. If the peak demand data is available for a sufficient number of years, weather correction and probability distributions for peak demand can be calculated, or this information could be directly provided by TSOs. Where a network has multiple pipelines, some average of the peak demands on each part of the network may be appropriate.

4.1.2 Gas Transmission Inputs

If more than one input is included in a benchmarking model, then the starting point for categorising inputs is usually the distinction between capital and non-capital inputs. Capital inputs are assets of the firm that provide services over several periods, whereas non-capital inputs are goods or services fully used within each period.

In gas TSO studies the non-capital inputs are often divided into compressor fuel and all other non-capital inputs. These inputs will not be used in the same proportions because different TSOs will have varying reliance on compressors, both compared to one another and over time. Furthermore, the prices of these two inputs are unlikely to be closely correlated in their movements, which can have implications for cost-efficiency analysis.

Usually, gas turbine compressors are used, and fuelled by a mixture of filtered air and gas taken off the pipeline (Peebles, 1992). Compressor fuel can be measured in energy units, which provides a common basis of measurement, given that different types of compressor fuel may be used.

In the studies reviewed, Other Non-capital Inputs are often measured by the number of employees. However, this approach has limitations because businesses may use different proportions of in-house and contracted services, and because other types of non-capital services may be significant. Another approach often used in benchmarking studies is to deflate operating expenses (after excluding compressor fuel expenses) using a deflator constructed as an average of several relevant input price indexes, including the most relevant index of wages and salaries and input price indexes for services commonly acquired by gas TSOs. The weights for the non-capital inputs price index should be based on the typical expenditure shares for each type of non-capital input.

Gas transmission capital inputs are often separated into pipelines and compressors. This is a useful distinction because these two inputs are, in part, substitutes. Adding compression capacity is an alternative to building additional pipeline capacity. When a pipeline is initially built it will typically have a minimal amount of compression capacity (to deal with pressure drop over long pipelines), and if demand is steadily increasing, this is augmented over time until a maximum feasible amount of compressor capacity is reached. Beyond that, some pipeline duplication or looping would be required. Hence, pipeline and compressor capacity may be used in different proportions over the lifecycle of a transmission system because compressor capacity is less ‘lumpy’ than pipeline capacity.

Pipeline inputs are often measured by length, but this is a crude measure because pipelines also vary in size. Transmission pipeline diameters vary from 40 centimetres (cm) to 100 cm or more, and maximum operating pressures may vary between 4,000 kilopascals (kpa) and 10,000 kpa. Both the diameter and the maximum operating pressure of a pipeline influence the amount of steel used to make it and the services it can provide. The discussion of the study by Granderson (2000) (in section 3.1.1) suggested that a useful measure of the capital input of a pipeline may be given by the formula: $K^{pipe} = A \times p \times d^2 \times l$, where A is a constant, p is pipeline maximum pressure, d is its diameter and l is its length. This type of measure can be aggregated if the TSO has multiple pipelines. This formula may be interpreted in terms of physical input or service potential. In terms of physical input, it approximates the amount of steel in a pipe. In terms of service potential, it also approximates the volume flow rate of gas.

A measure of pipeline inputs may also need to be supplemented by the number of regulator stations or ‘city gates’ (whether by some kind aggregation or as a separate input). This usually depends on the number of discrete urban centres that are supplied by the transmission network, and larger cities will generally have larger city gate facilities.

Although in some studies, compressor inputs are measured by the number of compressor stations or the number of compressor units, they are most commonly measured by capacity such as horsepower. Compressor stations usually consist of several compressors working in parallel, or in a series, and compressor turbines can be of varying capacities (Peebles, 1992). Hence measures based on the number of compressor stations or the number of units, are not ideal. A measure of aggregate capacity (more often nowadays measured in MW) would appear to be more useful.

4.2 Candidate Outputs and Inputs: Electricity TSOs

In Australia, an economic benchmarking framework for electricity transmission businesses using multilateral total factor productivity (MTFP) indexes was developed in 2013 and 2014, to assist in regulatory reporting and assessing forecasts of efficient costs. The output and input specification has recently been reviewed for the first time and a number of refinements recommended. Economic Insights advised (and continues to advise) the Australian Energy Regulator (AER) in relation to the development and review of the electricity TSO benchmarking methodology. This section uses that process as a case study to discuss some of the key issues confronted in relation to variable selection and measurement, drawing on Economic Insights (2013a, 2013b) where the framework was developed, and further developments in Economic Insights (2017). Many of the matters discussed here are also relevant to gas TSO benchmarking.

The AER consulted extensively with industry stakeholders in relation to the identification and selection of relevant variables to be included in the benchmarking analysis. This consultation went hand-in-hand with developing the requirements for information to be provided by industry participants, including precise definitions of all of the variables for which information was to be collected. The availability of robust and consistent data sufficient to support a range of feasible specifications is important for economic benchmarking.

The AER adopted the following criteria for variable selection.

- The outputs should reflect the services *directly* provided to customers, they should be individually significant (in terms of their influence on output or costs), and their inclusion should be consistent with the regulation framework and its objectives.
- The inputs should reflect the production function (i.e. the technology) and be mutually exclusive and as exhaustive as possible. Again, the choice of inputs should be consistent with the regulation framework and its objectives.
- Operating environment factors should be those outside the TSO's control and should have a material impact and be a primary driver of the TSO's costs.

4.2.1 Outputs

Several conceptual issues needed to be addressed in relation to the types of outputs to be considered. These included:

(1) Outputs can be identified with the quantities used for billing customers under the transmission tariffs, or they can be identified as variables that best reflect the services provided to users and that drive the costs of supply (i.e. 'functional' outputs).

This distinction arises because TSO charging practices have typically evolved on an ease of implementation and historical precedent basis rather than necessarily on a network cost reflective basis. For example, some businesses recover a large part of revenue from throughput-related charges, even though costs may not be highly sensitive to throughput; customer-specific charges may bear little relationship to direct customer-related costs; and some services that customers may value highly, such as reliability, may not be explicitly charged for at all. One advantage of defining outputs consistent with billing quantities, in a regulatory setting, is that it ensures the businesses are able to recover an appropriate level of revenue over time, since it reflects how the regulated businesses recover their costs. That said, regulatory revenue or price caps usually do not directly control the tariff structure, so it should not be regarded as a fixed constraint on the regulated businesses. A disadvantage of the billing quantities approach is that TSOs need not have the same tariff structures and in these circumstances it may be difficult to reconcile the billed outputs approach with like-for-like comparability between businesses and consistent data definitions within a benchmarking exercise. These considerations may, on balance, favour the use of 'functional' outputs rather than output measures aligned to billing quantities, but it was thought worthwhile to collect sufficient data to support both approaches where possible, to allow sensitivity analysis to be undertaken. It may also be preferable to include a particularly important billing quantity as an output in the benchmarking analysis, even if it does not have a substantial impact on costs—for example energy throughput.

(2) In principle, outputs should be limited to measures of the services provided *directly* to customers, but in some circumstances it may be appropriate to include 'secondary deliverables' or indirect outputs, such as the capacity required to deliver outputs now and in the future.

The rationale for including secondary deliverables is related to shortcomings in the scope of available measures of direct outputs. For example, the actual peak demand supplied by a TSO is heavily influenced by weather and other factors in a given year, and does not correspond to, and may be a poor guide to, the planning peak day (which has a particular very low probability of exceedance) that TSOs design networks to be able to meet. A more relevant measure of the security of supply service might be the lesser of the planning peak day and the peak day supply capacity of the network, because supply capacity in excess of the planning peak day may be inefficient. The capacity of the network is a secondary deliverable that is relevant to security of supply, but relying only on capacity as a measure of that service has shortcomings because it does not differentiate between capacity installed to meet the planning peak day, and capacity that exceeds it, and businesses may thereby be incentivised to 'gold plate' (i.e. invest inefficiently). For these reasons, both peak day demand and capacity appear relevant to the security of supply service, but care is needed in formulating how they are measured, combined and otherwise used in the analysis. This discussion also suggests that although it is desirable to focus on services directly supplied to customers, if their measurement or availability is inadequate, secondary deliverables may provide additional relevant information.

(3) A number of more specific measurement issues arise in relation to security of supply, peak demand and network capacity:

a. Whether security of supply service is mainly related to peak day supply capability or whether it is broader. If the latter, then a measure of network capacity may be relevant to a wider service than a measure of peak demand (or peak period deliveries).

b. Whether peak demand should be for the demand occurring in a particular year or a 'ratcheted' measure which is equal to the record peak demand up to that year. The peak demand for a given year may be quite volatile due to weather, and this could have an unwarranted impact in efficiency assessments. A ratcheted peak day measure gives the TSO credit for the capacity it has had to install to meet the highest observed peak demand in the sample period without providing an incentive to overinvest and gold plate. However, it may have shortcomings for two reasons. Firstly, if the planning peak day demand is trending upward steadily, the movement of the ratcheted measure may lag behind if the planning peak exceeds actual peak demand. Secondly, the planning peak day may be trending downward, in which case the ratcheted peak demand will remain constant, and again may not provide a good guide to the measure of interest. For these reasons, most available measures of peak demand will be imperfect, although a ratcheted measure is likely to provide a better trade–off than most.

c. How can capacity best be measured when it is the combined result of a number of different physical characteristics of the network? If the appropriate measure of system capacity reflects a number of important network features, how should those different aspects of, or contributors to, capacity best be combined?

d. A more general issue in relation to the services provided by TSOs is whether capacity availability should be viewed from a short-term or long-term perspective. There is a risk that too short-term a perspective may lead to excessive volatility of estimated efficiency scores. Because of the long-term nature of capacity planning, there is a likelihood that the existing physical assets of a TSO may not be optimal in a short-term perspective, either because of historical uncertainties affecting network planning, or because their optimality can only be fully assessed from a long-term perspective (Paulun et al., 2008).

