Productivity and Efficiency of US Gas Transmission Companies:

A European Regulatory Perspective

Tooraj Jamasb[a]

Michael Pollitt[b]

Thomas Triebs[c]

30 April 2008

---

[a] Faculty of Economics and ESRC Electricity Policy Research Group, University of Cambridge, Sidgwick Avenue, Cambridge CB3 9DE, UK.

[b] Judge Business School and ESRC Electricity Policy Research Group, University of Cambridge, Trumpington Street, Cambridge CB2 1AG, UK.

[c] Corresponding Author, Tel: +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.

The authors would like to thank the UK Economic and Social Research Council (ESRC) for supporting this study. Also, we would like to thank two anonymous reviewers.

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 productivity development of a panel of US interstate companies using Data Envelopment Analysis and Malmquist productivity indices. 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. However, we argue that a benchmarking-based regulation might have brought about stronger convergence. 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; total factor productivity

JEL classification: L51, L95, O57, D24

1 Introduction

Gas transmission networks are regulated rather differently in Europe and in the US. Both, traditional US cost-of-service or rate-of-return regulation and European incentive regulation are based on the notion of natural monopoly. However, while the 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), the European regulators treat gas transmission as incentive-regulated franchise monopolies (Makholm, 2007).

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 than 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.[1] A benchmarking approach, as explored in Jamasb et al. (2007)[2], can be 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.

---

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

[2] This report prepared by the authors for the Council of European Energy Regulators (CEER) benchmarks 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).

Makholm (2007) details 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 the productivity of US industry performance to assess the overall impact of various regulatory changes that took place both before and during our sample period. 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 several convergence tests. In doing so, we take a European regulatory perspective in the sense that we analyze US data with benchmarking techniques widely applied in Europe.

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

Thus, we ask the following two sets of questions. First, has average industry productivity increased and has firm-level technical efficiency converged?[3] We know

---

[3] We would like to stress that our analysis applies 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.

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 push for more competition in 1992 (FERC Order 636), we would expect to find increased efficiency and convergence in contrast to earlier periods. Also, even though we are not able to include a control group, 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 overall regulatory change has been in continuously improving performance. Second, what lessons can European regulators learn from the relative success (according to our judgment) 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 dialogue between regulators across the Atlantic.[4]

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; section 6 discusses the results and section 7 concludes.

2 Background

The challenge for any regulator is to increase efficiency and reduce price, as stated by the European Commission's second Gas Directive ("Acceleration Directive")[5]. Although the process of European gas market liberalization and integration commenced in the mid 1990's, the Commission acknowledges in its Acceleration Directive and a

---

[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.

[5] European Commission (2003).

recent Energy Sector Inquiry[6] that many obstacles remain. Also, in its Energy Sector Inquiry the Commission recognizes that the US gas market is more developed than the European markets, although it does not discuss the 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 wish to move towards incentive regulation for the (unbundled) pipeline part of the value-chain, the US regulator aims at promoting 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 to a large extent both economically and physically integrated.

According to Kalt and Schuller (1987), already in 1987 about one-third of city gate markets received services from multiple pipelines.[7] Doane et al. (2004) argue that regulatory change led to an integrated US market for gas, a competitive wholesale market, and competition among, often “virtual”, pipelines.[8] 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.[9] Unlike Europe, the US has a common market for gas with a single

---

[6] European Commission (2007).

[7] As cited by Ellig (1993).

[8] 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.

[9] 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.

federal regulator for all interstate commerce.[10] Tariff setting is, though still dominated by “original cost-of-service regulation” (O'Neill et al., 1996) 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 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 market institutions. Although the means differ, regulators on both sides of the Atlantic have the same objective – i.e. increasing efficiency[11] and passing on any resulting gains to consumers.[12] In the spirit of Shleifer (1985) both approaches should provide similar incentives for (static) efficiency increases.[13] As Figure 1 shows, regulatory change was accompanied by a large expansion in consumption which might be a partial indication of the success of the

---

[10] The picture is different for intrastate pipelines where regulation differs across states and generally lags behind federal regulation in terms of unbundling. For a broad overview of US gas regulation see the web page of the Natural Gas Supply Association (http://www.naturalgas.org/regulation/market.asp).

[11] 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 was 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.

[12] 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).

[13] 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).

overall regulatory change. However, our sample period is characterized by fluctuating and slightly downward trending consumption as well as increasing prices.

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.[14] We observe that the total number of companies is rather similar; measured by the physical characteristics US companies are larger; and the US network has more interconnection points. As to the last observation, it should be noted that Europe does not fare worse on the relation of interconnection points to the total length of pipelines.

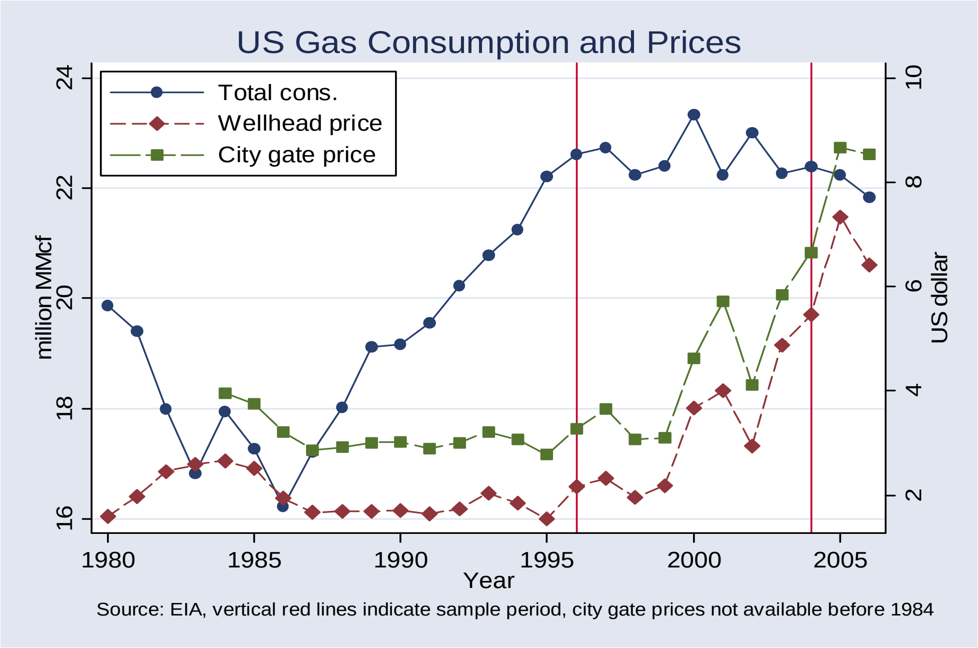


Figure 1: US Gas Consumption and Prices

---

[14] We assume 1bcf/d = 28.33 mcm/d. Because of the different sources the various figures for Europe are not necessarily for the same sample 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[15] Hubs: 14 (Energy Information Agency, 2003)	79 (GTE website)[16] 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. Finally, a major difference between the US and Europe is the tariff system, which is likely to affect incentives independently of the actual regulation and market structure. Hunt (2008) argues that the entry-exit tariff system in many European countries is a major stumbling block towards market integration.[17] Obviously, any results from an international benchmarking will also be driven by these differences in regulation, industry structure and tariff setting.

