Productivity and Efficiency of US Gas Transmission Companies:

A European Regulatory Perspective

Tooraj Jamasb [a]

Michael Pollitt [b]

Thomas Triebs [c]

1 March 2008

---

[a] Faculty of Economics, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DE, UK.

[b] Judge Business School, University of Cambridge, Trumpington Street, Cambridge CB2 1AG, UK.

[c] Corresponding author. Telephone: +44-(0)1223-339700, Fax: +44-(0)1223-339701, E-mail: t.triebs@jbs.cam.ac.uk, Judge Business School, University of Cambridge, Trumpington Street, Cambridge CB2 1AG, UK.

Abstract

On both sides of the Atlantic the regulation of gas transmission networks has undergone major changes since the early 1990's. Whereas in the US the long-standing regime of cost-plus regulation was complemented by increasing pipe-to-pipe competition, most European countries moved towards incentive regulation complemented by market integration. We study the impact of US regulatory reform using a Malmquist-based productivity analysis for a panel of US interstate companies. Results are presented for changes in productivity, as well as for several convergence tests. The results indicate that taking productivity and convergence as performance indicators, regulation has been rather successful, in particular during a period where overall demand was flat. Lessons for European regulators are twofold. First, the US analysis shows that benchmarking of European transmission operators would be possible if data were available. Second, our results suggest that, in the long-run, market integration and competition are alternatives to the current European model.

Keywords: Natural gas transmission; utility regulation; data envelopment analysis; total factor productivity; convergence

JEL classification: L51, L95, O57, D24

1 Introduction

Gas transmission networks are regulated rather differently in Europe and in the US. Whereas US regulation is shifting its focus from cost to value by complementing cost-of-service regulation with institutions fostering competition and market integration (O'Neill, 2005), European regulators treat gas transmission as incentive-regulated franchise monopolies (Makholm, 2007). Both, traditional US cost-of-service or rate-of-return regulation and European incentive regulation are based on the notion of natural monopoly.[1]

However, unlike electricity or gas distribution, gas transmission networks are not necessarily natural monopolies. The notion of natural monopoly relates to efficient service provision for a given market and not to the economies of scale at the level of a single pipeline. Kahn (1971) argues that although there exist economies of scale in relation to pipe diameter, markets for gas can easily be served by several pipelines, be it from different supply areas or not. O'Neill (2005) makes a similar point when stressing that natural gas pipelines are oligopolies rather then monopolies; and non-traditional approaches to regulation might provide superior results.

Whereas O'Neill (2005) makes the case against traditional cost-of-service regulation in the US, a similar case could be made against the European model of benchmarking-based incentive regulation for franchise monopolies.[2] A benchmarking-based model, as explored in a recent report commissioned by the Council of European Regulators

---

[1] The authors would like to thank the UK Economic and Social Research Council (ESRC) for supporting this study. Also, the authors would like to thank one anonymous reviewer.

[2] Clearly, there is not a single European approach to regulation but most regulators seem to follow the example of incentive regulation (RPI-X) regulation as set by the British and Dutch regulators.

(CEER)[3], is certainly a viable short to mid-term solution for the regulation of gas transmission in Europe. However, the characteristics of gas transmission and the, as we argue below, rather positive performance of the US industry under the new regulatory framework suggest that in the long-run a focus on market integration and competition might provide additional benefits.

Makholm (2007) describes in much detail the institutional framework that the Federal Energy Regulatory Commission (FERC), the US regulator for interstate transmission pipelines, put in place to complement its traditional rate-of-return regulation with competition between pipelines. Makholm argues that competition and market integration greatly improved the resilience of markets and therefore security of supply in the US and that Europe could achieve the same by similar means.

In this paper we look at another characteristic of US industry performance to assess recent regulatory change: productivity. We measure efficiency, efficiency change, technical change, total factor productivity change, and convergence for a sample of regulated US interstate gas transmission pipelines using Data Envelopment Analysis, Malmquist indices, and a regression analysis to test for convergence.

The productivity of US gas transmission pipelines has been explored in the literature using firm-level data for earlier periods. Aivazian (1987) measures productivity growth of the US gas transmission industry as well as its constituent parts (for labour productivity) including scale efficiency. The main finding is that the contribution of technical change is at least as large as the contribution of scale economies. There is also a literature on the effect of regulatory change on US transmission companies. Examples are Sickles and Streitwieser (1991), Sickles and Streitwieser (1998), and Granderson (2000). Together these papers show that technical efficiency fell after well-head price deregulation in 1978 due to increasing prices and falling consumption (Sickles and

---

[3] In our report for the CEER (Jamasb et al., 2007), we benchmarked several European gas transmission operators against a sample of US interstate transmission companies. Unfortunately, we were not able to include any data for European operators in this work. The often stunning differences in transparency between the US and Europe are discussed by Makholm (2007). The US approach on transparency is discussed for instance by Olsen (2005).

Streitwieser, 1991) and that the regulatory change requiring third-party access in the mid 1980's led to small average cost reductions and diverging performance (Granderson, 2000).

Thus, we ask the following two sets of questions. First, does average industry productivity increase and does technical efficiency converge at the firm level for our sample period? We know from the literature that when the old organization of the gas industry unravelled during the 1980s productivity of interstate transmission pipelines fell. Given that our sample (1996-2004) starts several years after the latest regulatory push for more competition in 1992 (FERC Order 636), we would expect to observe increasing efficiency and possibly convergence in contrast to earlier periods. Also, even though we are not able to include a control group here, we presume that original cost-of-service regulation provides only weak incentives for performance improvements. Our results, therefore, can give an indication of how successful regulatory change has been in bringing about such improvements. This brings us to our second question: What lessons can European regulators learn from the relative success of the changing US regulatory regime for gas transmission?

In particular, we hope that our results can help European regulators to define a long-term strategy for the regulation of gas transmission in Europe and contribute to the increasing dialogue between regulators across the Atlantic.[4] Lastly, we would like to stress that all our conclusions apply to gas transmission only. Both in the US and in Europe different energy networks are regulated in different ways and with varying levels of success.

This paper is organized as follows: Section 2 gives the background and in particular describes the development of the US market and regulation; section 3 describes our models and discusses variable selection; section 4 describes the data; section 5 presents the results; and section 6 discusses the results and concludes.

---

[4] Since 2000 a yearly US-European Roundtable of regulators takes place, where experiences are shared. See for instance: http://www.ceer-eu.org/portal/page/portal/CEER_HOME/CEER_PUBLICATIONS/CEER_DOCUMENTS/2007/EU-US%20Roundtable_closing_statement_final.pdf.

2 Background

The challenge for any regulator is to increase efficiency and reduce price, as stated by the European Commission in its second Gas Directive ("Acceleration Directive", 2003). Although the process of European gas market liberalization and integration commenced in the mid 1990's, the Commission acknowledges in its Acceleration Directive and its recent Energy Sector Inquiry[5] that many obstacles remain. Also, in its Energy Sector Inquiry the Commission recognizes that the US gas market is much more developed than European gas markets, although it does not discuss differences in regulation. It might be comforting however that the transition in the US from a vertically integrated, geographically fragmented, and heavily regulated industry to an increasingly integrated and lightly regulated industry has been a long tale of trial and error (Makholm, 2007).

Beginning with the deregulation of well-head prices in 1978, the US natural gas market and its regulation changed dramatically. Though there are many parallels with current efforts in Europe to unbundle, allow third-party access and integrate regional markets one difference is of particular importance here. Whereas most European regulators move towards incentive regulation for the (unbundled) pipeline bits of the value-chain, the US regulator aims at competition through a combination of unbundling, flexible short-term rate setting, strong property rights for holders of contractual capacity, and controlling the abuse of market power. Additionally, and unlike in Europe, the US market is both economically and physically integrated to a large extent.

Already in 1987 about one third of city gate markets received services from multiple pipelines according to Kalt and Schuller (1987)[6]. Doane et al. (2004) argue that regulatory change led to both an integrated US market for gas, a competitive wholesale

---

[5] European Commission (2007).

[6] As cited by Ellig (1993).

market, and competition among, often “virtual”, pipelines.[7] Transportation services have been fully unbundled since 1992 (FERC Order 636) though utility services are often integrated (vertically and horizontally) with other utility and non-utility services in the same firm.[8] Unlike Europe, the US has a common market for gas with a single federal regulator for all interstate commerce. Tariff setting, though still dominated by “original cost-of-service regulation” (O'Neill et al., 1996), is increasingly complemented by non-traditional tariff models (O'Neill, 2005). These fall into two categories: flexible short-term rates that allow the efficient allocation of capacity and incentive rates (e.g. indexed rates). Typically, fixed capacity is purchased long-term at regulated or negotiated rates (Granderson, 2000). Unused firm capacity is then released in the short-term in a secondary market, where prices are allowed to rise above the regulated rates and pipelines compete with released capacity in these markets. Hirschhausen (2006, p. 12) summarizes US regulation as follows:

> *"Contrary to Europe, where pipeline companies have a high degree of market power, the pipeline business in the US is competitive in many of the regions. Most destination markets are served by several competing pipelines. Thus, pipelines compete for shippers, and rates are negotiated in a competitive environment. On the other hand, there remains a formal cost-of-service regulation of interstate pipelines.”*

Thus, whereas European regulators aim at incentive regulation for monopolies, FERC aims at complementing traditional rate-of-return regulation with competition through encouraging (or mandating) the development of the necessary market institutions. Although the means differ, regulators on both sides of the Atlantic have the same