(4) The outputs to be included in a benchmarking analysis should have regard to the importance of different dimensions of service to customers. For example, the reliability of a TSO's network, and hence the reliability of supply to users, will be a key output that should be included in economic benchmarking if reliability is important to customers (or end-customers). If not, it may be a service of only secondary importance, and could be excluded from the analysis without any great detriment. The same principle applies to other aspects of service quality. However, given the critical role transmission plays in the electricity supply chain, transmission reliability is likely to be very important to end-users. However, there may be limited information on the importance of service quality to users because they are not separately priced or unbundled, so that revealed preference information is not available, and because willingness-to-pay studies may not be reliable if wholly dependent on stated preference data. It may also be necessary to cap the weight given to reliability lest the effect of a major outage at one terminal station swamp the TSO's other outputs that year.

(5) There are also issues relating to the choice and inclusion of service quality measures, in part because some of the commonly used measures increase when service quality worsens, i.e., they are measured as 'bads'. For example, the frequency and duration of outages are measured by indexes that decrease when there is an improvement in service reliability. Some of the alternative ways of dealing with this are:

a. A measure of this kind may be transformed so that an increase in the measure corresponds to an increase in quality. In some cases, a meaningful transformation may be available, but this approach may not be effective where it leads to nonlinearity, or where the transformation relies on *ad hoc* assumptions, or distorts the measure in some important way. For example, one can define a maximum level of outages and subtract the actual level of outages from this, to produce a variable where a higher value represents higher quality. However, there is the degree of arbitrariness in setting the maximum level, which would need to be common for all the TSOs being benchmarked and sufficiently high that it exceeded the worst observed performance to ensure the result of the subtraction was positive in all cases.

b. Some economic benchmarking studies have included reliability as an input rather an output in recognition of a DNSP's ability to substitute between using opex and capital, on the one hand, and reduced reliability and associated penalties on the other. However, it is perhaps more common to include a measure such as total outage minutes as an undesirable or 'bad' output. This involves constraining the weight for this variable to be non-positive rather than non-negative within the DEA program, so that a decreasing value of the 'bad' will be consistent with increasing overall output.

c. Some studies combine an output quality measure with a quantity measure for the same output to obtain a quality-adjusted output measure. Consequently, when output quality deteriorates it has a similar effect to a reduction in the output quantity. This approach has the advantage of greater parsimony than separately including the output quantity and quality measures, but implicitly an assumption is made about the rate of conversion between units of quality and quantity of the output.

4.2.2 Non-Capital Inputs and Input Prices

The main two groups of inputs are durable and non-durable inputs. Issues relating to measuring the quantities and prices of durable, or capital, inputs are addressed in our report on Capital Costs, and hence not addressed here. Non-durable, or non-capital, inputs are the inputs associated with operating and maintaining the network, including inspection, maintenance, repair, vegetation and emergency response, and in the case of gas pipelines the fuel used to operate compressors. The total cost of these inputs is operating and maintenance costs (i.e., not including depreciation or capital expenditure).

Some of the issues considered in relation to non-capital input quantities and prices (or together, the input cost) in the context of developing the electricity TSO benchmarking framework in Australia, are discussed in this section.

(1) Decisions need to be made in relation to the level of aggregation at which inputs are to be measured. In many benchmarking studies, just two inputs are included, capital and non-capital inputs. In some cases either capital or non-capital inputs (or both) may be disaggregated into their main components. For example, non-capital inputs may be separated into labour, materials and energy, or some other grouping. On the other hand, some benchmarking studies use the number of employees as a measure of non-capital inputs as a whole, which assumes that other components of non-capital inputs can be ignored. This would not be a realistic assumption if there was high variability in the degree of contracting out of operating and maintenance activities by TSOs in the sample. When the data sample size is small in DEA analysis it will be important to aggregate the non-capital inputs. If data is available at a disaggregated level for either input quantities or input prices it is usually feasible to derive an aggregate quantity or average input price measure using an index number method. In general, this should be a more suitable measure of the quantity of a group of inputs than simply choosing a measure representing one of its sub-groups (e.g. the number of employees).

(2) The quantity of a particular input may not be directly observable from available data, but may be estimated by dividing the cost of that input by the price, or an appropriate price index, for that input. For an aggregated group of inputs, such as non-capital inputs as a whole, a quantity index for that group can be obtained in the same way, by constructing an appropriate price index for that group of inputs, and dividing the total cost of that group of inputs by that price index. Because of the diverse composition of operating and maintenance costs, and differences between businesses in relation to contracting out or in-house provision of services, direct measurement of non-capital input quantities is often difficult and the method of deflating relevant costs is often used.

A price index for a group of inputs should be a weighted average of the prices of the key components of that group of inputs. Ideally, this weighted price index will reflect as closely as possible the prices faced by each TSO. However, in practice this information is usually not available, and economic benchmarking studies often use high-level input price indexes compiled by national statistical agencies. While objective and less subject to gaming than price information collected from TSOs, these indexes may not always accurately reflect the input prices faced by individual TSOs. In Australian electricity network benchmarking, a combined index for non-capital inputs has been constructed from a labour price index and five producer prices indexes (to represent different components of non-labour inputs), with fixed weights designed to represent the cost shares of the components of non-capital inputs.

(3) Consistency of the operating and maintenance cost ('opex') data collected from TSOs is necessary. Some areas where particular attention is needed to ensure consistency are:

a. capitalisation practices, for example in relation to isolated asset refurbishment;

b. cost-allocation methods, such as corporate overhead allocation in businesses that have other activities in addition to electricity (or gas) transmission;

c. related party services, such as a network operating agreement with a related company, can cause comparability difficulties if the transfer price is not cost reflective;

d. energy losses in transmission may be treated differently between jurisdictions. For example, in some cases generators may bear these costs and in others they may be borne by the TSO. Care is needed to ensure consistency.

4.3 Quality Variables

It can be important to take into account differences in the quality of service because the introduction of quality variables can significantly affect performance comparisons between utilities and over time. Measures of output quality have not been widely used in cost and productivity studies of electricity and gas TSOs, but are more used more widely in other utility sectors:

> Reliability of supply is clearly the big quality issue in electricity distribution. Typical quality measures in industrial countries include average length of supply interruption

> per customer per year, delays, cancellations, and average time taken to restore supply when it has been interrupted. … In addition, a quality variable that needs to be taken into account is the technical quality of the service. The distribution of water and electricity will invariably result in losses. The design and maintenance of the network will determine how much is lost. Some studies include these losses as a variable … (Coelli et al., 2003, pp. 90–91)

As indicated in section 4.1, benchmarking analysis should take into account the quality of service to customers, wherever the dimension of service quality is important to customers. This is likely to include the reliability of a TSO's network, which directly influences the reliability of supply to users. However, the interest for benchmarking is primarily in endogenously determined output quality, which may conflict with the chosen orientation in a DEA analysis, depending on the method chosen. Alternatively, exogenous factors that influence service quality can be included in a second-stage analysis to control for their effects. For example, severe weather events can have a substantial effect on outages, and this is clearly an exogenous factor, different in nature from service performance outcomes due to inadequate maintenance or response to outages.

Coelli et al suggest that the benefits of including quality of service variables needs to be assessed given the data sample at hand, and their likely importance:

> If degrees of freedom allow, and if a consistent measure of quality is available, regulators can include a quality variable; however, if it expects the effect to be limited to a few firms, or if degrees of freedom are tight, the best approach might be to omit the quality variable, and then ask the firms to discuss their individual situation with the regulator if they believe they have a case. (Coelli et al., 2003, pp. 90–91)

4.4 Operating Environment Differences

Utilities tend to operate in discrete geographical areas, and features of the geographical location, including topography, characteristics of the urban areas supplied (e.g., density) and climate in those locations, may all have an important influence on observed productivity, costs and profits. These operating environment characteristics essentially act as constraints, and can influence the ability of businesses to convert inputs into outputs. This section discusses a number matters relating to operating environment variables including:

- identification of candidate exogenous variables that reflect differing operating environments of energy networks, and the types of variables that have been used in previous studies, their rationales and definitions.
- available methods for determining which environmental variables are likely to be most important and controlling for their effects.

4.4.1 Operating Environment Variables

The aim of making like-for-like comparisons in benchmarking studies supports taking operating environment factors into account. Since they may impose constraints on the ability to achieve cost efficiency improvements, regulatory targets may be inappropriate if they are not taken into consideration. However, there is an issue of regulatory judgement around which types of factors to allow for, as emphasised by Hawdon:

> The environment in which the gas industry functions varies considerably from country to country, in terms of the terrain over which gas is transported, the geographic density of customers, and their economic characteristics. While this is easily recognized, treatment of such individual circumstances can be affected by strategic considerations. Producers have an interest in stressing the uniqueness of the conditions of supply since regulatory concessions often flow from such recognition. Any such concessions may however be welfare reducing as they remove pressure on producers to improve efficiency in the absence of properly functioning competitive markets. This creates a presumption against including measures of uniqueness where it is desired to assess relative performance unless a priori considerations are overwhelming. (Hawdon, 2003)

Since relatively few operating environment factors can be included in a benchmarking analysis when the sample size is small, it is best to concentrate on those that have the most significant effect and which vary the most across TSOs. The reason for including only those factors genuinely exogenous to the TSO (ie beyond management control[8]) as operating environment factors is that otherwise the incentives provided to the TSO to minimise costs and operate efficiently may be reduced. Where a number of operating environment factors are highly correlated, only the one with the most direct impact on TSOs' costs may be included.