---

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

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

[17] We would like to thank Jeff Makholm for pointing this out.

3 Model and variable selection

3.1 Model

We know from the vast literature on productivity and efficiency measurement that our results are likely to be sensitive to the choice of model and variables. We model efficiency for a given unit (here a pipeline company) in a given year as the distance to an input-oriented, constant or variable returns to scale (CRS or VRS), convex, free-disposable, non-parametric frontier. We use Data Envelopment Analysis (DEA) for the static efficiency scores (CRS and VRS) and a Malmquist Productivity Index (CRS)[18] for productivity change (TFPC) and its decomposition into technical efficiency change (TEC) and technical change (TC) as well as a further decomposition of technical efficiency change into pure technical efficiency change (PTEC) and scale efficiency change (SEC).[19] Here a SEC index greater than one indicates a move towards a more efficient scale. Thus scale efficiency change relates to the existing technology. 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 and total horsepower rating and, for some models, also total delivery volume. Table 2 summarizes our models.

---

[18] 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 index with changing base years where the two periods are 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). We found no systematic differences between the results and kept 1996 as our base year.

[19] The nonparametric results are obtained using the software package FEAR. See Wilson (2007).

Table 2: Models and Variables

	Model 1	**Model 2**	**Model 3**	**Model 4**
Technology	input-orientation, constant or variable returns 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

We use 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 et al., 2002). As we find autocorrelation and heteroskedasticity (the standard assumptions on the error structure are violated) 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). These two estimators follow different strategies to overcome the violations of the standard assumptions. Whereas the FGLS approach essentially tries to model the error structure the PCSE approach simply discounts the results for such possible violations. 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 a 5% level (for the PCSE estimation the qualitative results are the same at a 10% significance level). We use likelihood ratio tests to compare different models and test the significance of delivery volume as a cost-driver.[20] Admittedly, we do not test alternative cost-drivers but rather verify the econometric significance of the variables at hand.

Once we have constructed the technical efficiency scores and the productivity indices,

---

[20] 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.

we analyze convergence among firms. Following Färe et al. (1994) and Alam and Sickles (2000), we use two established convergence concepts from the macroeconomic growth literature often referred to as β-convergence and σ-convergence. β-convergence is the notion that companies that start from a lower base grow faster (they catch up with best practice), 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 several convergence tests. First, following Färe et al. (1994) our measure of (pure) technical efficiency change is used to measure β-convergence. We also 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 on the logarithm of the technical efficiency scores in the base year. For 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 address our assumptions for the technology in some detail. First, a non-parametric frontier was chosen because in a (rate of return) 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). Second, as the Malmquist indices allow for different returns to scale across periods the CRS assumption for TFPC 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). As to our choice of input-orientation we believe that in a regulated environment it is likely that pipelines will face a derived demand.[21] Also, though non-parametric approaches lower the risk of introducing specification errors, they have been criticised.

---

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

They presume to measure the true frontier whereas this is not necessarily the case. Thus, they do not account for measurement as well as sampling error and therefore do not allow the construction of confidence intervals or hypothesis testing. The seminal work of Simar and Wilson (1998 and 2000)[22] showed how to infer the statistical properties of the non-parametric estimators and thereby overcome the problem of sampling error.[23] Following Simar and Wilson (1999) we add confidence intervals to our set of Malmquist indices that measures the productivity change for the entire sample period (1996-2004). Also, we perform a nonparametric outlier analysis which is a strategy to address possible measurement errors.[24] 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.

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.[25] 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. We use an input-oriented model that treats output as the "right-hand side" and cost or revenue as input.

First, we turn to outputs or cost-drivers. Much of the literature on gas transmission uses

---

[22] See also Simar and Wilson (2008) for a good summary of statistical inference for non-parametric frontier models.

[23] Coelli et al. (2005, p. 202) stresses that the bootstrap technique does not address measurement errors or specification errors but only sampling error. As our data is based on a sample and not the entire population the bootstrap results are informative here.

[24] We search for possible outliers using the approach of Wilson (1993) and Simar (2003) which are both implemented in FEAR. Though these tests flag possible outliers we do not drop any observations for the following reasons: there seem to be no obvious errors (though we cannot exclude the possibility of errors in the original data), the two approaches do not necessarily flag the same observations, and the order-m approach Simar (2003) flags such a great number of observations that it would be impractical to delete all of them. In particular, the order-m approach though flagging some 30% of our data at $\alpha = 0.5$ shows no "elbow-effect" (the percentage of potential outliers is constant in m).

[25] 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 have different incentive properties.

production functions where the prime output is the volume of gas delivered and capital and labour are inputs. Callen (1978) uses an engineering Cobb-Douglas function where delivery is a function of compressor 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. Granderson (2000) uses compressor fuel as 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, and 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 rating and length of mains representing capacity or capital as outputs.[26] 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. We do not account for storage whose strategic use according to O'Neill (2005) does not infer an advantage. Hence, a company that uses storage more cost effectively than others use horsepower and mains will be disadvantaged.

When looking at the rationale for using delivery as an output one has to distinguish between total cost and revenue as input variables. 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). In cost models the arguments for inclusion of

---

[26] 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.

delivery seem weak as most costs are fixed. Even most operating and maintenance (O&M) costs (except for compressor fuel and compressor maintenance) are fixed as stated by the International Energy Agency (1994, p. 48).[27] When looking at the revenue models the main reason for 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. As argued by Alger and Toman (1990), with the development of secondary markets this is no longer necessary and tariffs increasingly reflect costs. Another reason for the inclusion of delivery is that increased competition and therefore increasingly diverging business models imply different approaches to increasing capacity and delivery in ways that we do not account for (e.g. better systems). Assuming that a company uses better management or trading to increase delivery with a given capacity, a model that excludes delivery would not account for this and less innovative companies would be rewarded.