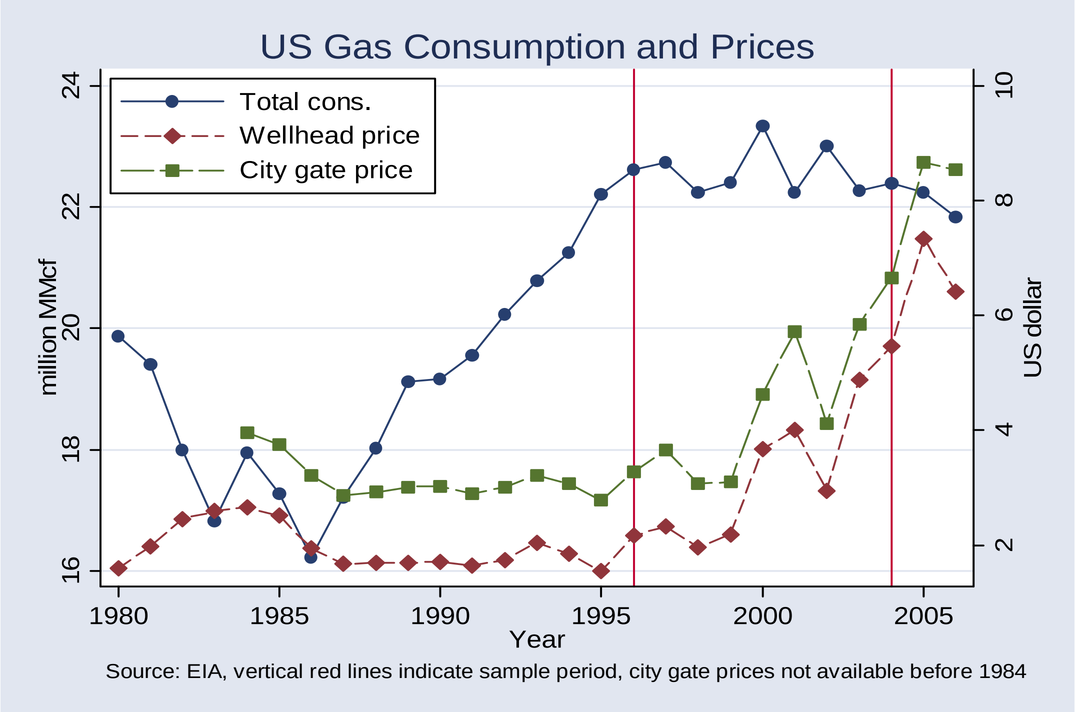
---

[7] The observation that there is pipe-to-pipe competition obviously runs counter to the natural monopoly argument. Here we do not argue the case for or against natural monopoly as done for instance by Ellig (1993), Aivazian (1987) or Hirschhausen et al. (2007) but simply take the observation from the literature that there is nascent competition.

[8] Johnson et al. (1999) show that in 1997 most interstate pipelines belong to about 14 parent companies. Of these 14 companies, 8 also own local distribution companies and 12 own energy marketing services as well.

objective: increasing efficiency[9] and passing on any resulting gains to consumers.[10] In the spirit of Shleifer (1985) both approaches should provide similar incentives for (static) efficiency increases.[11]

As Figure 1 shows, regulatory change was accompanied by a large expansion in consumption which might be a first indication of the success of the overall regulatory change in the US. However, our sample period is characterized by fluctuating and slightly downward trending consumption as well as increasing prices.



Source: EIA, vertical red lines indicate sample period, city gate prices not available before 1984

Figure 1: US Gas Consumption and Prices

---

[9] Alger and Toman (1990) report on auction experiments commissioned by FERC that investigated the effect of increasing competition and different ways to implement it. An important result from these experiments is that the introduction of small amounts of competition (i.e. adding alternative routes or more competitors on the same route) can lead to the much improved performance of a stylized network.

[10] Even without competitive pressures the regulatory lag might be sufficient to introduce incentives for cost reduction as argued by Sickles and Streitwieser (1998) and Schmidt (2000).

[11] However, a competitive and integrated market might provide additional advantages that incentive-regulated (and geographically non-integrated) markets do not supply. One example is resilience to shocks as argued by Makholm (2007).

Besides differences in regulation between the US and the EU there are also differences in industry structure. Table 1 provides an "ad hoc" comparison of the two industries.[12] We observe the following: The total number of companies is rather similar; measured by the physical characteristics US companies are bigger; and the US network has more interconnection points. As to the last observation one should note, however, that Europe does not fare worse on the relation of interconnection points to total length of pipelines.

Table 1: US-Europe Comparison of Industry Structure

	US	**Europe (EU-25)**
Number of companies	85 inter-state (Energy Information Agency, 2002) (our sample contains 39)	40 national, 38 regional (European Commission, 2005)
Length of pipeline (miles)	212.000 Mean: 2494 St. Dev.: 3775 (Energy Information Agency, 2002)	18.542 Mean: 515 St. Dev.: 608 (Makholm, 2007)
Capacity	133 bcf/d	57.3 bcf/d (European Commission, 2007)
Interconnection points	308[13] Hubs: 14 (Energy Information Agency, 2003)	79 (GTE website)[14] Hubs: 13 (however, almost all trading on only 6) (European Commission, 2007)

Another important point is the difference in size as measured by the length of pipelines. Though the US mean is several times larger then the European mean, the European mean is twice the minimum of our sample as shown in Table 5 below.

---

[12] We assume 1bcf/d = 28.33 mcm/d. Because of the different sources the various numbers for Europe are not necessarily for the same sample of pipelines.

[13] Counted as the number of pipeline interconnections at hubs and market centres.

[14] GIE system map at: http://gie.waxinteractive3.com/download/gridmap/GTE_OP_150.pdf

3 Model and variable selection

3.1 Model

Given the vast literature applying productivity and efficiency measurement we know that our results are likely to be sensitive to both the model we choose and the variables we select. To summarize, we model efficiency for a given unit (here a pipeline company) in a given year as the distance to an input-oriented, constant returns to scale (CRS), non-parametric frontier. Cumulative productivity change, efficiency change, and technical change between two years are modelled as weighted averages derived from the two respective efficiency scores (and two additional cross-period scores). In particular, we use the Data Envelopment Analysis (DEA) for the static efficiency scores and a Malmquist Productivity Index for productivity change (TFPC) and its decomposition into technical efficiency change (TEC) and technical change (TC). The technical details for the DEA and Malmquist indices can be found in Appendix A. Our single input variable is total cost or revenue. Output variables are total length of pipe, total horsepower rating and, for some models, total delivery volume. Table 2 summarizes our models.

We perform an econometric analysis to test our three candidate outputs or cost drivers. The use of econometrics to determine the relevance of variables prior to employing non-parametric techniques is common practice, see for instance Carrington (2002). As we find autocorrelation and heteroskedasticity in our panel we use the Feasible Generalized Least Squares (FGLS) estimator. Following the critique by Beck and Katz (1995) we also check our FGLS results against the estimation with panel-corrected standard errors (PCSE). As the results are very similar we only report the FGLS results below. Also, we do not include fixed-effects in our econometric model. Although a full set of firm fixed-effects is significant it renders all coefficients insignificant at our chosen 5% level (for the PCSE estimation the qualitative results are the same at a 10% significance level). We also use likelihood ratio tests to compare different models and test the significance of delivery volume as a cost-driver.[15] 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.

Table 2: Models and Variables

	Model 1	**Model 2**	**Model 3**	**Model 4**
Technology	input-orientation, constant-return to scale, non-parametric frontier			
Input	Totex	Revenue	Totex	Revenue
Outputs	Delivery	Delivery		
	Compressor capacity	Compressor capacity	Compressor capacity	Compressor capacity
	Network length	Network length	Network length	Network length

Once we have constructed the technical efficiency scores and the productivity indices, we use these to analyze convergence. Following Alam and Sickles (2000), we use convergence concepts from the macroeconomic growth literature which has established two measures, often referred to as β-convergence and σ-convergence. β-convergence is the notion that companies that start from a lower base grow faster, where β refers to the slope coefficient. σ-convergence assumes that changes in the moments of the distribution over time indicate convergence and σ stands for the variance. We perform a total of three convergence tests. First, we graphically relate the level of technical efficiency in the base year (i.e. the DEA scores) to our Malmquist-based measure of technical efficiency change. Second, following Alam and Sickles (2000) we regress the average year to year growth in the technical efficiency scores (i.e. the DEA score) on

---

[15] All regressions are performed in Stata. To test for autocorrelation and heteroskedasticity we use the *xtserial* and *xttest3* commands respectively. For the FGLS estimation we use the *xtgls* command and for the likelihood ratio test the *lrtest* command.

the logarithm of the technical efficiency scores in the base year. For both of these β-convergence tests a negative relationship would indicate convergence. Last, we analyze σ-convergence following Färe et al. (2006) and produce box plots for the static technical efficiency scores. A narrowing of the distribution is an indication of convergence. Next we discuss our model in more detail.

A non-parametric frontier was chosen because in a regulated environment it is not necessarily given that firms are cost-minimizing. In particular, the regression-based Stochastic Frontier Analysis (SFA) requires the assumption that efficient firms (i.e. the firms on the frontier) are cost-minimizing as the cost function is derived from the profit maximization problem (Button and Weyman-Jones, 1992). Arguably, increased competition between pipelines and the efficiency incentives inherent in the lag between rate cases would allow for maximizing behaviour (Sickles and Streitwieser, 1991). Also, unlike SFA, DEA does not account for measurement error but does not run the risk of introducing specification error as no functional form is required for the construction of the frontier. Regulators or regulated firms often take issue with the CRS assumption. However, as the Malmquist indices allow for different returns to scale across periods this assumption is not very strong. The intuition here is that although returns to scale are constant in any given period, returns can vary across periods (Grosskopf, 1993).[16] Last, we chose input-orientation since in a regulated environment it is likely that pipelines will face a derived demand.[17] Generally, Sickles and Streitwieser (1991) consider DEA (and SFA) a parsimonious way to model efficiency distortions under an "extremely complicated and often contradictory regulatory process", as they describe FERC regulation.

The advantage of the Malmquist Productivity Index (MPI) is that unlike other index number approaches it allows for a distinction between technical change and efficiency change. Note that our Malmquist index is cumulative – i.e. we use the first year in our sample, 1996, as the base year for all indices. For instance, the index for 2001 is based on the observations for 1996 and 2001. An alternative would be to use an incremental

---

[16] Also, when looking at our regression results below, we find that for Model 1 the implied economies of scale are about 0.7 and 0.9 for Model 2 (at the respective means).