In our previous work we identified the following main types of operating environment factors that can have an important bearing on energy network costs.

- *Climate:* Weather conditions such as storms, high wind and extreme heat can have a material impact on network operations, particularly for above-ground electricity transmission infrastructure. The weather conditions in particular regions will have an important influence on electricity transmission network costs because they influence engineering design requirements and maintenance activities. The incidence of severe storm events can also materially affect a TSO's costs from year–to–year and may make a TSO look inefficient in those years where it has had to restore services and clean up after severe weather events. Weather is clearly beyond management control,[9] and some benchmarking studies have found weather to be a decisive factor in explaining observed efficiency differences between energy networks (Yu et al., 2009).
- *Terrain:* The terrain and other physical features of a TSO's areas of operation can have a material and important impact on its costs. For example, mountainous areas will usually be more costly to traverse with infrastructure than flat areas, forested areas may incur higher vegetation management costs, and infrastructure may be more costly to establish through built-up urban areas than over farm land. In the case of gas

[8] By 'beyond management control' we mean that the operating environment factors are exogenous for the firm. Management can still make choices in how to deal with operating environment factors (which may be more or less effective), but these responses generally require resources to implement, so that differences in operating environments can affect the observed comparative productivity and cost efficiency of firms even when action is taken to mitigate their effects. The effects of operating environment factors are an empirical question.

[9] On the meaning of 'beyond management control', see foregoing footnote.

TSOs, ground conditions (e.g., sub-soil, geology and hydrology) will be important factors in pipeline construction and corrosion mitigation costs, since, for example, trenching in rocky ground will be more costly than in deep soil or clay. Topographic features such as river crossings can also be important. These characteristics are clearly beyond management control and are likely to be a primary driver of costs.[10] However, it is often difficult to obtain accurate indicators for the average terrain conditions applicable to each TSO and each TSO will typically have to deal with a range of terrain types.

- *Network configuration:* Characteristics, such as the concentration or dispersion of demand centres and distances between energy sources and demand centres will influence the design of networks including whether their configuration is linear or meshed, etc. These characteristics are sometimes referred to as 'network topography'. It is well established that these characteristics can have an important influence on network performance and failure risks, and are also likely to be an important driver of costs. To a large extent the topologies of electricity and gas transmission networks have evolved from land constraints and from locational and other decisions of upstream or downstream firms and of customers, and hence largely beyond TSO management control. However, there may be a range of actions that TSOs can take to modify the network topologies of existing assets to enhance their performance. Some measure of weighted average distances between transmission system entry and exit points is often used in benchmarking studies, but it is more challenging to obtain a single indicator that captures a wider set of network topography characteristics that have greatest influence on costs.
- *Regulations and Standards:* TSOs can face constraints on their operation from jurisdictional standards, regulations and environmental considerations. Standards can apply to design and construction. For example, TSOs may be obliged to adopt higher cost routes or undergrounding for new transmission lines in response to environmental issues. Standards may not only apply to the TSO, but directly to its employees and contractors. TSOs may face different technical standards or different environment protection requirements across different jurisdictions. These constraints are usually exogenous to TSOs, deriving from legislation, regulations, licenses, or standards bodies, and may have a material impact on costs but are difficult to quantify robustly and objectively.
- *Peak demand:* If peak demand is not included as an output, it (or the related load factor measure) may potentially be an operating environment factor. It is unarguably a significant and often a primary driver of TSO costs. Since transmission services providers are only one part of the supply chain in delivering energy to end-consumers, a TSO's ability to substantially change peak demand may be limited. However, it is not entirely outside the control of TSOs because tariff structures can be adopted which influence peak period demand. Including system peak demand as an operating environment factor in economic benchmarking may not incentivise TSOs to take

[10] On the meaning of 'beyond management control', see foregoing footnote.

actions which are under their control to smooth peaks and reduce the need for costly additional underutilised infrastructure.

4.4.2 Choosing and Controlling for Operating Environment Variables

Various methodologies can be used to control for the influences of non-discretionary or operating environment factors. At a broad level this may be done:

- before the DEA analysis, such as pre-analysis adjustment of data;
- during the DEA analysis, either by including operating environment variables in the DEA analysis alongside inputs and outputs or by using subsamples of like TSOs in the analysis (the latter approach is discussed further in section 4.4.4); or
- after the DEA analysis, such as by using 'second stage' approach to analyse and control for the influence of business environment factors on measured efficiency.

The approach to controlling for operating environment characteristics by treating them as additional inputs or outputs in the DEA analysis can be contentious because efficiency measurement in DEA assumes that the inputs produce the outputs, and there is no reason to expect that assumptions derived from production theory, such as monotonicity, convexity, etc., would apply to those variables. Furthermore, since operating environment factors are generally exogenous, they cannot usually be proportionately scaled down in input-oriented DEA (or scaled up in output-oriented DEA) through management discretion, as is typically assumed for regular inputs and outputs.

A more common approach is firstly to carry out the DEA analysis without controlling for the exogenous factors, and then conduct a 'second stage' analysis, in which the estimated efficiencies scores are used as the dependent variable in a regression against the operating environment factors. The model obtained from the second stage regression can be used calculate 'normalised' efficiency scores which control for differences in the exogenous factors. Early studies used a censored regression model (e.g. Tobit model), in which values of the dependent variable are not observed beyond a censoring threshold (here the maximum efficiency of 1 or 100%. This practice was criticised by Simar and Wilson (2007), who pointed out that (under certain assumptions) it may be more appropriate to use truncated regression, using bootstrapping. This can be combined with bias-correction (also using bootstrapping). Bias correction is used to derive better estimates of the (unconditional) efficiency scores relative to the best practice, whereas the second-stage truncated regression analysis is used to control for the effects of environmental factors to obtain (conditional) efficiency scores, given certain values of environmental factors.

It may be feasible (and probably desirable) to carry more than one approach to controlling for operating environment factors, and compare the results. It is also possible to combine some of these approaches, for example when some exogenous factors may be readily controlled-for through normalisation of variables before the DEA analysis, others can be addressed through second-stage analysis. One example of controlling for operating environment factors before undertaking the analysis is the process undertaken by Agrell and Bogetoft (2009), in which electricity TSOs were given the opportunity to submit operator specific claims—i.e. adjustments they believe should be made to the data to put them on a like-for-like basis—and these proposals were assessed on a case-by-case basis. Another example, is weather correction of energy data, which takes account of differences in weather patterns on gas or electricity deliveries.

4.4.3 Examples from TSO Studies

This section lists the operating environment variables that were used in the studies citied in chapter 3, and includes a few of the inputs and outputs that appear more like operating environment variables, because they are outside the influence of (i.e. exogenous to) the TSO. The examples are as follows:

- Growth in demand
- Reform (privatisation or deregulation)
- Responsiveness to EU gas directive
- Share of gas (or electricity) in total energy
- Capacity of connected power for renewable energy
- Average wind speed
- Minimum temperature
- Rainfall
- Capex/Opex (interacted with weather)
- Commercial (% of total)
- Distribution density
- Assets exposed to salinity
- Indicators for parent company
- Public / Private ownership
- Indicators for years
- Industrial energy use (% of total)
- Utility/municipal (% of total)
- Infrastructure Complexity index
- Regional indicator variables
- Connection density
- Supply area
- Land use
- Subsurface features
- Topography
- Soil humidity.

This is not an exhaustive list because, as noted in chapter 3, many studies did not control for operating environment factors.

4.4.4 Different network configurations

As discussed in section 4.4.1, different types of network configuration can arise due to historical or operating environment factors and can influence the efficiency achievable by businesses. That said, TSOs may be able to modify their network configurations over time to achieve greater technical or allocative efficiency. Questions that arise include: (a) whether the differences in network configuration are adequately reflected by suitable measures of network capacity and length, or not; and (b) whether these differences are correlated with other operating environment factors such as supply area or urban density measures. These are empirical questions.

In the context of small data samples, it is unlikely to be feasible to divide the sample into smaller sub-groups representing TSOs with similar network configurations. This is likely to increase the number of TSOs that are found to be efficient. However, in the non-parametric DEA framework, and in the context of small samples, the validity of decomposing the sample in this way could not be tested statistically. A more useful approach would be to use second-stage regression, as discussed in section 4.4.2 (and discussed in more detail in our report 'Choosing the model and explaining the results'), with a categorical variable for network configuration as one of the explanatory variables.[11] This can be used to:

- test whether network configuration has a significant effect on measured efficiency scores
- test whether the effect of one network configuration is significantly different from that of another, and
- quantify the effects of network configurations on efficiency scores (if their effect is statistically significant).

If the other candidate operating environment variables are included in the second-stage regression (as they should be), then questions (a) and (b) above can be addressed. If network configuration is already adequately reflected in other variables included in the DEA analysis, there will not be a significant effect on efficiency scores in the second-stage. If other operating environment variables measure a similar thing, this can be identified using tests for multicollinearity, and hypothesis tests can be used to decide which variables to retain in the model.

[11] Categorical variables are included in a regression analysis using dummy variables for observations in each category except the base category.

5 TECHNIQUES FOR VARIABLE SELECTION OR REDUCTION

In any application of benchmarking it is necessary to select among alternative variables to find those that best measure the inputs, outputs, input prices and operating environment variables for the industry being examined. The need to narrow down the number of variables is especially acute in the context of DEA analysis (and other nonparametric approaches). An important limitation of DEA is that, as more input or output variables are added to the model, and so its dimensionality increases, it loses some ability or power to discriminate between efficient and inefficient DMUs (and the confidence intervals around efficiency estimates also widen). This is an especially important problem when data samples are small. It is also often the case that there are only a small number of really important variables, and the challenge is to identify them.