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.[28] Second, a decrease in volume might be interpreted as technical regress. Although our cross-section is larger then the one used by Nghiem and Coelli (2002) we observe technical regress for some models when including delivery volume.

Next, we turn to our input measures discussing total cost first. Unlike most of the literature we use total cost as in, for instance, Jamasb and Pollitt (2003), Hess (2008), and Edvardsen et al. (2006). The latter study uses a Malmquist cost index including

---

[27] 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.

[28] An alternative would be to estimate MPI using SFA 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).

input quantities and prices. Rouse and Swales (2006) describe how total expenditure is used as 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 for 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 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.[29] 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. 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 non-reporting.[30] Last, consumers and regulators are not interested in technical efficiency as such but the cost of the service. Hence we also use total revenue as an alternative measure of input.

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 can serve as a proxy especially where

---

[29] Traditionally prices could not be included in Malmquist indices and one would have to resort to parametric techniques. See e.g. Farsi and Filippini (2004).

[30] Additionally FERC Form 2 does not report all physical quantities of interest as for instance, the number of employees.

regulatory accounting procedures are not well-established. Second, in regulatory practice throughout Europe the rate-of-return is set in lengthy procedures reminiscent of US rate-cases and often seems to be set rather arbitrarily.[31] To some extent this defeats the purpose of incentive regulation by introducing a cost-plus element as the German monopoly commission also believes (Monopolkommission, 2006, p. 64). The use of revenue might address the critique of European 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. However, these desirable features of revenue are not fully reflected in our data as it is influenced by the rates of return as set by the regulator. Also in practice, a "benchmarkable" revenue has to be defined somehow before the benchmarking takes place.

We do not include non-discretionary variables in our analysis but suggest considering the following issues. First, the way the systems cope with peaks can differ and might not be at the discretion of management. Another 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 are relative to a particular point in the past only (Edvardsen et al., 2006). Other non-discretionary variables include 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 productivity and efficiency studies of the US gas transmission industry.

---

[31] Joskow (1972) observes that (as in Europe) there are complex rules on how the rate base is set but there is little guidance for the rates. Also unlike other components of total cost, cost of capital is unobserved.

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 in 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
Hess (2008)	47 US inter-state pipeline companies in 1996-2005	Operating expense Total transmission assets (in $)	Total revenue	SFA

4 Data and Measurement

The data used in this study is from the Federal Energy Regulatory Commission (FERC) which requires all inter-state transmission companies above a certain size to file yearly regulatory reports containing financial and operating data (FERC Form 2). As far as possible all data is confined to the transmission function. Though the FERC data is not gathered for efficiency and productivity analysis, the large number of studies (previously referred to) that uses the data testifies to its general adequacy.

Several missing values had to be estimated from adjacent periods as MPI does not tolerate any missing values. Some observations where the data did 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.[32] Revenue was adjusted such that no company had a rate-of-return lower then 6% in any year. This adjustment was to prevent the frontier to be made up of firms with sub-normal rates of return. Six percent was chosen to be slightly higher than the US 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.[33]

---

[32] For asset values for instance, that means that we do not account for inflation from the date of purchase, which would amount to current asset valuations, but from the date of reporting. For example, a firm might report a certain historic asset value in 2002, which was purchased in 1980. Our inflation adjustment only covers the period from 2002 to 2004 and not from 1980 to 2004. We used consumer price indices published by the US Bureau of Labor Statistics.

[33] 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)[34]
Mains	Total length of pipes (mains)	Miles
Horsepower	Total horsepower rating at compressor stations	HP
Age	Accumulated depreciation at mid-year / Annual depreciation	Years
Load factor	Delivery/Capacity (max. past single-day peak*365)	%
Rate of return	(Revenue – O&M – Dep.)/Average written-down value	%

Summary statistics are given in Table 5. It is evident that the size of the companies 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. The three variables age, load factor, and ROR (before adjustment) are not included in our analysis but help to

---

[34] 1 therm is equal to 100000 British thermal units (BTU).

describe the sample. For instance, the average age is 27 years which is three times our sample length.[35] As discussed below, this discrepancy can weaken some of our results.

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

We use observations for a balanced panel of 39 pipelines over 9 years. 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”.

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

---

[35] 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.

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.

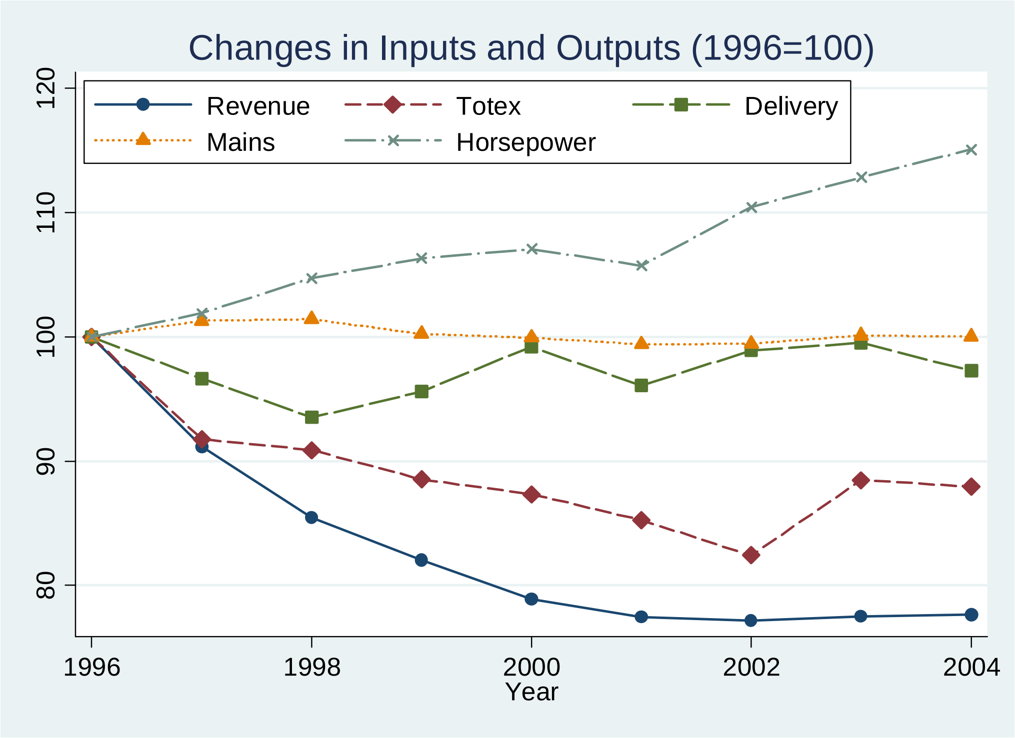


Figure 2: Changes in Inputs and Outputs (1996=100)