[17] Sickles and Streitwieser (1991) also use input orientation in their DEA model.

index with changing base years where the two periods are always adjacent. We opted for the cumulative index because, given our relatively short time series, it provides a smoother path. To confirm that our base year, 1996, is not affected by any event specific to that year, we also calculated cumulative indices based on 1997 (i.e. by dropping 1996). There are no systematic differences between the results and we keep 1996 as our base year.

3.2 Variable selection

Although gas transportation technology is not necessarily complex from an engineering perspective, variable selection from an economic and regulatory perspective is not obvious as different choices produce different results.[18] For this reason we contrast several variable specifications. Generally, our choice of variables is informed by the actual US regulatory framework, the literature and our discussions with regulators. As mentioned above we use a model that for the purpose of our optimization is input-oriented and therefore treats output as the "right-hand side" of our cost model and cost as input. We now discuss our variables one at a time.

First, we turn to outputs or cost-drivers. Much of the literature on gas transmission uses production functions where the prime output is gas delivered and inputs are capital and labour. Callen (1978), for instance, uses an engineering Cobb-Douglas function where delivery is a function of horsepower and line-pipe capacity and a scale factor. Line-pipe capacity is measured in tons of steel as a function of length, diameter, and an assumption of wall thickness. Aivazian (1987) and Sickles and Streitwieser (1991) use delivery volumes weighted by transport distances as output. Alternatively, Granderson (2000) uses compressor fuel as a proxy for output. Construction cost drivers identified by the International Energy Agency (1994) are: length of pipeline, maximum flow required for a day of peak demand, the trade-off between diameter and compressor power rating, as well as the terrain and right of way. We exclude all exogenous factors, as well as right of way. As we have no measure of diameter we use total horsepower

---

[18] Also, if the results were to be used to implement a form of incentive regulation it would have to be taken into account that different models imply different incentive properties.

rating and length of mains as outputs representing capacity or capital.[19] The importance of horsepower is that it allows increasing capacity on a given line. Aivazian and Callen (1981, p. 147) state that, "[…] the line-pipe may take months or years to construct and is clearly the most inflexible input. However, once the line-pipe is in the ground, horsepower capacity may be added fairly continuously to the line to build up capacity."

Note that the exclusion of other capacity measures might affect comparability. According to the International Energy Agency (1994) the "peak problem" might be solved differently in different countries and at different times as well as by different firms. Spare capacity, storage and demand response can all address the issue but might be of different importance under different regulatory regimes and for different companies. For instance, as we do not account for storage, its strategic use as addressed by O'Neill (2005) does not infer an advantage. A company that uses storage more cost effectively than others use horsepower and mains will be disadvantaged. Next, we discuss delivery which we alternatively include or exclude.

When looking at the rationale for including delivery as an output one has to distinguish between total cost and revenue as input variables because as a first approximation it is likely that revenue is more closely related with delivery volumes than total cost (though we find no statistical evidence for this below). When looking at the cost models the arguments for the inclusion of delivery seem weak as most costs are fixed. Even most O&M costs (except compressor fuel and compressor maintenance) are fixed as stated by the International Energy Agency (1994, p. 48).[20] When looking at the revenue models the main reason for the inclusion of delivery is that tariffs include a volume element. However, this element was drastically reduced in 1992 to better reflect the cost decomposition. Before 1992, FERC attempted to restrain market power by forcing companies to recoup their fixed cost via a volume charge. With the development of secondary markets, this is no longer necessary and tariffs increasingly reflect costs as argued by Alger and Toman (1990). Another reason for the inclusion of delivery is that

---

[19] As shown by International Energy Agency (1994, Fig. 2), there is a clear relationship between pipeline diameter and compressor power for a given transport volume per year.

[20] The IEA estimates O&M costs for onshore pipelines to be about 2% of investment cost. Maintenance costs for compressor stations run at a relatively high load factor are estimated to be 3-6% of investment costs.

increased competition and therefore increasingly diverging business models imply different approaches to increasing capacity and delivery in ways that we do not yet account for (e.g. better systems). Assuming a company uses better management or trading to increase delivery with given capacity, a model excluding delivery would not account for this and less innovative companies would be rewarded. Last, our econometric results below do not unambiguously reject delivery as a cost or revenue driver.

On a different note, including delivery causes a technical problem related to our use of Malmquist indices and the fact that delivery shows a rather high year-to-year variability. Coelli et al. (2005, p. 306) and Nghiem and Coelli (2002) explain why variables that fluctuate on a year-to-year basis potentially cause problems for efficiency measurement results. First, as the frontier is calculated using two years only, DEA-based MPI are influenced by stochastic factors.[21] Secondly, a decrease in volume might be interpreted as technical regress. Although our cross-section is much larger then the one used by Nghiem and Coelli (2002) we do observe some technical regress when including delivery volume.

Next, we turn to our input measures discussing total cost first. Here unlike most of the literature we use total cost as in, for instance, Jamasb and Pollitt (2003) and Edvardsen et al. (2006). The latter however use a Malmquist cost index including input quantities and prices. Also, Rouse and Swales (2006) describe how total expenditure is used as an input in DEA for the pricing of health services in New Zealand. In their case, output which is typically measured in number of discharges is cost-weighted to account for the difficulty of different treatments. This illustrates the desirable property of total cost as an input: the proper economic weighting of all inputs.

The choice of inputs changes the interpretation of our results as compared to standard measures of technical efficiency. Following Maniadakis and Thanassoulis (2004), our

---

[21] An alternative would be to estimate MPI using Stochastic Frontier Analysis where the frontier is based on the entire sample and year-to-year fluctuations affect the technical efficiency change component rather then the technical change component as explained by Coelli et al. (2005, p. 306).

measure might be referred to as "cost technical efficiency" as it implicitly includes allocative efficiency. As we do not have unit prices we cannot distinguish between technical and allocative efficiency as done by Maniadakis and Thanassoulis (2004) who constructed a new Malmquist index that allows for the inclusion of prices.[22] Though one might expect similar input prices across the US (except for labour), the International Energy Agency (1994, Fig. 1 Chapter 3) shows that for construction projects in 1990/91 costs differed for a given pipeline diameter. Sickles and Streitwieser (1991) calculate input prices from revenue and physical quantities. This has the problem that higher margins translate into higher input prices. And in particular with recent increases in rate flexibility it is likely that margins differ across firms.

The main advantage of a single monetary input measure from a regulator's point of view is that correct physical measures are difficult to obtain due to outsourcing, quality differences, or simple non-reporting. Also, as mentioned by Jamasb and Pollitt (2003), our input measure accounts implicitly for all possible trade-offs between the various inputs. Last, consumers (and hopefully regulators too) are interested not in technical efficiency as such but the cost of the service. In this light we also use an alternative input measure: total revenue.

First we should stress that although revenue is influenced by the regulatory regime, we do not use a bottom-up measure (like our Totex variable) as used by many regulators. Our motivation for using revenue as an alternative input measure is twofold. First, total cost, like physical inputs, might be difficult to measure and thus revenue might serve as a proxy especially in countries where regulatory accounting procedures are not as well-established as in the US. Second, in regulatory practice throughout Europe the rate-of-return is set in lengthy procedures reminiscent of US rate-cases. And, like in the US, returns often seem to be set rather arbitrarily.[23] To some extent this practice defeats the

---

[22] Traditionally prices could not be included in Malmquist indices and one would have to resort to parametric techniques to account for prices in productivity measurement, as done for instance by Farsi and Filippini (2004).

[23] Joskow (1972) illustrates how rates are set in the US. He observes that (not unlike in Europe) there are complex rules on how the rate base is set but little formal guidance is given for actual rates. Also unlike the other items that make up total cost, cost of capital is unobserved.

purpose of incentive regulation by introducing a cost-plus element. Also, the use of revenue might address the critique of European-style incentive regulation by Shuttleworth (1999) who argues that regulators use an average rate of return to reward efficient performance. Thus, investors have no incentive to invest in companies that are regulated in this way. Benchmarking revenue (instead of setting a WACC ex-ante) would allow for the trade-off between high efficiency and high returns to be reflected in actual frontier performance. A practical problem is obviously how to arrive at a "benchmarkable" revenue in the first place.

Last, we do not include non-discretionary variables in our analysis but suggest considering the following issues. First, as mentioned above, the way the systems cope with peaks might differ and might not be at the discretion of management. Another important exogenous variable is the age of the network, particularly because we rely on historic book values as a measure of capital expense. Though age might affect our *static* efficiency measures it should have a lesser impact on our measures of *change* as these measures are relative to a particular point in the past only. Possible other non-discretionary variables might be the end-use of deliveries (heating vs. industrial) and layout (trunk-line vs. radial grid). Table 3 summarizes the variables and methods used by several studies on productivity and efficiency of the US gas transmission industry.

Table 3: Summary of the Literature

Author	Data	Inputs	Outputs	Method
Callen (1978)	28 US inter-state gas transmission companies in 1965	horsepower weight of pipeline steel	delivery volume	Econometric production function
Aivazian (1987)	14 US inter-state gas transmission companies in 1953-1979	horsepower weight of pipeline steel compressor fuel labour	delivery volume multiplied by length of delivery	Econometric production function
Sickles and Streitwieser (1991)	14 US inter-state pipeline companies in 1977-1985	horsepower weight of pipeline steel compressor fuel labour	delivery volume multiplied by length of delivery	DEA, SFA
Ellig (1993)	50 Texan gas transmission companies, 1989	sales (commercial, industrial, resale) third-party delivery volume total throughput length of pipes gas purchasing cost	O&M expense	Econometric cost function
Granderson (2000)	20 US inter-state pipeline companies in 1977-1987	horsepower weight of pipeline steel compressor fuel labour	compressor fuel	SFA

4 Data and Measurement

The data used in this study is taken from the Federal Energy Regulatory Commission (FERC) which requires all inter-state transmission companies above a certain size to file a yearly regulatory report containing both financial and operating data (FERC Form 2). As far as possible all data is confined to the transmission function.