The focus of this section of the paper is on techniques for variable selection and search for parsimony. Section 6.1 considers a number of methods and principles for screening and selecting the most suitable variables to use in a benchmarking study. Section 6.2 considers an approach that involves using principal components analysis (PCA) to transform the set of original variables into a smaller group of derived variables that contain much of the information in the original variables, thereby reducing dimensionality with minimal loss of information, and hence minimal bias to the efficiency estimates.

5.1 Techniques for Variable Selection

In addition to well-established approaches that rely on economic theory and industry-specific engineering knowledge, there are a number of more recently developed approaches that make use of tools of inference and sensitivity analysis to screen variables and improve DEA models. Techniques of this kind are briefly surveyed, their strengths and weaknesses and the extent to which they have been used in the literature are considered.

The approaches discussed below include:

- ‘First principles’ approach
- Reliability Assessments
- Partial Correlations and Preliminary Regressions
- Efficiency Contribution Measure
- A Regression-based Test; and
- Bootstrapping.

5.1.1 *The ‘first principles’ approach*

In any parametric or nonparametric modelling task, knowledge of the industry being examined is important to ensure that the variables used and the specifications employed are likely to be sufficiently representative, at least as a starting point for analysis. This is one reason why benchmarking exercises usually involve consultation with industry participants to assist in ensuring that the most appropriate variables and definitions are considered. Economic theory also has an important role to play in understanding the way that the industry works, how the variables are likely to be interrelated, and how operating environment factors are likely to influence the production process. It can be useful to consider a conceptual or theoretical model that is believed to describe the functions of the industry, and derive from it the relationships to be included or tested in the model. The techniques that are discussed in the remainder of this chapter are not a substitute for the appropriate use of industry knowledge and economic theory, but can be helpful as complementary considerations in practical applications of DEA.

5.1.2 Reliability Assessments

One approach to checking the reliability of a proposed choice of variables in a DEA context is to undertake trial analysis and analyse the implicit weights of the DEA model, which reflect observed marginal rates of transformation between inputs and outputs, or if there are multiple inputs, the marginal rates of substitution between them (Thanassoulis et al., 2008). These results may be compared to estimates of marginal rates of transformation or substitution obtained from the literature on energy network cost analysis and benchmarking. This may help to assess the plausibility of candidate input or output variables.

If data is available over several years, then it may be useful to examine the extent to which variables perform well when the model is applied to different periods, allowing an assessment of the stability of the implied marginal rates of transformation or substitution. The underlying assumption is that the production process is stable so that the variables measuring the inputs and outputs should perform similarly in each period. A similar approach may be taken using sub-samples of TSOs.

5.1.3 Correlations and Scatter Plots

Correlation coefficients and scatter plots between pairs of variables can be used as a guide to the degree of association between them, and studies such as Jamasb *et al* (2007) have made use of such information to guide variable selection. For example, if an input has little correlation with any output it might be omitted, and where two input variables are highly correlated with each other (or two output variables are highly correlated with each other) then one of those variables is often omitted or they can be aggregated into one.[12]

A limitation of such approaches, as Jenkins and Anderson noted, is that "the interrelation between a number of variables that are partly correlated is rarely obvious, and which ones can be eliminated with least loss of information cannot be determined just by looking at the correlation matrix" (Jenkins and Anderson, 2003, p. 54). They developed a more systematic statistical method of variable selection using partial correlations. The method involves testing whether a subset of the original variables account for the great majority of the total variation in the original data set by testing all possible partitions of the variables into two groups (those to be retained and those to be omitted) to determine which partition works best in representing as much as possible of the information in the original variables, while achieving greater parsimony. The variables that will be dropped will be those with partial correlations

[12] There is no hard rule, but a typical rule of thumb is that correlation of above 0.8 is considered as highly correlated.

(conditional on the included variables) that are close to zero.

A remaining concern is that the DEA efficiency scores can still be sensitive to the omission of variables with very small partial correlation. It should also be noted that Adler and Yazhemsky (2010) find that PCA-DEA (discussed in section 5.2) performs better than Jenkins and Anderson's variable reduction method, especially with small datasets.

5.1.4 Efficiency Contribution Measure

Pastor *et al* (2002) developed another statistical approach to variable selection in DEA. It involves comparing pairs of nested radial DEA models, one with a candidate variable (input or output) and one without. The incremental effect of the candidate variable on the efficiency measures is quantified using the efficiency contribution measure (ECM). If the estimated efficiencies are not *expected* to be substantially affected by including a candidate variable in the model, then that variable is considered not relevant. A variable will be relevant if its inclusion affects the estimated efficiencies of at least a certain proportion (say 15%) of DMUs by at least a certain amount (say 10%). Using bootstrapping a statistical test of the null hypothesis that the candidate variable is irrelevant can be carried out.

Formal iterative variable selection procedures, such as forward selection (progressive inclusion of variables until the significance test is not met) or backward elimination (progressive elimination of variables until all the retained variables meet the significance test), or both, can be used. But the authors indicate that: "in general, it is not recommended that these kind of automated procedures be used blindly to identify a 'best' model because they can never replace professional judgement in the matter field" (Pastor et al., 2002, p. 732). Nataraja and Johnson (2011) test several alternative methods and conclude that the ECM method worked well in most of the scenarios tested, and is best suited to larger samples (>100 observations) where there are low correlations between the variables.

5.1.5 Regression Methods

Another type of approach used to narrow down a long list of candidate variables is to use preliminary statistical regression analysis, an approach used in some of the previous benchmarking studies of gas and electricity TSOs undertaken for European regulators. For example, OLS regression can be used to estimate a production, cost or distance function to identify whether there is a statistically significant relationship between a dependent variable, such as cost, and the explanatory variables, such as the candidate cost drivers. This approach has been frequently used in energy TSO benchmarking, for example: Jamasb et al (2007), Jamasb et al (2008), and Frontier, Consentec & Sumicsid (2013), among others. Alternatively, SFA models could be estimated to determine the statistical significance of alternative variables.

One concern about this approach is the need to assume a specific functional form for the production or cost function. Indeed, very simple functional forms are typically assumed, such as linear, log-linear or log-log, without quadratic and interaction terms necessary to be classed as flexible functional forms. This means there is a risk of bias in tests of the statistical significance of variables due to the risk of misspecification error. This raises a question as to whether some of the benefits of using DEA, in regard to not imposing a particular functional form on the technology, may be mitigated when the selection of variables has been made using an assumed functional form.

An alternative regression-based approach was developed by Ruggiero (2005). In this approach, a minimal DEA model is initially computed using only the output(s) and one input. The resulting technical efficiency scores are then used as dependent variables in a linear regression against all other candidate input variables that could have been included in the production or cost function. For each candidate input x_i, there is an estimated coefficient β_i, and x_i is only to be added to the DEA model if β_i is statistically significantly different from zero (e.g., at a suggested 90% confidence level). After the significant variables are added to the model, the procedure is repeated with the new technical efficiency scores regressed against the remaining candidate variables that were not previously included. This procedure is repeated until the regression yields no additional significant explanatory variables.

Nataraja and Johnson (2011) found that this method worked well with relatively large samples, and when there was very little correlation among the candidate inputs. It is also relatively easy to implement. That said, in network benchmarking applications, input and output variables are often highly correlated with one another, in part reflecting substantial differences in relative sizes of DMUs in the samples.

5.1.6 Bootstrapping

An alternative method is to carry out preliminary DEA analysis with and without a candidate variable, and use bootstrapping to develop statistical test procedures for the significance of the individual candidate variables. Simar and Wilson (2001) developed a bootstrap method of this kind, and proposed various test statistics for determining:

- whether some outputs or inputs in the model are irrelevant; and
- whether some of the inputs or some of the outputs can be aggregated, thereby reducing the number of independent inputs and outputs.

The bootstrapping method is used to obtain critical values of the test statistics for the significance of overall change in the estimated distance functions, and forward selection or backward elimination may be used. Nataraja and Johnson (2011) found that this method did not generally work as well as either the ECM or Ruggerio regression-based methods, in the tests they carried out.[13] That study also examined the PCA-DEA method, and the findings are discussed in section 5.2.

5.2 Methods for Improving Parsimony

The previous sections have examined variable selection from the perspective of identifying the most appropriate measures of the outputs, inputs and input prices for the industry being benchmarked (assuming operating environment variables are only included in the second-stage analysis). The desirability of selecting only those variables that have an important

[13] The reason for this might be due to the restricted version of the bootstrap used in Simar and Wilson (2001), relative to more recent developments (e.g. Kneip et al., 2016, 2015, 2008).

bearing on cost or input use has been balanced against the need to ensure completeness in the representation of the technology of the industry under study. This section considers methods aimed at achieving a greater degree of parsimony in the variables used in a DEA study for reasons of improving the robustness of the efficiency measures derived.