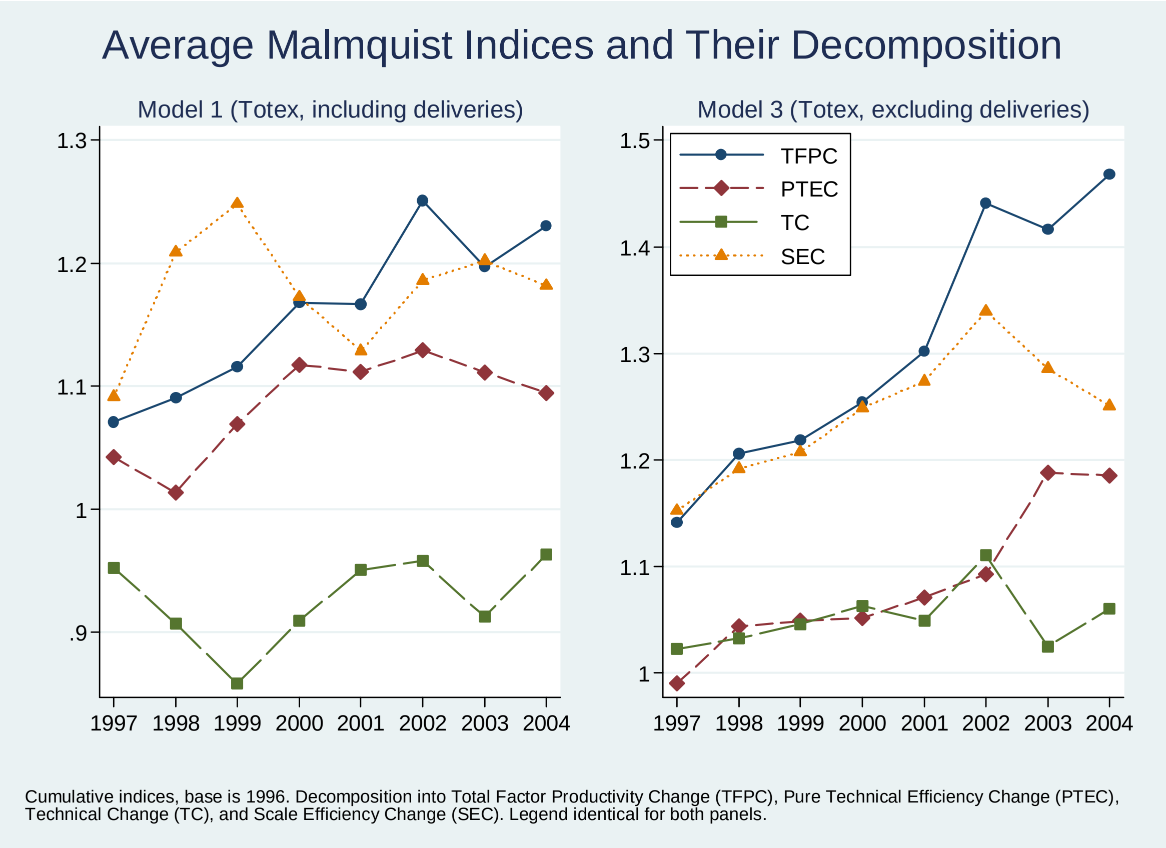
5 Results

First, Table 6 gives the results for our econometric cost driver analysis. We started our model selection 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. As we are particularly interested in the significance of delivery as a cost driver we performed the relevant likelihood ratio test. But even though the two coefficients for delivery are individually insignificant, the test does not suggest that we can drop them. Thus, the evidence on delivery volume as a cost driver is not entirely conclusive but there is a strong indication that all costs are essentially fixed (recall that compressor fuel is not included in our total cost measure). Though the results for Model 1 and Model 2 are similar, delivery has a slightly higher coefficient for the revenue model and the opposite is true for the length of pipe. It is likely that as tariffs are not fully cost reflective, revenue varies relatively more with delivery. Also, we added time trends whose coefficients confirm that cost and revenue are falling over the sample period. Last, 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).

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	

* p<0.05, ** p<0.01, *** p<0.001



Average Malmquist Indices and Their Decomposition

Model 1 (Totex, including deliveries)

Model 3 (Totex, excluding deliveries)

Cumulative indices, base is 1996. Decomposition into Total Factor Productivity Change (TFPC), Pure Technical Efficiency Change (PTEC), Technical Change (TC), and Scale Efficiency Change (SEC). Legend identical for both panels.

Figure 3: Average Malmquist Indices and their Decomposition

Now we turn to our main results – i.e. the Malmquist productivity indices. Figure 3 and Figure 4 show the cumulative (and averaged across firms) Malmquist indices (TFPC) and their decomposition into technical change (TC), pure technical efficiency change (PTEC), and scale efficiency change (SEC). Whereas the left panel in each figure includes delivery, the right panel excludes delivery as an output. Additionally, Table 11 and Table 12 in Appendix B give technical efficiency change TEC (the combination of PTEC and SEC) as well as the yearly technical efficiency scores (TE) which are inputs to the Malmquist indices. The last rows in both tables give the implied average yearly growth rates by taking the index for the last year (2004) and dividing it by the number of years elapsed since the base year.

First, we look at the industry's average performance across time. There is a clear upward trend for TFPC for all models. Also, the overall increase in PTEC is positive. Thus, firms are catching-up with best practice. However, a closer look at PTEC shows an initial increase and a decrease towards the end for PTEC and SEC. This reversal of performance is most striking for the revenue models where the turning point is the year 2000 for both models.

Across most years and models TC is very low. The technical regress for Model 1 might be caused by the fluctuations in delivery as explained above, even though the revenue model that includes delivery has the strongest growth in TC. There are other differences across the models. Models excluding delivery show a higher overall growth. This is not surprising as the length of mains and horsepower (unlike delivery) are virtually non-decreasing (see Figure 2). It is also interesting to note that the "opposite" is true for the static efficiency scores where the inclusion of delivery always increases efficiency as more variables allow more firms to be efficient.