Though the data is not explicitly gathered for efficiency and productivity measurement purposes, the large number of studies (previously referred to) that relies on the data testifies to its general adequacy. However, several missing values had to be estimated from adjacent periods as MPI does not tolerate any missing values. Also, some observations where the data do not appear to be correct were excluded and several obvious errors were corrected. The data was corrected for inflation during the sample period. All monetary values are in 2004 US dollars. Revenue was adjusted such that no company had a rate-of-return lower then 6% in any year. This adjustment was made to prevent frontier firms from being those with sub-normal rates of return. Although six percent is essentially an arbitrary choice, it is chosen to be slightly higher then US interest rates (i.e. the risk free opportunity cost). The adjustment was necessary for five observations. The detailed definitions of the various variables and their measurements are given in Table 4.

Note that we exclude the cost of fuel as most pipelines withhold fixed percentages of the gas actually delivered as compensation for compressor fuel usage.[24] Also, we use historic book value as a cost measure which may be open to criticism. But Edvardsen et al. (2006) note that historic book value is a reasonable measure as we analyze "efficiency improvement and not static individual scores".

Summary statistics are given in Table 5. It is evident that the size of the companies included varies greatly. In terms of pipeline length the largest company is about sixty times larger then the smallest company. This reflects the fact that the nature of the companies differs. Whereas some connect several other pipelines in a particular region to benefit from arbitrage, others deliver gas over long distances from the main production regions in Canada and the Gulf of Mexico. For this reason the two largest hub operators were excluded as their delivery to cost ratios are by far the largest.

---

[24] This assumption is based on private communication with two companies whose data we include.

Table 4: Variable Description

Variable Name	Description	Measurement
Totex	O&M (less fuel, including labour) + Deprecation + Cost of Capital (written-down value multiplied by 6 percent)	2004 $
Revenue	Revenue from transportation of gas of others through transmission pipes.	2004 $
Delivery	Yearly total of gas transmitted for others (excluding losses).	Dth (decatherm)[25]
Mains	Total length of pipes (mains)	Miles
Horsepower (HP)	Total horsepower rating at compressor stations	HP
Age	Accumulated deprecation at mid-year / Annual depreciation	Years
Load factor	Delivery/Capacity (max. past single-day peak*365)	%
Rate of return (ROR)	(Revenue – O&M – Dep.)/Average written-down value	%

Also note that the last three variables age, load factor, and ROR (before adjustment) are not included in our analysis but help to describe the sample. For instance the average age is 27 years which is three times our sample length.[26] As discussed below this discrepancy is likely to weaken some of our results.

We use observations for 39 companies per year. Button and Weyman-Jones (1992) state that for DEA an approximate "degree of freedom" is the number of observations less the number of variables in the model. As we have four variables (one input and three outputs) our degrees of freedom are within the suggested limit of "about 35".

---

[25] 1 therm is equal to 100000 British thermal units (BTU).

[26] The unrealistic measure for maximum age should be due to measurement error or non-linear depreciation practice. Only two observations for age are above 83 years. The outliers are characterized by above average values for accumulated depreciation and below average values for annual depreciation.

Table 5: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max
Obs.: 351, Years: 1996-2004, Firms: 39				
Totex (m$.)	137	112	7.88	540
Revenue (m$.)	263	223	14.2	1100
Delivery (Dth.)	715	589	59	2840
Mains (miles)	4,645	4,117	269	16,666
Horsepower (HP)	395,553	399,938	5,200.00	1,600,000
Age (years)	27	31	4	508
Load factor (%)	0.67	0.19	0.25	1.15
ROR (%)	26	12	4	98

Figure 2 gives the changes in the yearly sample totals for the variables that are included in the calculation of the MPI. We observe that delivery volume fluctuates on a yearly basis, whereas total length of pipelines stays virtually constant and total horsepower is continuously increasing.

It is interesting that while capacity is added, total cost and revenue are falling (though they increase slightly towards the end of the sample period). This might be explained by pipelines expecting demand to pick up, pipelines taking advantage of arbitrage opportunities, or falling returns or other costs.

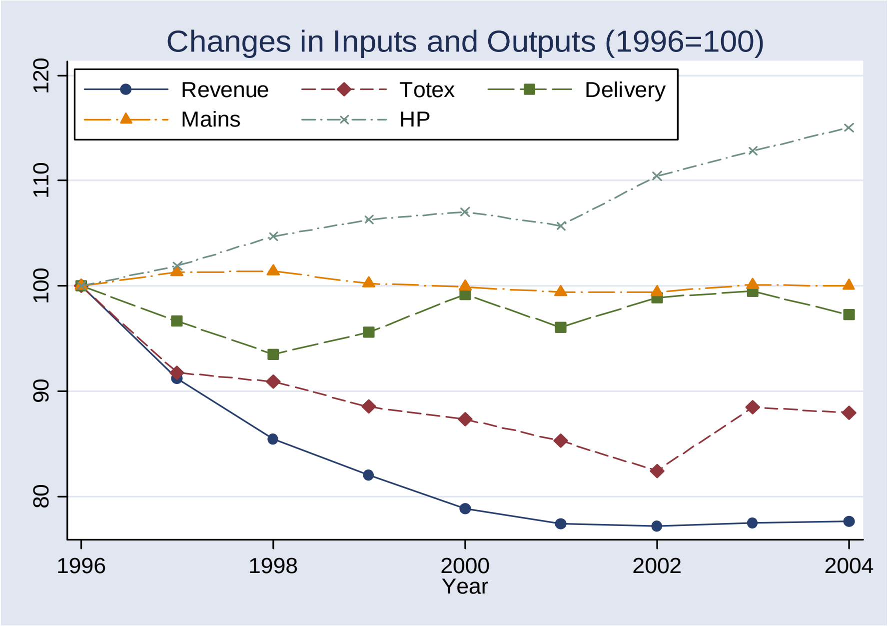


Figure 2: Changes in Inputs and Outputs (1996=100)

5 Results

Table 6 gives the regression results of the cost driver analysis. As already discussed above, there are reasons to believe that total cost does not vary with delivery once we account for capacity. Our results below seem to confirm this. However, the coefficient for delivery is larger (and the coefficient for mains smaller) in the revenue model, possibly reflecting that tariffs are not entirely cost-reflective in their decomposition.

We started with a full translog specification but dropped the interaction terms which are not individually significant (possibly due to multi-collinearity). Also, the model reported is the same as the full translog according to a likelihood ratio test._However, even though the two coefficients for delivery are individually insignificant, a likelihood ratio test does not suggest that we can drop them. Thus, the empirical evidence on delivery volume as a cost driver is not conclusive. Also, we added a time trend (year) whose coefficient confirms that costs (and revenue) are falling over the sample period. Last, it is important to stress that these particular results are likely to be influenced by the actual tariff regime in place (especially for revenue).

Table 6: Cost Driver Test Results

	Model 1	**Model 2**
Dependent variable	ln(Totex)	ln(Revenue)
ln(Delivery)	0.267	0.482
	(0.810)	(0.642)
ln(HP)	-0.529**	-0.508*
	(0.201)	(0.248)
ln(Mains)	1.558***	0.842*
	(0.343)	(0.342)
ln(Delivery)^2	0.003	0.000
	(0.020)	(0.016)
ln(HP)^2	0.032***	0.031**
	(0.009)	(0.011)
ln(Mains)^2	-0.086***	-0.042
	(0.022)	(0.022)
Year	-0.016***	-0.025***
	(0.003)	(0.003)
CONSTANT	39.607***	56.250***
	(10.250)	(8.105)
LL	327.72	279.64
AIC	-639.43	-543.29
obs.	351	351
* p<0.05, ** p<0.01, *** p<0.001		

Now we turn to our main results: the Malmquist productivity indices. Table 7 and Table 8 show the cumulative (and averaged across firms) Malmquist indices (TFPC) and their decomposition into technical efficiency change (TEC) and technical change (TC). Additionally, the yearly technical efficiency scores (TE) which are inputs to the Malmquist indices are reported. The last row in both tables gives the implied average yearly growth rates by taking the index for the last year (2004) and dividing it by the number of years that elapsed since the base year. Table 7 gives the results for the two Totex models.

Table 7: Average Malmquist Indices and their Decomposition

	Model 1 (Totex, incl. delivery)				**Model 3 (Totex, excl. delivery)**			
Year	TE	TFPC	TEC	TC	TE	TFPC	TEC	TC
1996	0.52	1	1	1	0.45	1	1	1
1997	0.58	1.07	1.13	0.95	0.49	1.14	1.12	1.02
1998	0.61	1.09	1.22	0.91	0.50	1.21	1.18	1.03
1999	0.65	1.12	1.33	0.86	0.50	1.22	1.18	1.05
2000	0.63	1.17	1.30	0.91	0.49	1.25	1.20	1.06
2001	0.61	1.17	1.24	0.95	0.51	1.30	1.26	1.05
2002	0.64	1.25	1.33	0.96	0.52	1.44	1.32	1.11
2003	0.63	1.20	1.32	0.91	0.54	1.42	1.40	1.02
2004	0.64	1.23	1.28	0.96	0.53	1.47	1.41	1.06
growth rate p.a. (%)	-	2.9	3.5	-0.5	-	5.9	5	0.8

First, we observe that the MPI and its components have larger values when delivery is excluded. This is not surprising as the length of mains and horsepower (unlike delivery) are virtually non-decreasing. The technical regress for Model 1 might be caused by the fluctuations in delivery as explained above. However, for the static technical efficiency scores the values for Model 1 and Model 2 are higher as the additional variable allows more firms to be relatively efficient.

Table 8 gives the same results for our two revenue models. Here TFP growth is stronger because revenue falls more quickly then Totex as shown in Figure 2 above. Also, for three of the four models TEC dominates TC. Generally, absolute numbers vary greatly across our four models.