The need for parsimony in the variables used in a DEA analysis arises especially in regulatory benchmarking contexts, in which data sample sizes are often small. This introduces two problems. Firstly, resulting efficiency scores may not be meaningful when the number of variables used is large relative to the number of observations (Simar and Wilson 2008). This is because the efficient frontier is likely to be further away from the true efficient frontier with a small sample than with a large sample (i.e. bias is greater in small samples as there is convergence toward the 'true' frontier as the sample increases) and this problem is accentuated when there are more dimensions to the frontier when a larger number of variables are used (i.e. rate of convergence to the 'true' frontier is much weaker) (Simar and Wilson, 2004). This means that as the number of variables increases relative to sample size, the degree of bias is likely to be increased, impairing the accuracy and reliability of the estimated efficiency scores. This is referred to as the *curse of dimensionality problem*. Second, for technical reasons in DEA analysis using small data samples, the more variables that are included, the more DMUs are found to be efficient. This is because the frontier is defined in more dimensions and the heterogeneous characteristics of DMUs means that more are used to define the frontier. When data samples are quite small, even a few included variables may be sufficient to result in a significant proportion of the DMUs being efficient. This is referred to as the *low discriminative power problem*.

This raises a conundrum for regulatory applications. On one hand, it is desirable to have a sophisticated model that includes all relevant variables to better identify the 'true' levels of efficiency. As Adler and Yazhemsky (2010, p. 282) have observed "the omission of relevant variables leads to under-estimation of the mean efficiency, while the inclusion of irrelevant variables leads to over-estimation." A model that excessively restricts the number of variables may result in efficiency measures that are downwardly biased (in the input orientation). On the other hand, as Nieswand *et al* have observed, "the request for a realistic representation of company structures easily increases the number of variables substantially and hence, harms the ratio between observations and variables. The known consequence is a deteriorated discrimination capability of DEA" (Nieswand et al., 2010, p. 14). The efficiency scores may become unreliable.

This section examines methods of achieving greater parsimony, including through aggregation methods, such as the construction of indexes or variables that otherwise combine variables; and through the use of Principal Components Analysis (PCA) in conjunction with DEA.

5.2.1 Aggregation

It may be feasible to aggregate some of the inputs or outputs to reduce the dimensionality of the DEA program. Some of the available methods of aggregation include:

- Summation, when variables to be aggregated are in monetary units.

- Constructing indexes that combine other variables. An example from the TSO benchmarking literature is the development of aggregate measures of the physical capital stock from a set of variables that measure the quantities of different types of assets, and weights indicating the relative importance of the different asset types, such as cost shares.[14] In this way, the input measures used in the DEA analysis may essentially be indexes of groups of inputs. This method is used in all index-based methods of measuring aggregated inputs, outputs and productivity, but can be applied to construct sub-aggregates, such as an aggregate for all capital inputs. The Törnqvist index method is often used in the context. The Törnqvist index rate of change is defined as: $\delta_{t,t-1} = \sum_{i=1}^{n} \frac{1}{2}\left(w_{i,t} + w_{i,t-1}\right). \ln\left(x_{i,t}/x_{i,t-1}\right)$, where w's are value-share weights, x's are quantities, $i = 1 \ldots n$ is the number of measures being aggregated and t and t-1 are two consecutive periods. For example, if $x_{i,t}$ is the quantity of an input, then $w_{i,t}$ will be the cost share of that input (out of all the inputs included in the aggregate, such that $\sum_{i=1}^{n} w_{i,t} = 1$). If $x_{i,t}$ is an output quantity, then $w_{i,t}$ may be a revenue share or it may be constructed using marginal costs as unit values (Denny et al., 1981). The Törnqvist index, in index number form, in period t is: $\exp\left(\prod_{\tau=0}^{t} \delta_{\tau,\tau-1}\right)$, assuming the base period value of the index equals one.
- Computing a new variable by combining other variables in a meaningful way. For example, some TSO benchmarking studies have made use of engineering formulas to construct variables that may better represent the capacity or output of a facility. Alternatively formulas may be derived from commercial practice. For example, gas transmission pipeline services are often charged-for based on an energy × distance measure (particularly in the USA), so this type of combined variable may sometimes be more meaningful than using energy throughput and pipeline length as separate outputs. Another kind of formula may come from other research. For example, willingness-to-pay studies that provide guidance on customer valuation of quality attributes of a service may be used, together with output quantity and quality variables, to construct a quality-adjusted output measure.

  The use of quality-adjusted output measures is common in water utility benchmarking because substantial quality improvements have been made over time at considerable cost (see: Hunt and Lynk, 1995; Saal and Parker, 2000). This involves constructing a service quality index for each output, relative to a base period, and multiplying the measured quantity of the output by the quality index. For example, a quality index for

[14] An aggregate measure of capital assets, as described here, is an index of quantities of different types of capital assets using value-based weights such as the proportion of total asset value accounted for by each type of asset. This sums to an index-based measure of the physical capital stock. This differs from the 'Normalized Grid' variable used in some benchmarking studies reviewed here, which is an aggregate of two components. In one component, the quantities of assets of each type are multiplied by the relative operating and maintenance cost associated with an asset of that type. The other component is an annuity value associated with a series of investment-value related aggregates, in which the quantities of each asset type installed in each year are multiplied by the relative cost of installing one unit of an asset of that type. The 'Normalized Grid' variable appears to be related to efficient costs but does not appear to be a measure of the physical capital stock.

sewage treatment may be based on the average level of treatment (primary, secondary or tertiary). Where water utilities have environmental responsibilities, the quality index may be a measure of the environmental health of the waters under management. Quality-adjustment of output measures is analogous to quality-adjustment of pricing (e.g. using hedonic pricing methods) and is used in preference to including separate quality variables in order to preserve degrees of freedom.

5.2.2 Principal Components Analysis DEA (PCA-DEA)

Principal components analysis (PCA) is a data reduction method. It is used to evaluate whether the values of a set of variables can be explained by a smaller number of *latent variables* called 'components'. Data reduction is feasible when the original variables are correlated with each other, so that some of the information they contain is the same. It is assumed that the variables in the dataset (i.e. the set of inputs, or alternatively the set of outputs) can be interpreted as linear combinations of one or more underlying components, which are uncorrelated with each other. The aim of PCA is to reduce the number of variables by deriving those components. Although the number of components is the same as the original set of variables, and contain the same information, the way that the information is separated into the components differs, with the earlier components containing more information than the later components. The idea is that the leading components (say, the first two) may do a reasonably good job of reconstructing the correlations among a larger number of measured variables (Warner, 2013).[15]

PCA-DEA involves using the leading components (or principal components) as the variables in the DEA analysis rather than the original variables. Since a smaller number of components (e.g. two) can contain much of the information in a larger set of original set of variables, this means more variables to be taken into account in the analysis, without excessively multiplying the number of variables used in the DEA analysis, and hence without such a detrimental effect on the quality and reliability of the DEA efficiency estimates. Daraio and Simar (2007) was one of the first DEA studies to use this method. Nieswand *et al* (2010) applied the PCA-DEA method to benchmarking natural gas transmission businesses.

Given a set of variables, none of which is a dependent variable, and where there are no assumed groupings of observations or any partitioning of variables into subsets, PCA involves finding particular linear transformations of those variables which are uncorrelated with each other (or dimensions that are orthogonal to each other), and contain most of the variance in the dataset. The original data is first centred and standardised, which does not affect DEA scores.

The principal components (PCs) are new variables that are weighted sums of the original variables, where the weights are obtained from the elements of the eigenvectors of the dataset. The PCs are uncorrelated with each other (by construction) and ranked by their variances, or contributions to the overall variation in the original data. The first PC is the linear combination of the variables in the dataset having maximum variance. The second PC is the linear combination having maximum variance of those completely uncorrelated with

[15] PCA can only be used on interval data. It cannot be used on categorical data.

(i.e. orthogonal to) the first PC. The third PC is uncorrelated with the first and second PCs, and so on.

> In some applications, the principal components are an end in themselves and maybe amenable to interpretation. More often they are obtained for use as input to another analysis. For example, two situations in regression where principal components may be useful are (1) if the number of independent variables is large relative to the number of observations … and (2) if the independent variables are highly correlated … In such cases, the independent variables can be reduced to a smaller number of principal components … (Rencher and Christensen, 2012, p. 406)

Although the complete set of PCs is as large as the original set of variables (and if all used in the DEA analysis the result would be identical to the DEA model using the original variables), the idea is to remove the PCs that account for the least amount of variation one-by-one until the remaining PCs explain at least 80 per cent of the variance in the original data (as a rule of thumb). A Scree graph for eigenvalues is also often used to look for a natural break between the large eigenvalues and the small eigenvalues (Rencher and Christensen, 2012).

The application of PCA in conjunction with DEA has the particular advantages of:

- allowing a richer set of input and output variables to be used in the overall analysis (thereby improving the ability to identify 'true' efficiency); while also
- enabling a reduced number of variables used in the DEA analysis (thereby mitigating the dimensionality and discrimination problems).

In their testing of several variable reduction methods Nataraja and Johnson concluded that:

- PCA-DEA is a robust technique in which some amount of information is retained from each of the original variables, unlike the other three methods which select or discard one entire variable. PCA-DEA also has the smallest run time, works best with smaller sample sizes, and is robust to the high correlations between inputs and irrelevant variables. (Nataraja and Johnson, 2011, p. 668)
- As sample sizes increase (e.g. over 300), and correlation between inputs diminishes, PCA-DEA was found to underperform against the ECM method, discussed in section 5.1.4, and the regression-based method of Ruggiero discussed in 5.1.5. This is because the PCA-DEA method is ultimately limited by the fact that only 80 per cent of the variance in the original data is retained.