Next, we turn to scale efficiency (change). The TE scores for VRS and CRS indicate that in each year firms do not operate at the most efficient scale. Moreover, SEC indices indicate that whereas initially firms move towards their most efficient scale this trend reverses towards the end of the sample period. Also, the inclusion of delivery seems to add volatility to the SEC measure. As SEC is also the "difference" between PTEC and TEC it is interesting to observe that the gap is larger for the Totex models.

Looking at the implied average yearly growth rates for PTEC (Appendix B), the Totex models give a range of percentage growth rates between 1.13% and 2.38%. When SEC is included the range is 3.5% to 5%. The respective ranges for the Revenue models are 0.75-1.5% and 1.75-4.3%.

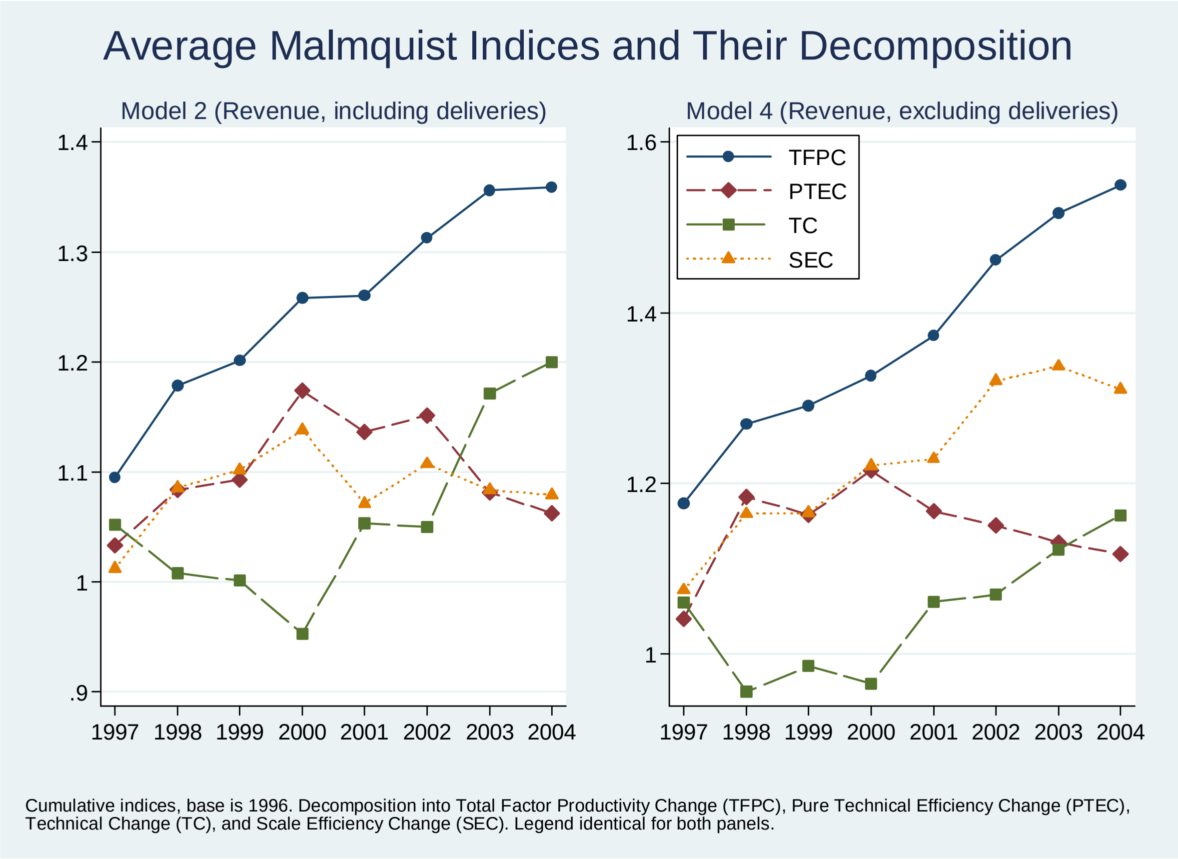


Figure 4: Average Malmquist Indices and Their Decomposition

Table 7 gives the TFPC measure and its decomposition for our last year 2004 for each individual firm (Model 1). The purpose of this table is twofold. First we show that in spite of the upward industry trend firm-level performance varies greatly. Second, we provide bootstrap based significance levels indicating whether a particular index is significantly different from 1, i.e. whether there is a significant change over the entire period or not. It appears that both TFPC and SEC are significant for almost all firms whereas technical change is mostly insignificant. The significance of the measure for (pure) technical efficiency change on the other hand varies greatly across firms.

Table 7: Malmquist Productivity Change (1996-2004) and its Decomposition

Unit	TFPC	TEC	TC	PTEC	SEC
1.	0.890***	0.902	0.987	1.000	0.902*
2.	0.842***	0.947	0.889	0.869	1.090*
3.	1.155***	1.048	1.102	1.000	1.048*
4.	0.894***	1.139	0.785	0.967	1.177*
5.	1.143***	1.306	0.875	1.082	1.207*
6.	1.646***	1.312**	1.255*	1.000	1.312*
7.	1.545***	1.496***	1.033	1.234**	1.212*
8.	0.963***	1.083	0.889	0.868	1.247*
9.	1.178***	1.454***	0.810	1.187	1.224*
10.	1.423***	1.143	1.245*	0.968	1.181*
11.	1.235***	1.554***	0.795	1.266**	1.228*
12.	1.182***	1.353*	0.873	1.000	1.353**
13.	1.765***	1.702***	1.037	1.070	1.591*
14.	1.335***	1.340***	0.996	1.191	1.125*
15.	1.015**	0.959	1.058	1.000	0.959*
16.	1.689***	1.602***	1.054	1.327***	1.207*
17.	1.134***	1.463**	0.775	1.268*	1.154*
18.	1.027	1.170	0.878	1.012	1.156*
19.	1.037	0.878	1.181	1.000	0.878*
20.	0.933***	0.818**	1.140	0.665***	1.230*
21.	0.826***	0.751**	1.099	0.600***	1.253*
22.	0.974	0.868	1.122	0.740***	1.174*
23.	0.999	1.434*	0.697*	1.499**	0.956*
24.	1.513***	1.681***	0.900	1.074	1.564*
25.	0.853***	1.281	0.666*	1.160	1.105*
26.	1.215***	1.189	1.022	0.836	1.422**
27.	0.541***	0.836	0.648*	1.000	0.836*
28.	0.993	1.146	0.866	0.994	1.153*
29.	0.823***	1.011	0.814	1.000	1.011*
30.	1.939***	2.086***	0.929	2.029***	1.028*
31.	1.158***	1.227	0.944	1.042	1.177*
32.	1.425***	1.225	1.163	1.000	1.225*
33.	0.763***	0.830	0.920	1.032	0.804*
34.	0.736***	0.799***	0.921	0.918	0.871*
35.	1.117***	1.030	1.084	0.869	1.185*
36.	0.924**	0.824*	1.121	0.737*	1.119*
37.	0.923*	1.043	0.884	1.285**	0.812***
38.	0.796***	0.742***	1.073	0.710***	1.045*
39.	1.199***	1.292	0.929	0.959	1.348*