Table 8: Average Malmquist Indices and their Decomposition

	Model 2 (Revenue, incl. delivery)				**Model 4 (Revenue, excl. delivery)**			
Year	TE	TFPC	TEC	TC	TE	TFPC	TEC	TC
1996	0.53	1	1	1	0.44	1	1	1
1997	0.54	1.09	1.04	1.05	0.48	1.18	1.11	1.06
1998	0.59	1.18	1.17	1.01	0.55	1.27	1.34	0.95
1999	0.60	1.20	1.20	1.00	0.54	1.29	1.32	0.98
2000	0.65	1.26	1.32	0.95	0.57	1.33	1.38	0.96
2001	0.60	1.26	1.20	1.05	0.54	1.37	1.32	1.06
2002	0.61	1.31	1.25	1.05	0.55	1.46	1.38	1.07
2003	0.57	1.35	1.16	1.17	0.54	1.52	1.36	1.12
2004	0.56	1.36	1.13	1.21	0.53	1.55	1.35	1.16
growth rate p.a. (%)	-	4.5	1.6	2.6	-	6.9	4.3	2

Looking at the implied *yearly* growth rates, Model 1 for instance would produce an average yearly productivity increase of 2.9 percent. These growth rates are higher then in earlier periods as reported by Sickles and Streitwieser (1991) and Granderson (2000). Also, as an example for a US-EU comparison this might be contrasted with the result of a recent report by the German network regulator (Bundesnetzagentur, 2006) which found an average yearly TFP growth of 2.19 percent for the entire German energy industry for the years 1977-1997 (using a Törnquist index). However, the different methodologies used by these authors and the different market environment at the time make comparisons difficult. Appendix B gives the yearly DEA efficiency scores by firm for Model 1.

Next, the same numbers are presented graphically. When looking at the results across time there appear to be more similarities across models. Figure 3 shows Malmquist indices and their decomposition into technical efficiency change and technical change for the Totex models. Whereas the upper panel includes delivery, the lower panel excludes delivery as an output.

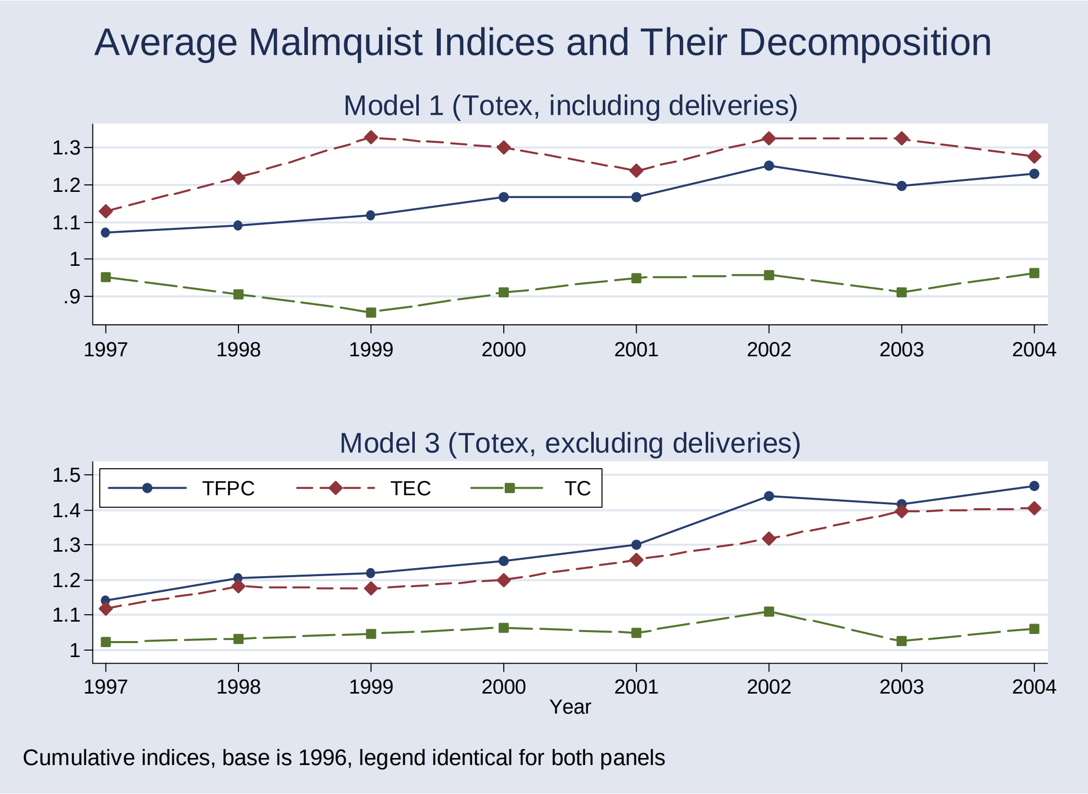


Figure 3: Average Malmquist Indices and their Decomposition

Figure 4 shows a similar pattern for the revenue models. In the second half of the sample period, technical efficiency change is falling whereas technical change is much stronger than for the Totex models. While we do not have a good explanation for this, we note that a merger wave in gas distribution and transmission occurred around the year 2000 as shown by Moss (2005) which may have increased market power. Increased market power would also explain the discrepancy between the path of the Totex- and Revenue-based technical efficiency changes.[27]

---

[27] In the US, like in Europe, mergers are decided by the antitrust authorities and not the regulators. In the US the former use a less stringent "no harm" benchmark as discussed by Balto and Mongoven (2001).

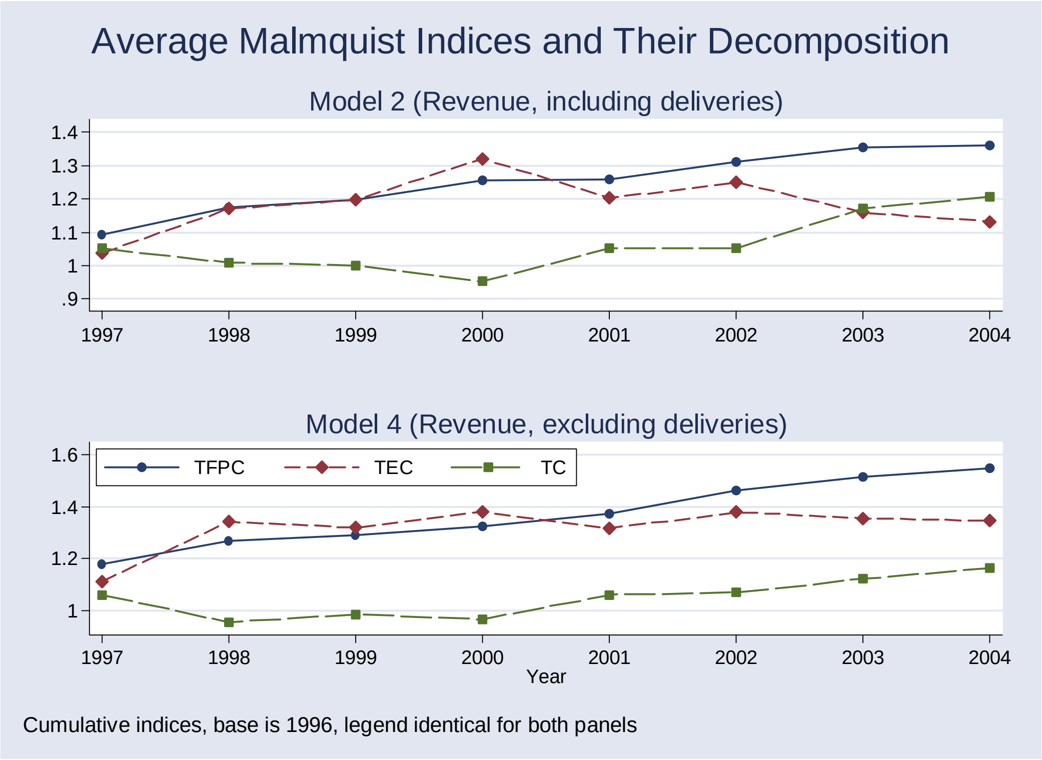


Figure 4: Average Malmquist Indices and Their Decomposition

Table 9 and Table 10 give the Pearson correlation coefficients for the results from the various models. In particular, we are interested in the correlation between the Revenue and Totex models where the results are highlighted along the diagonals in the lower left of the two tables.

When focusing on these diagonals, two observations can be made. First, the correlation is higher for the level TE scores as compared to the various change measures. Second, the correlations are higher for the models excluding delivery in Table 10. This may be due to the presence of volume-related charges affecting revenue more than total cost. Thus, when delivery is excluded the remaining explanatory variables (horsepower and length of mains) have the same relative effect on Totex and Revenue leading to higher correlations.

Table 9: Pearson Correlation Coefficients, Models including Delivery

		Model 1 (Totex)				Model 2 (Revenue)			
		TE	TFPC	TEC	TC	TE	TFPC	TEC	TC
Model 1	TE	1.000							
	TFPC	-0.134*	1.000						
	TEC	-0.166*	0.729*	1.000					
	TC	0.003	0.403*	-0.310*	1.00				
Model 2	TE	0.849*	-0.130*	-0.082	-0.090	1.000			
	TFPC	-0.241*	0.589*	0.414*	0.376*	-0.037	1.000		
	TEC	-0.301*	0.395*	0.428*	-0.011	0.049	0.858*	1.000	
	TC	0.031	0.293*	0.013	0.546*	-0.166*	0.329*	-0.180*	1.000

*indicates significance at 5%

Table 10: Pearson Correlation Coefficients, Models excluding Delivery

		Model 3 (Totex)				Model 4 (Revenue)			
		TE	TFPC	TEC	TC	TE	TFPC	TEC	TC
Model 3	TE	1.000							
	TFPC	-0.215*	1.000						
	TEC	-0.267*	0.886*	1.000					
	TC	0.071	0.262*	-0.194*	1.000				
Model 4	TE	0.876*	-0.167*	-0.199*	0.071	1.000			
	TFPC	-0.230*	0.804*	0.682*	0.289*	-0.091	1.000		
	TEC	-0.278*	0.683*	0.748*	-0.103	-0.084	0.890*	1.000	
	TC	0.069	0.238*	-0.143*	0.848*	0.005	0.280*	-0.169*	1.000

*indicates significance at 5%

Next, we turn to the results for the convergence test and look at β-convergence first. Figure 5 plots the DEA technical efficiency score in the base year against the technical efficiency change component of the MPI (for Model 1). Second, to examine the effect of the sample period length on convergence, the MPI calculations were repeated for all possible sample lengths moving the base year up by one each time. The different runs are represented by the differently shaped markers and lines. For each value on the x-axis there are several values along the y-axis. These are values for a given firm across the years. The fitted lines show that the rate of technical efficiency *change* tends to be higher the lower the *level* of technical efficiency in the base year. Also, this negative relationship weakens as the sample period becomes shorter. Although we provide no formal test, the relationship between the length of the sample period and the β-convergence potentially has implications for the length of the regulatory period.