5.3 The Regulatory Context

Objectives within the regulatory framework will also be relevant considerations in the selection of the set of candidate variables. This is because the candidate variables used for a benchmarking study will influence the dimensions in which services and inputs are viewed, which will in turn affect the way in which efficiency targets are formulated for regulated businesses. Target levels of inputs for given levels of outputs are implied by the point on the DEA best practice efficiency frontier that a particular inefficient TSO is projected onto. It may be desirable that output targets reflect a balance of regulatory priorities. If variables that measure the relevant dimensions of service are not included in the candidate variables, this will not be achieved in the benchmarking study. The choice of variables included in the final analysis may have an influence on the incentives of regulated businesses. Agrell and Bogetoft have emphasised that more research is needed into how to identify and quantify the incentives for regulated businesses produced by different benchmarking models, which would assist to compare the likely effectiveness of alternative models if applied (Agrell and Bogetoft, 2016, p. 33).

APPENDIX A: DATA COLLECTION

This appendix presents a summary of the types of information gathered by ACM and other regulators from gas and electricity TSOs. The aim is to comment on the types of data that may need to be collected in an electricity or gas TSO benchmarking study, and particularly the additional information that could usefully be collected, compared to previous data collected from European TSOs.

A.1 Introductory Comments

The data collected should include, at a minimum, information corresponding to that previously collected *and used*. This would enable previous methodologies to be applied and tested. Even if a different methodology were used in a new benchmarking exercise, it would be useful to see how previously used methods perform with more recent data, since this would also assist model development.

Some other points to note:

- If there is a widening of the data reporting requirements to include information not previously collected, then it is desirable that the regulated businesses should report the same wider set of information for the previous collection period. This would enable the benchmarking analysis to use data for both periods, which would reduce the impact of idiosyncratic factors that are only applicable to a single year.
- Since there is a gap of several years between the previous and current data collection periods, it would be preferable that the businesses report the same information for each of the intervening years, thereby developing a continuous data series over time. This information is likely to be important for modelling accuracy, and assists future benchmarking by developing a continuous data series over time.
- In previous information collections, capital expenditure and some other information has been collected for an extended period from the 1960s or 1970s, and presumably the historical information will not need to be collected again. Operational data was only collected for one year, whereas the financial data was collected for two years. We would suggest that the operational and financial data should both be collected for the same periods, otherwise the more limited availability of the operational data may be a constraint on the observations that can be used in analysis.
- It would be preferable for future electricity and gas TSO data gathering to take into account a similar range of information.

It should be noted that good data collection practice requires templates circulated for the collection of subsequent years' data to include the data supplied by the TSO for earlier years. This maximises the opportunity for TSO staff supplying the data and analysts processing the data to ensure consistency in the data supplied across years. This is particularly an issue when staff completing the templates change over time. If the latest year's data is collected in the template without listing of earlier years' data supplied by the TSO there is a greater chance for errors and inconsistencies to enter the benchmarking database, because it provides the means of cross-checking against data previously reported and identifying the nature of discrepancies. If discrepancies arise because incorrect data was previously reported then the historical data should be corrected by the TSO. This process enables identified errors in historical data to be corrected, which would not be possible if only the data for the latest year is collected.

The following sections compare ACM data collection for European TSOs (mainly gas TSOs) and the Australian Energy Regulator's data collection from electricity TSOs. The information to be collected is grouped under the following headings:

(i) Financial data

(ii) Physical capital data

(iii) Operational data

(iv) Service quality data

(v) Operating environment data.

A.2 Financial Data

This section considers three main types of financial data collected from TSOs: revenues; operating expenses; and capital expenditure and asset data. It is important to ensure that the data collected covers a common range of activities across TSOs and that differences in functional boundaries across TSOs are adjusted for. It is also important to ensure common cost allocation methodologies are used as much as possible to ensure like is being compared with like.

A.2.1 Revenue

ACM's information collection templates for gas TSOs include data for total direct revenues by 10 activity categories, namely: (1) grid ownership & financing; (2) grid planning; (3) grid construction; (4) grid maintenance; (5) gas transport and metering: gas storage operations; (6) gas transport and metering: LNG terminal operations; (7) gas transport and metering: system operations; (8) gas transport and metering: market facilitation; (9) administration; and (10) any other activity. Several of these activities are excluded from benchmarking. For electricity TSOs the activity categories were: (a) market facilitation; (b) system operations; (c) grid planning; (d) grid construction; (e) grid maintenance; and (f) general admin & overhead.

By way of contrast, this section describes the revenue data collected from electricity TSOs by the Australian Energy Regulator (AER), whilst noting that it uses different benchmarking methodologies (including multilateral TFP indexes). The AER collects electricity TSO revenue data using two different breakdowns. The first revenue breakdown is by chargeable quantity using the following headings.

- Fixed Customer (Exit Point)
- Variable Customer (Exit Point)
- Fixed Generator (Entry Point)
- Variable Generator (Entry Point)

- Fixed Energy Usage Charges (Charge per day basis)
- Variable Energy Usage charges (Charge per kWh basis)
- Energy based Common Service and General Charges
- Fixed Demand based Usage Charges
- Variable Demand based Usage Charges
- Revenue from other Sources
- Total revenue.

The second revenue breakdown is by type of connected equipment, which is collected by the AER in the following categories.

- Other connected transmission networks
- Distribution networks
- Directly connected end–users
- Generators
- Other revenue
- Total revenue.

In addition to the foregoing, the AER collects information on TSO revenues allowed, or penalties deducted, due to incentive schemes with detail by incentive scheme.

As indicated in the discussion of functional versus billed output measures in in chapter 4, the billing quantities associated with revenues can be useful candidates as outputs, so information on revenue broken down by billing categories may be useful within a benchmarking exercise. However, if revenue by chargeable quantity were collected, one would need to collect both the monetary amounts and the associated billing unit quantities in each category. Similarly, if information on revenue by type of connected equipment were collected, one would need to collect the number of customers in addition to the monetary revenue amounts, for each category.

A.2.2 Operating Expenses

ACM's data templates for gas TSOs specify for the same ten activity classes previously listed, for both electricity and gas TSOs data is collected for the following expenditure categories.[16]

- Direct manpower cost
- Direct cost of purchased services
- Direct cost of expensed goods (excl. energy)
- Direct cost of energy

[16] The electricity template has an additional category for 'depreciation of non-benchmarking related assets'.

- Depreciation of non-grid-related assets
- Depreciation of grid-related assets
- Leasing fees
- Indirect cost and overhead
- Other costs
- Total costs.

In general, these opex categories seem to be sufficiently detailed, but note:

(a) It may be useful to specifically include Amortisation with depreciation to ensure it is not reported under ‘other costs’;

(b) Greater detail on Changes in Provisions may be needed. Although changes to provisions for employee entitlements would presumably be included under manpower costs, changes to other provisions are sometimes excluded from the costs used in benchmarking analysis.

(c) It may also be desirable to include separate line items for interest expenses and tax expenses to ensure that they are treated consistently between TSOs. These are usually excluded from costs used in benchmarking analysis.

(d) It may be useful for the direct cost of energy to be separated into: (i) gas purchased for on-sale; (ii) gas cost associated with compressor fuel use; (iii) gas cost associated with line losses; (iv) other energy purchases. This is because it would be desirable to exclude the cost of gas purchased for on-sale (if any) from expenses, and to identify any differences between TSOs in relation to expenses incurred in relation to line losses.

A.2.3 Capital Expenditure & Asset Values

The data collected previously by ACM for gas TSOs provided for a breakdown of historical capital expenditure into:

- Pipelines
- Pressure regulator/controller stations and compressor stations, connection points
- SCADA, telecom
- Other equipment
- Other categories of capitalised expenses such as capitalised tax.

For the electricity TSOs, the operational data includes asset-by-asset data on: asset definition; asset specification; cost; voltage code; capital expenditure; number of units; and the periods in which the assets were installed.

This section discusses collection of capital expenditure and asset values by broad categories. Financial information on assets at a detailed or individual asset level is discussed in the context of physical capital data in section A.3.

Again, the AER's data collection for electricity TSOs can be used as a comparator. The AER collects data for the regulatory asset values and the full roll-forward calculations from year-to-year in the following asset classes:

- Overhead transmission assets
- Underground transmission assets
- Switchyard, substation and transformer assets
- Easements
- Other assets with long lives
- Other assets with short lives.

The roll-forward calculations show: the opening asset value; revaluations (ie the inflation adjustment); additions (capital expenditure); disposals; depreciation (separately for regulatory depreciation and straight line depreciation); and closing value. In addition, for each of these asset classes, the average total asset life and the average remaining asset life are reported.

Although regulatory asset data may not be sufficiently comparable in a multi-jurisdictional setting such as Europe, the decomposition of changes in asset value from year-to-year shown in the roll-forward calculations, and breakdowns of asset values into asset classes, can be useful information for further analysis, including, for example, developing alternative measures of capital inputs.

A.3 Physical Capital Data

The ACM's previous data collection for gas TSOs included physical data for the following asset types: pipelines; pressure regulator/controller stations and compressor stations; and connection points. The types of data collected for each of these asset classes is summarised in Table A.1.

In this section a comparison is again made with the AER's data collection for electricity TSOs, mainly to compare the overall scope of the information collected. There are three broad groups of assets for which physical data is collected: overhead transmission assets; underground transmission assets; and switchyard, substation and transformer assets. Table A.2 shows the physical capital data collected by the AER.

(a) The ACM data collects a small number of physical characteristics variables for each asset class, but supplements this with a great deal of data reported for individual assets. The data reported at the aggregate level (but by two gas quality categories) include:

  - o the total number, power and energy use of compressors
  - o the aggregate length and volume of pipelines
  - o the number of injection points, delivery points (by broad class) and interfaces with storage facilities.