Decomposition into Technical Efficiency Change (TEC), Technical Change (TC), Pure Technical Efficiency Change (PTEC), and Scale Efficiency Change (SEC). Stars indicate that estimates are significantly different from 1 at significance levels 1% (*), 5% (**), and 10% (***).

To further investigate the relationships between our models and various productivity (change) measures we provide pair wise correlation coefficients in Table 8 and Table 9. 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, we can make two observations. First, the correlation coefficient is higher for the level TE scores as compared to the various change measures. Second, the correlations are higher for the models that exclude the volume of delivery in Table 9. This may be due to the presence of volume-related charges affecting revenue more than total cost. Thus, when delivery is excluded from the model the remaining explanatory variables (horsepower and length of mains) have the same relative effect on Totex and Revenue leading to higher correlations.

Table 8: Pearson Correlation Coefficients, Models Including Delivery

	Model 1 (Totex)				**Model 2 (Revenue)**			
	TE (VRS)	**TFPC**	**TEC**	**TC**	**TE (VRS)**	**TFPC**	**TEC**	**TC**
TE (VRS)	1							
TFPC	-0.1292*	1						
TEC	-0.2917*	0.7283*	1					
TC	0.1993*	0.4032*	-0.3110*	1				
TE (VRS)	0.9134*	-0.1354*	-0.3354*	0.2458*	1			
TFPC	-0.0705	0.5870*	0.3973*	0.2876*	0.0374	1		
TEC	-0.0709	0.4053*	0.4363*	-0.0078	0.0451	0.8512*	1	
TC	-0.0161	0.3783*	-0.0209	0.5602*	-0.0246	0.3318*	-0.1908*	1

*indicates significance at 5%

Table 9: Pearson Correlation Coefficients, Models Excluding Delivery

	Model 3 (Totex)				**Model 4 (Revenue)**			
	TE (VRS)	**TFPC**	**TEC**	**TC**	**TE (VRS)**	**TFPC**	**TEC**	**TC**
TE (VRS)	1							
TFPC	-0.2435*	1						
TEC	-0.2850*	0.8860*	1					
TC	0.0599	0.2601*	-0.1966*	1				
TE (VRS)	0.8894*	-0.2268*	-0.2892*	0.1187*	1			
TFPC	-0.1652*	0.8043*	0.6831*	0.2381*	-0.0728	1		
TEC	-0.1936*	0.6823*	0.7485*	-0.1434*	-0.0788	0.8898*	1	
TC	0.0614	0.2904*	-0.1029	0.8490*	0.0451	0.2818*	-0.1676*	1

*indicates significance at 5%

Next, we turn to the results for the convergence test and look at β-convergence first. The increases in PTEC and TEC as shown in Figure 3 and Figure 4 are evidence for firms catching up with best practice. Figure 5 illustrates this by plotting the TE (VRS) scores for all firms in the base year against TEC component (for Model 1). Second, in order to examine the effect of the sample length on convergence, the MPI calculations were repeated for two shorter samples (moving the base up). 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. For instance, for the full sample each observation on the x-axis has eight observations along the y-axis as there are eight years till the end of the sample and therefore eight TEC measures. The fitted lines (quadratic fit) 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.

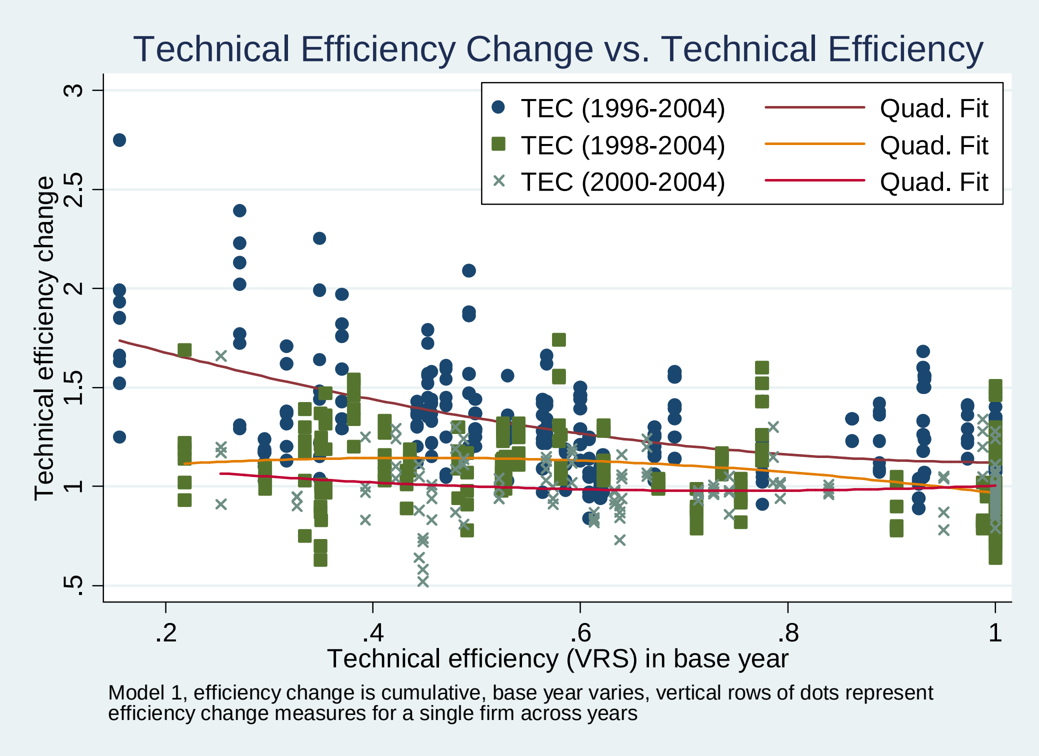


Figure 5: Technical Efficiency Change vs. Technical Efficiency

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 10 presents the results for the CRS and VRS versions of our four models. The negative coefficients confirm that there is convergence in the efficiency scores. However, the slope coefficient is most significant for models that exclude delivery. Thus, both tests indicate that there is some β-convergence.