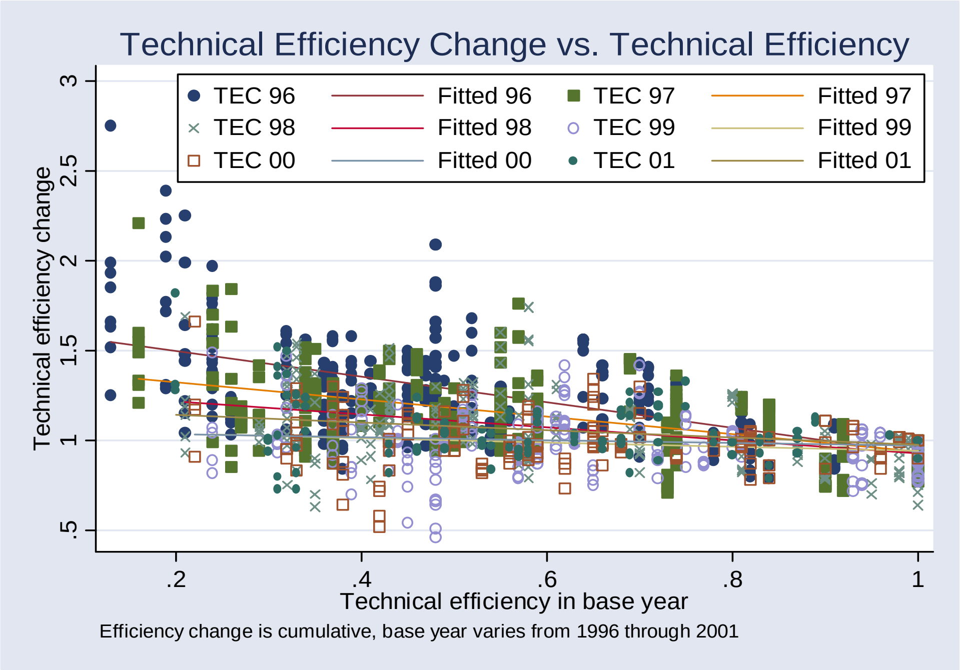


Figure 5: Technical Efficiency Change vs. Technical Efficiency

Next, we present more formal evidence on β-convergence. Following Alam and Sickles (2000), we regress the average year to year growth in the technical efficiency scores on the logarithm of the technical efficiency scores in the base year (i.e. 1996). Table 11 presents the results. The negative coefficient confirms that there is convergence in the efficiency scores. However, the slope coefficient for the revenue model is not significant at a 5% level. Thus, both tests indicate that there is β-convergence.

Next, we turn to our results for σ-convergence. Figure 6 gives the box plots for the static efficiency scores for our models. We observe that for all models the variance decreases slightly over the sample period. Again, this fall in variance is more pronounced for the models excluding delivery. Thus, we conclude that there is evidence for both β-convergence and σ-convergence. However, given the very low minimum values for the static efficiency scores, the decrease in the variance does not seem particularly strong.

Table 11: Results of β-Convergence Test

	Model 1	**Model 2**
Dependent variable	Avg. TE growth for Totex	Avg. TE growth for Revenue
ln(TE96)	-0.019*	-0.019
	(0.009)	(0.011)
Constant	0.003	-0.014
	(0.008)	(0.010)
Prob>F	4.55	3.21
R-squared?	0.09	0.05
No. of obs.	39	39
* p<0.05		

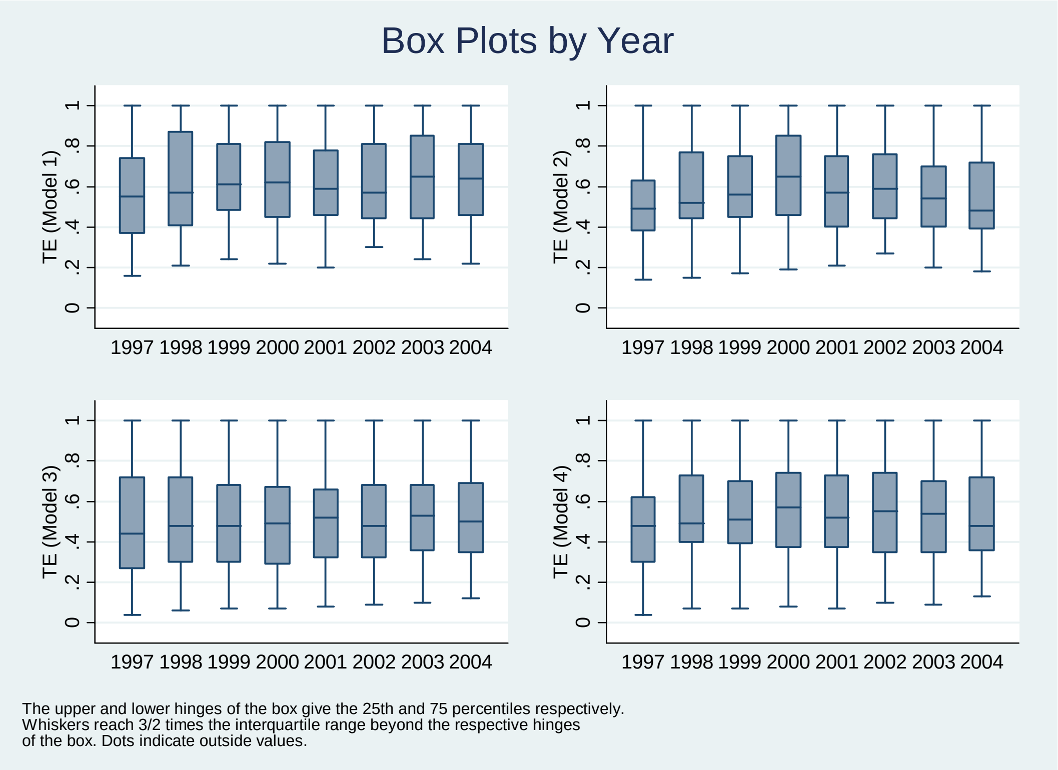


Figure 6: Box Plots by Year

6 Discussion and Conclusion

6.1 Discussion of Results

Since the liberalization of the well-head price for natural gas in the US and subsequent regulatory change for the pipeline business, several studies have measured firm-level efficiency and productivity and in particular the effect of regulatory change. These studies have found that efficiency declined and convergence did not occur in the first decade after well-head price deregulation. In this study, we focus on more recent years and find modest improvements in technical efficiency (and stronger improvements in overall productivity). Moreover, we find evidence of convergence over time.

As for our models, we opted for a single monetary input measure (total cost or revenue) which is not a common approach in the literature but offers several advantages: only standard accounting data is required; trade-offs between the various inputs are accounted for; and monetary measures have the advantage that they account for outsourcing and quality differences in inputs. Also, our use of revenue as an alternative input has two desirable features from a regulatory perspective: (1) revenue is the total cost to consumers; and (2) aggregate revenue measures are readily available. As we use a monetary input measure but do not include prices (and thus do not distinguish between technical and allocative efficiency), our efficiency measures have incentive properties different from standard technical efficiency measures. Our measure would have the same incentive properties as a standard technical efficiency measure if firms were allocatively efficient and faced the same input prices. The results for Totex and Revenue models are similar but there are differences. Unlike cost, revenue is more likely to be driven by the particular tariff regime, market power, etc.

On the output side, we examined models which exclude gas delivery volumes because the literature on costs and our understanding of tariff setting indicate that both are largely driven by capacity rather then volume. Our efforts to confirm this empirically are not conclusive and the productivity results differ depending on whether delivery is included as an output or not.

We report both DEA-based technical efficiency scores and Malmquist productivity indices. Though most regulators rely on cross-section based static scores,[28] it might be interesting to contrast these with observed productivity changes. According to our results, TFP change and TEC seem higher than one would expect for a rate-of-return regulated, natural monopoly industry. For our two Totex models we observe yearly average growth rates of 2.9 and 5.9 percent. For our revenue models it is 4.5 and 6.9 percent respectively. For all models except Model 2, TEC dominates TC. For all models, TFPC is upward-trending over the entire sample period. For all models except Model 3, TEC is flattening or declining from the year 2000 onwards. Though we have no good explanation for this, it may be related to merger activity as discussed above.

Next, we observe some convergence in relative performance. Firms starting at a lower efficiency level grow faster (β-convergence) and the dispersion of static efficiency scores (σ-convergence) declines over time. Again, under rate-of-return regulation we would expect only limited convergence as, unlike competition, it exerts no pressure to converge. In order to come to stronger conclusions on convergence, a sample length equivalent to the investment cycle would be necessary. Also, our results show averages and are therefore likely to gloss over regional differences. We know that pipeline competition is foremost a regional phenomenon and would expect the same to be the case for efficiency trends and convergence.

Thus, our results show that regulatory change in the US is followed by "cost productivity" and "revenue productivity" improvements. What changed is not so much the rate-of-return regulation but the building of competitive markets and increased tariff flexibility (which can be obtained even under an unchanged rate-of-return regulation). Encouraging competition through creation of necessary institutions might be more important in the long-run than the prevailing form of tariff regulation. Also, increased competition might explain why we observe that in a mature industry with a long history of rate-of-return regulation, technical efficiency change dominates technical change.