  This aggregate data may not be sufficient as a physical description of the network, so that some of the information in the more detailed data may need to be aggregated, such as pipeline pressures or maximum hourly capacity at delivery points. It is likely that the Normalised Grid variables are calculated using the detailed by-asset data.

(b) The AER's electricity TSO data collection also includes a small number of physical characteristics variables for each asset class, and most of these are broken down into seven voltage classes. The aggregate measures include:

- overhead transmission lines circuit length and capacity in MVA, and similar data for underground transmission lines

Table A.1: **ACM Gas TSO Physical Asset Data**

Data item	**Aggregate info. (by H and L gas)**	**Asset detail**
Compressors	No. Total power (MW) Total energy use (MWh)	*Pressure regulator & compressor station detail:* Station ID Location Nominal pressure (bar) Pressure controlled (bar) Upstream pressure (bar) Downstream pressure (bar) Flow rate (m^3/h) Mechanical efficiency (%) Share of external usage (%)
Pipeline network	Length (km) Volume (m^3)	*Pipe section detail:* ID / name Length (km) Volume (m^3) External use share Gas quality (H/L) Pressure range (LP/MP/H1/H2/H3/H4) Material Class (S1/S2/S3/I/PE/PVC) Diameter Class (A/B/C/D/E/F/G) Location - pipe section start Location - pipe section end
Connection points	Injection points (#) Delivery points: - downstream network (#) - final customer (#) - neighbouring network (#) Storages (#)	*Connection point detail:* Connection point ID Location Point also regulator? (yes/no) Pressure level (bar) Type of connection point (I/D/C/N/S)

Data item	Aggregate info. (by H and L gas)	Asset detail
	Total connection points (#)	Gas quality (H/L) Single or shared use Ownership Max Pressure (bar) Min Pressure (bar) Capacity injection peak (m^3/h) Capacity delivery peak (m^3/h) Commercial settlement

- capacities of transmission substations; terminal points to distribution networks; transformers for directly supplied end-users; and (inter-state) interconnectors.
- number of connection points (entry and exit).

Some of the surveyed studies preferred route line length to circuit length, and this information is collected by the AER among the operating environment characteristics.

(c) The key difference in scope between the ACM's and the AER's data collection is the detailed by-asset physical data collected by the former. The AER data is broken down into voltage classes of assets, but does not drill down to specific assets. Experience suggests the AER's asset-related information could be more detailed for some asset classes. That said, there may be an intermediate option that provides sufficient data by asset type and class (eg voltage class for electricity or pressure range for gas), and perhaps by location, while requiring less detailed by-asset data provision by the TSOs.

Table A.2: **AER Electricity TSO Physical Asset Data**

Data item	Aggregate info.	Asset detail
Overhead transmission assets	Circuit length (km) at each voltage class* Weighted average capacity (MVA) by voltage class*	
Underground transmission assets	Circuit length (km) at each voltage class* Weighted average capacity (MVA) by voltage class*	
Switchyard, substation and transformer assets	Transmission substations capacity (MVA) Terminal points to distribution network systems – total capacity (MVA) Transformer capacity for directly connected end–users owned by the	

Data item	Aggregate info.	Asset detail
	TNSP (MVA) Transformer capacity for directly connected end–users owned by the end–user (MVA) Interconnector capacity (MVA)	
Connection points	Number of entry points at each transmission voltage level* Number of exit points at each transmission voltage level*	

* 500 kV, 330 kV, 275 kV, 220 kV, 132 kV, 66 kV, 33 kV.

A.4 Operational Data

The operational data collected for European gas TSOs included:

- Gas quantities (annually) injected in total, and from storage
- Gas quantities (annually) delivered to:
  - DSOs
  - end-users
  - other TSOs in the national market area
  - neighbouring countries, and
  - storage.
- Gas quantities (annually) used for own consumption; and network losses
- Peak injections (maximum hour concurrent)
- Peak deliveries (maximum hour concurrent).

These data were reported in kWh and cubic metres (m^3). The former is an unconventional unit for the energy content of gas. Terajoules (TJ) are a more conventional measure of energy used in the gas industry. Further, the peak day is a more common measure for planning purposes in the gas transmission industry than is peak hour. Perhaps because of the effects of line pack and variations in line pressure, the timeframes within which system balance is maintained are somewhat different to electricity networks.

The electricity TSO operational data collected by the AER includes:

- energy deliveries (annually) in MWh:
  - to other connected transmission networks
  - to distribution networks
  - to directly connected end–users, and
  - the total energy transported.
- system maximum hour demand (annually), in both MW and MVA, including:

- coincident maximum demand
  - coincident weather adjusted maximum demand 10% POE
  - coincident weather adjusted maximum demand 50% POE
  - non-coincident summated maximum demand
  - non-coincident weather adjusted summated maximum demand 10% POE
  - non-coincident weather adjusted summated maximum demand 50% POE.
- power factor (for conversion between MVA and MW) by line voltage class.

Peak demand is an important cost driver, but peak demand in a particular year is only useful for benchmarking purposes if there is little volatility from year to year. Since energy demand is generally sensitive to weather, and peak demand especially so, there is usually volatility in the peak demand from year to year. The weather-adjusted measure of maximum demand with a relatively small probability of exceedance (POE) is the more relevant measures for benchmarking purposes (although for network planning purposes a much smaller POE, such as 2%, is more relevant).

If the peak demand data is available for a sufficient number of years, weather correction and probability distributions for peak demand can be calculated, but usually TSOs should be able to provide this information.

A.5 Quality data

The AER collects a range of service quality data from electricity TSOs. This type of data does not seem to be collected for the European TSOs. Quality data collected from Australian electricity TSOs include:

- Average Circuit outage rates (number and per cent):
  - Line fault outage rate
  - Transformer fault outage rate
  - Reactive plant fault outage rate
  - Lines forced outage rate
  - Transformer forced outage rate
  - Reactive plant forced outage rate.
- Loss of supply event frequency (number):
  - Number of events greater than target minutes per annum (more than one target)
- Average outage duration (minutes)
- Proper operation of equipment – number of failure events
  - Failure of protection system
  - Material failure of SCADA system
  - Incorrect operational isolation of primary or secondary equipment.
- System losses (per cent).

Quality-of-service measures have been used in some energy network benchmarking studies, sometimes as ‘bads’ (for example outages). It is desirable to collect data of this kind because without it the measurement of output characteristics would be incomplete, which may be detrimental to benchmarking analysis.

It will usually be necessary to also allocate a value to energy not supplied due to outages. This is often done by using estimates of the value consumers place on reliability although this is sometimes capped to avoid distortions to the total output measure resulting from unusual one–off outages.

A.6 Operating Environment Characteristics Data

The literature survey has shown that many studies have not included operating environment characteristics and among those that have, a variety of variables have been used. It will be difficult to specify with confidence the data needed with regard to operating environment characteristics, without unnecessary data collection and/or the omission of important data. The main types of operating environment variables are: climate, terrain, network configuration, and regulations and standards.

The ACM gas TSO data templates require information by asset, which includes information on terrain related to pipeline sections, namely:

- Land use
- Soil subsurface
- Topography, and
- Soil humidity.

The detailed by-asset data also includes the locations of network assets from which network configuration indicators can be calculated.

The AER's electricity TSO data includes some indicators of network configuration such as: route line length; variability of despatch; concentrated load-distance; and total number of spans. Information is provided by TSOs on the IDs of all weather stations in their network areas, and weather data can be obtained for these weather stations from the national weather bureau. The terrain data includes:

- Total number of vegetation maintenance spans
- Average vegetation maintenance span cycle (years)
- Average number of trees per vegetation maintenance span
- Average number of defects per vegetation maintenance span
- Tropical proportion
- Vehicle access (km)
- Altitude (km)
- Bushfire risk (number of spans).

The terrain information relevant to overhead electricity transmission networks will differ from that relevant to gas transmission pipelines. Some measure of vegetation management is relevant in the Australian electricity TSO context.

At present, network configuration indicators for European TSOs appear to be calculated from detailed location data by asset. With enough analysis this information might enable a number of useful measures to be calculated, but with a risk of lack of information on some attributes, such as the density of demand centres. Consideration could be given to developing and defining useful network configuration indicators and collecting these measures directly from the TSOs, rather than calculating them from the disaggregated data. The literature review suggests that a measure of the size of the TSO's supply area may be relevant, and information about the downstream markets, such as their size and density, may also be relevant.

The collection of weather station identification numbers from TSOs, rather than weather data, seems to be a useful approach that could presumably also be adopted for the European TSOs. In addition to the more commonly used climate variables, some studies of electricity TSOs have used the amount of atmospheric salinity in particular areas as a relevant factor.

It is also notable that one of the studies discussed in the literature review used the mix of types of generation, particularly the amount of (centrally generated) renewable versus non-renewable energy, as a relevant operating environment variable for electricity transmission.

REFERENCES

Adler, N., Yazhemsky, E., 2010. Improving discrimination in data envelopment analysis: PCA–DEA or variable reduction. Eur. J. Oper. Res. 202, 273–284. https://doi.org/10.1016/j.ejor.2009.03.050

Agrell, P., Bogetoft, P., 2009. International Benchmarking of Electricity Transmission System Operators, e3Grid Project: Final Report.

Agrell, P., Bogetoft, P., 2007. Project Gerner/AS6: Development of benchmarking models for German electricity and gas distribution: Final Report.

Agrell, P., Bogetoft, P., 2006. ECOM+ Results 2005: Final Report (Sumicsid).

Agrell, P., Bogetoft, P., Beaussant, H., Talarmin, J., 2014. Project PE2GAS: Benchmarking European Gas Transmission System Operators: A Feasibility Study, Final Report.