Table 10: Results of β-Convergence Test

	Model 1 (CRS)	**Model 1 (VRS)**	**Model 2 (CRS)**	**Model 2 (VRS)**	**Model 3 (CRS)**	**Model 3 (VRS)**	**Model 4 (CRS)**	**Model 4 (VRS)**
ln(respective DEA score in 1996)	-0.010 (0.008)	-0.020* (0.009)	-0.023** (0.008)	-0.017 (0.010)	-0.026** (0.009)	-0.041*** (0.006)	-0.024* (0.011)	-0.028*** (0.007)
CONSTANT	-0.003 (0.005)	0.003 (0.008)	-0.015** (0.005)	-0.011 (0.009)	-0.009 (0.008)	-0.017* (0.007)	-0.020 (0.010)	-0.011 (0.008)
Prob>F	1.53	4.99	9.36	2.72	8.86	50.52	4.55	16.57
R-squard	0.01	0.10	0.18	0.04	0.17	0.57	0.09	0.29
obs.	39							

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Next, we turn to our results for σ-convergence. Figure 6 gives the box plots for the static efficiency scores for the VRS models. For all models the variance decreases slightly over the sample period. Looking both at the results for β-convergence and σ-convergence two interpretations seem possible. First, the catch-up effect might be so weak that performance differentials are hardly reduced. Alternatively, assuming that β-convergence is relatively strong one might conclude that even though the average firm catches up with best practice the worst performing firms do not.

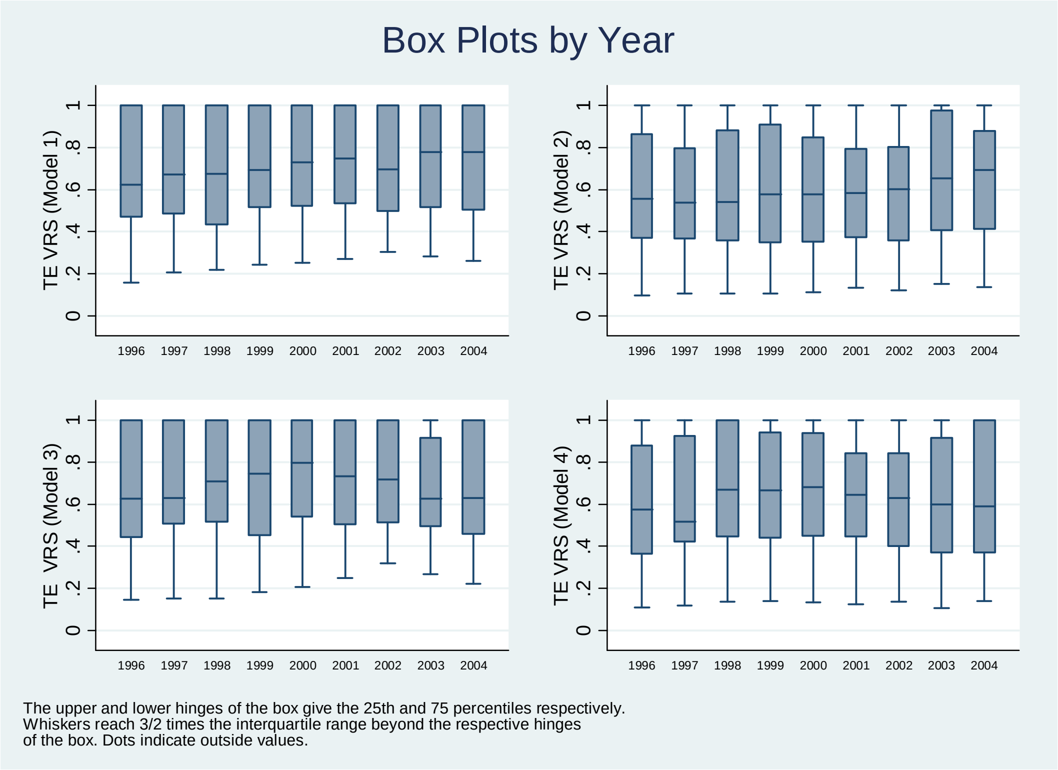


Figure 6: Box Plots by Year

6 Discussion

6.1 Model and Variable Choice

For our analysis of productivity (change) we choose non-parametric frontier models for two reasons. First, these models provide intuitive results that are meaningful to regulators and can be decomposed into various elements of total factor productivity growth. Second, the necessary assumptions on firm behaviour are less stringent than for most parametric models. To address the “deterministic” nature of our models we also report bootstrap-based significance levels for some of our results. We report both DEA technical efficiency scores and Malmquist productivity indices. Though most regulators rely on static scores from cross-sections, it is useful to contrast these with productivity changes.[36]

As to our variables, we opted for a single monetary input (total cost or revenue) which offers several advantages: only standard accounting data is required; trade-offs between the various inputs are accounted for; and 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: revenue is the total cost to consumers including cost of capital; and 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), the incentive properties of our efficiency measures are different from those of standard technical efficiency measures. Our measure would have the incentive properties of 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 observe that though the results produced by the different models are often (highly) correlated there are important differences when looking at for instance, time trends.

Again from a regulatory perspective the choice of model is important not only because the model might mis-represent how firms operate but in particular because once a model is adopted firms will start to behave accordingly.

---

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

6.2 US results

For our sample of US gas transmission companies on average, we observe modest and statistically significant improvements in technical efficiency (and stronger improvements in overall productivity). Moreover, we find some evidence of convergence over time. However, we find that these average results mask strong differences at the firm level. Also, firms do not operate at the most efficient scale. But it appears that the industry performs better than it did before.

Since the liberalization of the well-head price of natural gas in the US and subsequent regulatory change for the pipeline business, several studies (see Table 3) have measured firm-level efficiency, and productivity and in particular the effect of regulatory change. Some studies have found that efficiency declined and convergence did not occur in the first decade after well-head price deregulation.

Looking towards Europe our results might be contrasted with the result of a recent report by the German 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). This is rather low compared to our range of 2.9-6.9 percent across all models. However, the different methodologies used and the different market environment at the time make a comparison difficult.