---

[28] It is interesting that Ofwat the UK water regulator performs its benchmark on a cross-section even though it has a panel at its disposal (Weeks and Lay, 2006). At a presentation of their draft the authors commented that a possible reason for the continued use of a cross-section is that management does not consider itself responsible for the performance of past management teams.

6.2 Lessons for Europe

Many European countries recently embarked on a route of regulatory change heading towards incentive regulation. It might be worthwhile for European regulators to consider the insights that US experience and data offer.

First, our work points towards issues for data collection. Though FERC data collection is driven by the needs of elaborate rate-cases, its overall requirements on transparency and rigour in data collection are an important point of reference. However, FERC recognizes that a move away from rate-of-return regulation shifts the emphasis from quantity to quality for data collection (O'Neill et al., 1996).

Broadly speaking, our analysis points towards a short-run and a long-run lesson for European regulators. In the short-run FERC data provides the opportunity for individual European regulators to benchmark national gas transmission companies without there being a standardized European data set. In the long-run European regulators should consider giving more emphasis to market integration and competition since these arguably lead to productivity increase and convergence, as in the US. As for benchmarking individual European companies, we remark the following: Although we do not address comparability in detail we believe (and other regulators have shown[29]) that a sufficient degree of comparability can be obtained with FERC data.

We would like to make two general points. First, once European regulators begin to collaborate on gathering data in a systematic and comparable way there will be enough data to produce robust results from European data alone. However in the meantime, comparing European companies to US companies might provide at least some guidance for regulators that often face difficult-to-verify claims from industry. An added advantage of using US data is that a panel is available that allows for more robust conclusions on performance changes since single cross-sections are likely to be affected by measurement error.

---

[29] In particular, regulators in New Zealand and Australia have been keen users of US data to benchmark their regulated companies. See, for instance, IPART (1999), Pacific Economics Group (2004), and Carrington (2002).

In the long-run, even if sufficient European data were available, international benchmarks still have an important role to play. It is possible that US companies embody best international practice. Also, there is no reason to believe that firms under incentive regulation should fare worse than under rate-of-return regulation (complemented by competition or not). To exclude US management performance from any European benchmark could amount to forfeiting consumer surplus.

References

Aivazian, V. A. and J. L. Callen (1981). Capacity Expansion in the U.S. Natural-Gas Pipeline Industry. Productivity Measurement in Regulated Industries. T. G. Cowing and R. E. Stevenson, Academic Press.

Aivazian, V. A., Callen, J.L., Chan, M. W. L., Mountain, D. C. (1987). "Economies of Scale Versus Technological Change in the Natural Gas Transmission Industry." The Review of Economics and Statistics **69**: 55-561.

Alam, I. M. S. and R. C. Sickles (2000). "Time Series Analysis of Deregulatory Dynamics and Technical Efficiency: The Case of the U.S. Airline Industry." International Economic Review **41**(1): 203-218.

Alger, D. and M. Toman (1990). "Market-Based Regulation of Natural Gas Pipelines." Journal of Regulatory Economics **2**(3): 263-280.

Balto, D. and J. Mongoven (2001). "Deregulation and Merger Enforcement in the Natural Gas Industry." Antitrust Law Journal **69**(527-568).

Beck, N. and J. N. Katz (1995). "What to do (and not to do) with Time-Series Cross-Section Data." The American Political Science Review **89**(3): 634-647.

Bundesnetzagentur (2006). 2. Referenzbericht Anreizregulierung – Generelle sektorale Produktivitätsentwicklung im Rahmen der Anreizregulierung. Bonn, Bundesnetzagentur.

Button, K. J. and T. G. Weyman-Jones (1992). "Ownership structure, institutional organization and measured X-efficiency." American Economic Review **82**(2): 439-445.

Callen, J. L. (1978). "Production, Efficiency, and Welfare in the Natural Gas Transmission Industry." American Economic Review **68**: 311-323.

Carrington, R., Coelli, T., Groom, E. (2002). "International Benchmarking for Monopoly Price Regulation: The Case of Australian Gas Distribution." Journal of Regulatory

Economics: 191-216.

Caves, D. W., L. R. Christensen and W. E. Diewert (1982). "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity." Econometrica **50**(6): 1393-1414.

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

Coelli, T. J., D. S. P. Rao, C. J. O'Donnell and G. E. Battese (2005). An Introduction to Efficiency and Productivity Analysis, Springer.

Doane, M. J., R. P. McAffee and M. A. Williams (2004). "Evaluating and Enhancing Competition in the Interstate Natural Gas Transportation Industry." Natural Resources Journal **44**(3): 761-808.

Edvardsen, A., D. Fjeld, F. R. Forsund, W. Hansen, S. A. C. Kittelsen and T. Neurauter (2006). Productivity and Regulatory Reform of Norwegian Electricity Distribution Utilities. Performance Measurement And Regulation Of Network Utilities. T. Coelli and D. Lawrence, Edward Elgar.

Ellig, J., Giberson, M. (1993). "Scale, Scope, and Regulation in the Texas Gas Transmission Industry." Journal of Regulatory Economics **5**: 79-90.

Energy Information Agency (2002). Expansion and Change on the U.S. Natural Gas Pipeline Network. Washington D.C., Energy Information Agency.

Energy Information Agency (2003). Natural Gas Market Centers and Hubs: A 2003 Update. Washington D.C., Energy Information Agency.

European Commission (2005). Technical Annexes to the Report from the Commission on the Implementation of the Gas and Electricity Internal Market. Brussels, European Commission.

European Commission (2007). DG Competition Report on Energy Sector Inquiry. Brussels, European Commission.

Färe, R., S. Grosskopf, B. Lindgren and P. Roos (1989). Productivity developments in Swedish hospitals: A Malmquist output index approach. Discussion Paper No. 89-3, Southern Illinois University.

Färe, R., S. Grosskopf and D. Margaritis (2006). "Productivity growth and convergence in the European Union." Journal of Productivity Analysis **25**: 111-141.

Farrell, M. J. (1957). "The Measurement of Productive Efficiency." Journal of the Royal Statistical Society. Series A (General) **120**(3): 253-281.

Farsi, M. and M. Filippini (2004). "Regulation and Measuring Cost-Efficiency with Panel Data Models: Application to Electricity Distribution Utilities." Review of Industrial Organization **25**(1): 1-19.

Granderson, G. (2000). "Regulation, Open-Access Transportation, and Productive Efficiency." Review of Industrial Organization **16**: 368-385.

Grosskopf, S. (1993). Efficiency and Productivity. The Measurement of Productive Efficiency. H. O. Fried, C. A. K. Lovell and S. S. Schmidt, Oxford University Press.

Hirschhausen, C. v. (2006). Infrastructure Investments and Resource Adequacy in the Restructured US Natural Gas Market – Is Supply Security at Risk? Working Paper, MIT Center for Energy and Environmental Policy Research.

Hirschhausen, C. v., A. Neumann and S. Rüster (2007). Competition in Natural Gas Transportation? Technical and Economic Fundamentals and an Application to Germany. Globalization of Natural Gas Markets Working Papers WP-GG-21b, DIW Berlin.

International Energy Agency (1994). Natural Gas Transportation: Organization and Regulation. Paris, International Energy Agency.

IPART (1999). Benchmarking the Efficiency of Australian Gas Distributors. Sydney, IPART.

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

Jamasb, T. and M. Pollitt (2003). "International benchmarking and regulation: an application to

European electricity distribution utilities." Energy Policy **31**(15): 1609-1622.

Johnson, S., J. Rasmussen and J. Tobin (1999). Corporate Realignment and Investments in the Interstate Natural Gas Transmission System, Energy Information Administration.

Joskow, P. L. (1972). "The Determination of the Allowed Rate of Return in a Formal Regulatory Hearing." The Bell Journal of Economics and Management Science **3**(2): 632-644.

Kahn, A. (1971). The Economics of Regulation: Principles and Institutions. New York, Wiley & Sons. **II:** 153-154.

Kalt, J. P. and F. C. Schuller (1987). Drawing the Line on Natural Gas Regulation. New York, Quorum Books.

Makholm, J. D. (2007). Seeking Competition and Supply Security in Natural Gas: The US Experience and European Challenge. Boston, National Economic Research Associates.

Maniadakis, N. and E. Thanassoulis (2004). "A cost Malmquist productivity index." European Journal of Operational Research **154**: 396-409.

Moss, D. L. (2005). Natural Gas Pipelines: Can Merger Enforcement Preserve the Gains from Restructuring? Working Paper No. 05-08, American Antitrust Institute.

Nghiem, H. S. and T. Coelli (2002). "The Effect of Incentive Reforms Upon Productivity: Evidence from the Vietnamese Rice Industry." Journal of Development Studies **39**(1): 74-93.

O'Neill, R. P. (2005). Natural Gas Pipelines. Network Access, Regulation and Antitrust. D. L. Moss, Routledge**:** 107-120.

O'Neill, R. P., C. S. Whitmore and M. Veloso (1996). The Governance of Energy Displacement Network Oligopolies. Washington, D.C., Office of Economic Policy, Federal Energy Regulatory Commission.

Olsen, W. P. (2005). "Secrecy and Utility Regulation." The Electricity Journal **18**(4): 48-52.

Pacific Economics Group (2004). New Zealand Gas Distribution Cost Performance: Results from International Benchmarking, commissioned by Vector Networks.

Rouse, P. and R. Swales (2006). "Pricing public health care services using DEA: Methodology versus politics." Annals of Operations Research **145**(1): 265-280.

Schmidt, M. R. (2000). Performance-based Ratemaking: Theory and Practice. Vienna, Virginia, Public Utilities Reports, Inc.

Shleifer, A. (1985). "A theory of yardstick competition." Rand Journal of Economics **16**(3): 319-327.

Shuttleworth, G. (1999). Regulatory Benchmarking: A Way Forward Or A Dead-End? Energy Regulation Brief, NERA, London.

Sickles, R. C. and M. L. Streitwieser (1991). Technical inefficiency and productive decline in the U.S. interstate natural gas industry under the natural gas policy act. Discussion Paper No. 91-6, Center for Economic Studies.