Agrell, P.J., Bogetoft, P., 2016. Regulatory Benchmarking: Models, Analyses and Applications.

Agrell, P.J., Bogetoft, P., 2014. International benchmarking of electricity transmission system operators, in: 2014 11th International Conference on the European Energy Market (EEM). IEEE, pp. 1–5.

Agrell, P.J., Bogetoft, P., Trinkner, U., 2016. Project E2GAS: Benchmarking European Gas Transmission System Operators (Sumicsid/Swiss Economics).

Aivazian, V., Callen, J., Chan, L., Mountain, D., 1987. Economies of Scale versus Technological Change in the Natural Gas Transmission Industry. Rev. Econ. Stat. 556–561.

Burns, P., Weyman-Jones, T.G., 1996. Cost functions and cost efficiency in electricity distribution: a stochastic frontier approach. Bull. Econ. Res. 48, 41–64.

Cadena, A., Marcucci, A., Perez, J.F., Duran, H., Mutis, H., Palacios, F., 2009. Efficiency Analysis in Electricity Transmission Utilities. J. Ind. Manag. Optim. 5, 253–274.

Caves, D.W., Christensen, L.R., Diewert, E., 1982. Multilateral Comparisons of Output, Input, and Productivity Using Superlative Index Numbers. Econ. J. 92, 73–86.

Coelli, T., Estache, A., Perelman, S., Trujillo, L., 2003. A Primer on Efficiency Measurement for Utilities and Transport Regulators, WBI Development Studies.

Cooper, W.W., Seiford, L.M., Tone, K., 2006. Introduction to data envelopment analysis and its uses: with DEA-solver software and references. Springer, New York.

Daraio, C., Simar, L., 2007. Advanced robust and nonparametric methods in efficiency analysis: methodology and applications, Studies in productivity and efficiency. Springer, New York.

Denny, M., Fuss, M., Waverman, L., 1981. The Measurement and Interpretation of Total Factor Productivity in Regulated Industries, with an Application to Canadian Telecommunications, in: Cowing, T.G., Stevenson, R.E. (Eds.), Productivity Measurement in Regulated Industries. Academic Press.

Dismukes, D., Cope, R., Mesyanzhinov, D., 1998. Capacity and economies of scale in electric power transmission. Util. Policy 155–162.

Economic Insights, 2017. Review of Economic Benchmarking of Transmission Network Service Providers – Issues Paper (Report prepared for Australian Energy Regulator).

Economic Insights, 2014. Economic Benchmarking Assessment of Operating Expenditure for NSW and Tasmanian Electricity TNSPs (Report prepared for the Australian Energy

Regulator).

Economic Insights, 2013a. Measurement of Outputs and Operating Environment Factors for Economic Benchmarking of Electricity Transmission Network Service Providers (Briefing Notes prepared for Australian Energy Regulator).

Economic Insights, 2013b. Measurement of Inputs for Economic Benchmarking of Electricity Network Service Providers (Briefing Notes prepared for Australian Energy Regulato).

ERGEG, 2007. Gas Transmission Tariffs: An ERGEG Benchmarking Report (No. C06-GWG-31-05).

Frontier Economics, Consentec, 2016. Gas TSO efficiency analysis for the Dutch transmission system operator.

Frontier Economics, Consentec, Sumicsid, 2013. E3GRID2012 – European TSO Benchmarking Study.

Gilsdorf, K., 1994. Vertical Integration Efficiencies and Electric Utilities: A Cost Complementarity Perspective. Q. Rev. Econ. Finance 34, 261–282.

Granderson, G., 2000. Regulation, open-access transportation, and productive efficiency. Rev. Ind. Organ. 16, 251–266.

Hawdon, D., 2003. Efficiency, performance and regulation of the international gas industry—a bootstrap DEA approach. Energy Policy 31, 1167–1178.

Huettner, D., Landon, J., 1978. Electric Utilities: Scale Economies and Diseconomies. South. Econ. J. 883–912.

Hunt, L.C., Lynk, E.L., 1995. Privatisation and efficiency in the UK water industry: an empirical analysis. Oxf. Bull. Econ. Stat. 57, 371–388.

Jamasb, T., Newbery, D., Pollitt, M., Triebs, T., 2007. International Benchmarking and Regulation of European Gas Transmission Utilities: Final Report.

Jamasb, T., Pollitt, M., Triebs, T., 2008. Productivity and efficiency of US gas transmission companies: A European regulatory perspective. Energy Policy 36, 3398–3412. https://doi.org/10.1016/j.enpol.2008.05.001

Jenkins, L., Anderson, M., 2003. A multivariate statistical approach to reducing the number of variables in data envelopment analysis. Eur. J. Oper. Res. 147, 51–61.

Kim, T.-Y., Lee, J.-D., Park, Y.H., Kim, B., 1999. International comparisons of productivity and its determinants in the natural gas industry. Energy Econ. 21, 273–293.

Kneip, A., Simar, L., Wilson, P.W., 2016. Testing Hypotheses in Nonparametric Models of Production. J. Bus. Econ. Stat. 34, 435–456. https://doi.org/10.1080/07350015.2015.1049747

Kneip, A., Simar, L., Wilson, P.W., 2015. When Bias Kills the Variance: Central Limit Theorems for DEA and FDH Efficiency Scores. Econom. Theory 31, 394–422. https://doi.org/10.1017/S0266466614000413

Kneip, A., Simar, L., Wilson, P.W., 2008. Asymptotics and Consistent Bootstraps for DEA Estimators in Nonparametric Frontier Models. Econom. Theory 24, 1663–1697. https://doi.org/10.1017/S0266466608080651

Lee, J.-D., Oh, K., Kim, T.-Y., 1999a. Productivity growth, capacity utilization, and technological progress in the natural gas industry. Util. Policy 8, 109–119.

Lee, J.-D., Park, S.-B., Kim, T.-Y., 1999b. Profit, productivity, and price differential: an international performance comparison of the natural gas transportation industry. Energy Policy 679–689.

Llorca, M., Orea, L., Pollitt, M., 2013. Efficiency and Environmental Factors in the US Electricity Transmission Industry (No. EPRG Working Paper 1310). University of Cambridge.

Llorca, M., Orea, L., Pollitt, M.G., 2014. Using the latent class approach to cluster firms in benchmarking: An application to the US electricity transmission industry. Oper. Res. Perspect. 1, 6–17. https://doi.org/10.1016/j.orp.2014.03.002

Meyrick and Associates, 2004. Comparative Benchmarking of Gas Networks in Australia and New Zealand.

Nataraja, N.R., Johnson, A.L., 2011. Guidelines for using variable selection techniques in data envelopment analysis. Eur. J. Oper. Res. 215, 662–669. https://doi.org/10.1016/j.ejor.2011.06.045

Nieswand, M., Cullmann, A., Neumann, A., 2010. Overcoming data limitations in nonparametric benchmarking: Applying PCA-DEA to natural gas transmission. Rev. Netw. Econ. 9, 1–20.

Ofgem, 2011. Decision on strategy for the next transmission price control - RIIO-T1 Tools for cost assessment.

Pastor, J.T., Ruiz, J.L., Sirvent, I., 2002. A statistical test for nested radial DEA models. Oper. Res. 50, 728–735.

Paulun, T., Haubrich, H.-J., Maurer, C., 2008. Calculating the efficiency of electricity and natural gas networks in regulated energy markets, in: Electricity Market, 2008. EEM 2008. 5th International Conference on European. IEEE, pp. 1–5.

Peebles, M.W.H., 1992. Natural Gas Fundamentals. Shell International Gas Limited.

Pollitt, M.G., 1995. Ownership and Performance in Electric Utilities: The International Evidence on Privatization and Efficiency. Oxford University Press.

Rencher, A.C., Christensen, W.F., 2012. Methods of multivariate analysis, Third Edition. ed, Wiley series in probability and statistics. Wiley, Hoboken, New Jersey.

Saal, D.S., Parker, D., 2000. The impact of privatization and regulation on the water and sewerage industry in England and Wales: a translog cost function model. Manag. Decis. Econ. 21, 253–268. https://doi.org/10.1002/mde.988

Sickles, R., Streitwieser, M., 1992. Technical Inefficiency and Productive Decline in the U.S. Interstate Natural Gas Pipeline Industry Under the Natural Gas Policy Act. J. Product. Anal. 3, 119–133.

Simar, L., Wilson, P.W., 2004. Performance of the Bootstrap for DEA Estimators and Iterating the Principle, in: Cooper, W.W., Seiford, L.M., Zhu, J. (Eds.), Handbook on Data Envelopment Analysis. Kluwer Academic.

Thanassoulis, E., Portela, M., Despic, O., 2008. Data Envelopment Analysis: The Mathematical Programming Approach to Efficiency Analysis, in: Fried, H.O., Lovell, K., Schmidt, S.S. (Eds.), The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press.

von Geymueller, P., 2009. Static versus dynamic DEA in electricity regulation: the case of US transmission system operators. Cent. Eur. J. Oper. Res. 17, 397–413.

von Geymueller, P., 2007. The Efficiency of European Transmission System Operators: An Application of Dynamic DEA (No. Working Paper No. 3). Research Institute for Regulatory Economics, Vienna University Economics and Business Administration.

Warner, R.M., 2013. Applied Statistics: From Bivariate Through Multivariate Techniques, 2nd ed. SAGE Publications Inc.

Yu, W., Jamasb, T., Pollitt, M., 2009. Does weather explain cost and quality performance? An analysis of UK electricity distribution companies. Energy Policy 37, 4177–4188. https://doi.org/10.1016/j.enpol.2009.05.030