Sickles, R. C. and M. L. Streitwieser (1998). "An Analysis of Technology, Productivity, and Regulatory Distortion in the Interstate Natural Gas Transmission Industry: 1977-1985." Journal of Applied Econometrics **13**: 377-395.

Weeks, M. and H. Lay (2006). Efficiency Measurement in the Privatised English and Welsh Water and Sewerage Industry 1992-2004. 13th International Conference on Panel Data, Cambridge, UK.

Appendices

Appendix A: DEA and Malmquist Techniques

One way to account for changes in productivity is to combine single and mixed-period distance functions into an index as pioneered by Caves et al. (1982) and Färe et al. (1989). Next, we present the methodology formally following Grosskopf (1993).

At each time period $t = 1,..,T$ there are $k = 1,\dots,K$ firms (i.e. decision units) that use a single input $x^{k,t} = (X_k)$ to produce n outputs $y^{k,t} = (Y_{1k},\dots,Y_{nk})$. For each time period a production technology is constructed using DEA following Farrell (1957) and Charnes et al. (1978). For a given period t the constant returns to scale (CRS) frontier technology is given by:

$$S^t_{CRS} = \left\{ \begin{array}{l} \left(x^t, y^t\right) : \sum_{k=1}^{K} z^{k,t} y_n^{k,t} \geq y_n^t, \sum_{k=1}^{K} z^{k,t} x^{k,t} \leq x^t, \\ n = 1,\dots N, z_k \geq 0, k = 1,\dots K \end{array} \right\}$$

where the upper boundary of this set represents the best practice frontier. Relative to this frontier technology S^t one may define an input distance function for company k:

$$D_I^t\left(x^{k',t}, y^{k',t}\right) = \left\{\theta : \left(\theta x^{k',t}, y^{k',t}\right) \in S^t\right\}$$

Following Färe et al. (1989) and given two time periods (and thus two technologies), four input distance functions can be calculated. Two of these functions evaluate a period's observations against its respective reference technology and two evaluate its observations against the technology of the other period. The MPI is the geometric mean of these four distance functions

$$M_I\left(k',t,t+1\right)=\left[\frac{D_I^t\left(x^{k',t+1},y^{k',t+1}\right)}{D_I^t\left(x^{k',t},y^{k',t}\right)}\frac{D_I^{t+1}\left(x^{k',t+1},y^{k',t+1}\right)}{D_I^{t+1}\left(x^{k',t},y^{k',t}\right)}\right]^{1/2}$$

As mentioned above, an important feature of the Färe et al. (1989) version of the Malmquist index is that it can be decomposed, namely into

$$Technical\ Efficiency\ Change\left(TEC\right)=\frac{D_I^{t+1}\left(x^{k',t+1},y^{k',t+1}\right)}{D_I^t\left(x^{k',t},y^{k',t}\right)}$$

and

$$Technical\ Change\left(TC\right)=\left[\frac{D_I^t\left(x^{k',t+1},y^{k',t+1}\right)}{D_I^{t+1}\left(x^{k',t+1},y^{k',t+1}\right)}\frac{D_I^t\left(x^{k',t},y^{k',t}\right)}{D_I^{t+1}\left(x^{k',t},y^{k',t}\right)}\right]^{1/2}$$

and thus

$$M_I\left(k',t,t+1\right)=TFPC=TEC*TEC$$

Noting that the input distance function is the reciprocal of the Farrell (1957) input-oriented measure of technical efficiency we calculate the distance function for period *t* as

$$
\begin{aligned}
&\left[D_I^t\left(x^{k',t}, y^{k',t} \mid CRS\right)\right]^{-1} = \min \theta \ s.t. \\
&\sum_{k=1}^{K} z^{k,t} y_n^{k,t} \geq y_n^t, \\
&\sum_{k=1}^{K} z^{k,t} x^{k,t} \leq \theta x^t, \\
&z^{k,t} \geq 0, \\
&n = 1, \ldots N; k = 1, \ldots K
\end{aligned}
$$

Further details and the equivalent formulae for the mixed-period distance functions are given in Grosskopf (1993). The Malmquist index and its decomposition are illustrated in Figure A1. The two lines from the origin give the technological frontiers in the two periods. For both periods their respective observations lie somewhat below the frontier.

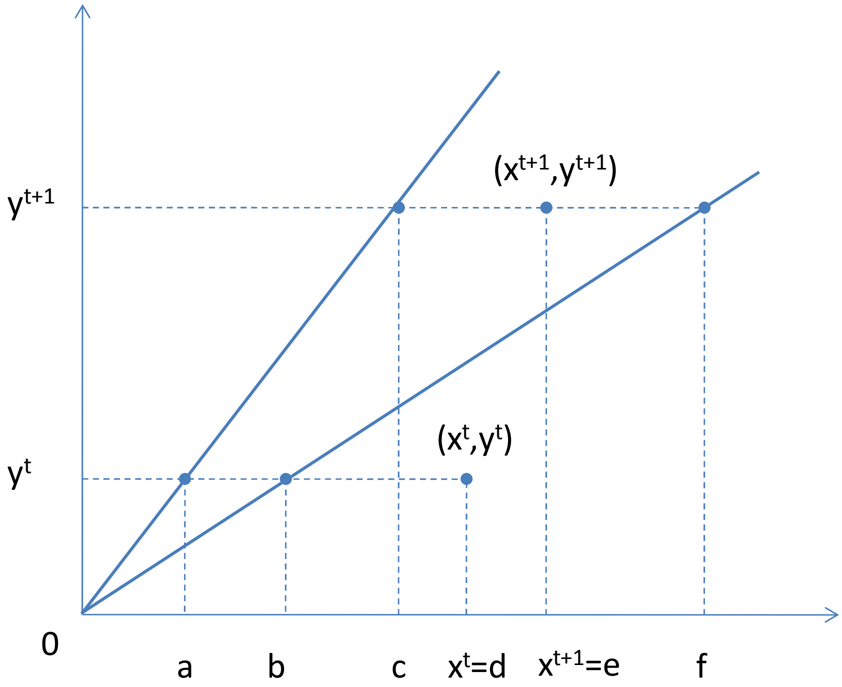


Figure A1: Illustration of a Malmquist Decomposition

Technical efficiency change is given as

$$TEC = \frac{0c}{0e}\frac{0b}{0d}$$

and technical change as

$$TC = \left[\frac{0c}{0f}\frac{0a}{0b}\right]^{1\backslash 2}$$

and hence total factor productivity change as

$$TFPC = \frac{0c}{0d}\left[\frac{0b}{0e}\frac{0a}{0f}\right]^{1\backslash 2}$$

Appendix B: Individual Efficiency Scores

Table B1: Firm-level DEA efficiency scores by year

No./Year	1997	1998	1999	2000	2001	2002	2003	2004
1	0.55	0.55	0.62	0.65	0.69	0.78	0.87	0.83
2	0.29	0.33	0.39	0.33	0.33	0.42	0.32	0.27
3	0.59	0.61	0.62	0.65	0.68	0.69	0.80	0.78
4	0.49	0.52	0.59	0.56	0.53	0.51	0.56	0.53
5	0.34	0.33	0.40	0.45	0.49	0.51	0.51	0.46
6	0.84	0.98	1.00	0.93	0.89	0.98	1.00	1.00
7	0.34	0.32	0.32	0.38	0.31	0.44	0.43	0.48
8	0.90	0.81	0.77	0.74	0.68	0.67	0.66	0.68
9	0.35	0.42	0.44	0.53	0.46	0.44	0.44	0.43
10	0.27	0.29	0.31	0.32	0.30	0.30	0.29	0.30
11	0.37	0.41	0.48	0.37	0.32	0.40	0.44	0.48
12	1.00	1.00	1.00	0.96	0.87	0.91	0.88	0.81
13	0.24	0.32	0.45	0.38	0.33	0.40	0.42	0.24
14	0.16	0.21	0.24	0.22	0.20	0.36	0.25	0.26
15	1.00	0.87	0.80	0.82	0.83	0.77	0.82	0.84
16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.90
17	0.81	0.99	0.94	0.98	0.97	1.00	1.00	1.00
18	0.26	0.35	0.48	0.42	0.31	0.31	0.24	0.22
19	0.40	0.43	0.49	0.50	0.49	0.47	0.51	0.47
20	0.69	0.80	0.96	0.99	1.00	1.00	1.00	1.00
21	0.37	0.32	0.37	0.33	0.34	0.36	0.41	0.42
22	0.92	0.95	0.93	0.84	0.84	0.85	0.66	0.72
23	0.37	0.40	0.43	0.43	0.43	0.42	0.40	0.36
24	0.42	0.58	0.57	0.49	0.46	0.51	0.52	0.57
25	0.73	0.69	0.81	0.78	0.80	0.79	0.74	0.79
26	0.48	0.48	0.57	0.62	0.54	0.56	0.52	0.45
27	0.55	0.51	0.59	0.68	0.66	0.63	0.65	0.66
28	0.51	0.51	0.51	0.59	0.58	0.57	0.55	0.54
29	0.55	0.63	0.63	0.65	0.63	0.62	0.65	0.64
30	0.74	0.90	0.94	0.70	0.72	0.71	0.91	0.81
31	0.73	0.58	0.59	0.58	0.57	0.53	0.54	0.52
32	0.43	0.43	0.52	0.51	0.56	0.54	0.63	0.65
33	0.46	0.52	0.61	0.58	0.59	0.64	0.68	0.67
34	0.57	0.58	0.70	0.90	0.75	0.89	1.00	1.00
35	0.50	0.57	0.65	0.56	0.49	0.51	0.53	0.54
36	0.59	0.70	0.72	0.66	0.69	0.57	0.65	0.64
37	1.00	0.98	1.00	0.81	0.78	0.81	0.77	0.80
38	0.92	1.00	1.00	1.00	1.00	1.00	1.00	1.00
39	0.74	1.00	0.75	0.82	0.72	0.86	0.85	0.64