Also, the initially strong growth of PTEC and SEC softens or reverses for some models around the year 2000. 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. This hypothesis is supported by Hess (2008) who concludes that M&A activity for US gas transmission pipelines has not lead to substantial improvement in efficiency. Increased market power would also explain the discrepancy between the paths of the Totex- and Revenue-based technical efficiency changes.

When looking again at the firm level and decomposing TEC into pure technical efficiency change and scale efficiency change we observe that SEC change dominates PTEC and that for the latter differences across firms are greater. An intuitive explanation might be that the hurdle to improve scale efficiency is lower (e.g. by installing additional compressor horsepower) than the hurdle for genuine improvements in technical efficiency.

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 somewhat over time. The convergence is not great and a sample length equivalent to the investment cycle might produce stronger results. 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. Last, not accounting for individual pipeline's particularities is likely to affect our convergence results.

Thus, our results show that regulatory change in the US is followed and accompanied by "cost productivity" and "revenue productivity" improvements. What changed is not so much the actual 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 the creation of the necessary institutions might be more important in the long-run than the prevailing form of tariff regulation.

Though we have no counterfactual our results might be used to assess the relative success of this particular regulatory regime. First, TFP change and TEC seem higher than one would expect for rate-of-return regulated natural monopolies. However, given the very low minimum values for the static efficiency scores, the reduction in the variance does not seem particularly strong. This might suggest that even though firms are catching-up on average this is not true for the worst performers. This is not entirely surprising as even though the regulatory regime seems to increase average performance, firms that are the least exposed to competition fare worst and continue to do so if local competitive conditions do not change. Here fully fledged benchmarking-based incentive regulation might have brought about stronger convergence because it explicitly targets worst performers as well as mimicking competition among all the firms.

6.3 Lessons for Europe

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

First, our work points towards issues related to data. Though FERC data collection is driven by the needs of elaborate rate-cases, its overall requirements on transparency and rigour 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 of 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 companies without a standardized European data set.[37] There is reason to believe (and other regulators have shown)[38] that a sufficient degree of comparability can be obtained with FERC data. 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.

We would like to make two general points. First, once the 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 some guidance for

---

[37] Obviously, international benchmarking requires careful analysis of the differences across firms and strategies to address these. For some of these issues see for instance Jamasb and Pollitt (2003).

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

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.

There is a finite number of gas transmission companies worldwide. 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 the US regulatory and management performance from any European benchmark could amount to forfeiting consumer surplus.

7 Conclusion

This paper investigates the performance of a sample of US interstate gas transmission pipelines from a benchmarking perspective. We assess the performance of the US industry and therefore also of the regulator using frontier productivity and efficiency measures that are widely applied in Europe today. Also, we show that European regulators can learn both from the data analysis as well as recent US regulatory reform. The lessons we draw from the latter are exploratory in the sense that we extrapolate from past US performance to future European performance. Nevertheless we believe that this study helps regulators on both sides of the Atlantic to learn from each other.

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.

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 (2003). "Directive 2003/55/EC concerning common rules for the internal market in natural gas and repealing Directive 98/30/EC."

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.

Färe, R., S. Grosskopf, M. Norris and Z. Zhang (1994). "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries." American Economic Review **84**(1): 66-83.

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.

Hess, B. (2008). Evaluating the Efficiency Effects of Industry Consolidation – Evidence from US Interstate Pipeline Companies. Efficiency Analysis Working Papers, WP-EA-14, Dresden University.

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.

Hunt, P. (2008). Entry-Exit Transmission Pricing with Notional Hubs: Can it Deliver a Pan-European Wholesale Market in Gas? Oxford Institute for Energy Studies, NG 23.

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.

Monopolkommission (2006). Sechzehntes Hauptgutachten der Monopolkommission 2004/2005.

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.

Ray, S. C. and E. Desli (1997). "Productivity growth, technical progress, and efficiency change in industrialized countries: Comment." American Economic Review **87**(1033-1039).

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.

Simar, L. (2003). "Detecting Outliers in Frontier Models: A Simple Approach." Journal of Productivity Analysis **20**(3): 391-424.

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

Simar, L. and P. W. Wilson (1999). "Estimating and bootstrapping Malmquist indices, European Journal of Operational Research." European Journal of Operational Research **115**: 459-471.

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

Simar, L. and P. W. Wilson (2008). Statistical Inference in Nonparametric Frontier Models: Recent Developments and Perspectives. The Measurement of Productive Efficiency and Productivity Change. H. O. Fried, C. A. K. Lovell and S. S. Schmidt, Oxford University Press.

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.

Wilson, P. W. (1993). "Detecting Outliers in Deterministic Nonparametric Frontier Models with Multiple Outputs " Journal of Business & Economic Statistics **11**(3): 319-323.

Wilson, P. W. (2007). "FEAR: A software package for frontier efficiency analysis with R." Socio-Economic Planning Sciences **In Press**

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,…,*K* firms (i.e. decision units) that use a single input $x^{k,t} = (X_k)$ to produce n outputs $y^{k,t} = (Y_{1k},\ldots,Y_{nk})$. 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:

$$S_{CRS}^{t}=\left\{\begin{array}{l}\left(x^{t},y^{t}\right):\sum_{k=1}^{K}z^{k,t}y_{n}^{k,t}\geq y_{n}^{t},\sum_{k=1}^{K}z^{k,t}x^{k,t}\leq x^{t},\\ n=1,\ldots N,z_{k}\geq 0,k=1,\ldots K\end{array}\right\}$$

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

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

A further decomposition of technical efficiency change into pure technical efficiency change (PTEC) and scale efficiency change (SEC) was proposed by Färe et al. (1994), where the PTEC is calculated relative to a variable returns to scale technology and the scale component “captures changes in the deviation between the variable-returns and the constant-returns-to-scale technology” (p. 74). Where scale efficiency change is

$$Scale\ Efficiency\ Change\left(SEC\right)=\left[\frac{D_{ICRS}^{t+1}\left(x^{k',t+1},y^{k',t+1}\right)/D_{IVRS}^{t+1}\left(x^{k',t+1},y^{k',t+1}\right)}{D_{ICRS}^{t}\left(x^{k',t},y^{k',t}\right)/D_{IVRS}^{t}\left(x^{k',t},y^{k',t}\right)}\right]$$

The VRS addition to the subscript now marks a set of distance functions relative to a variable returns to scale frontier; and pure technical efficiency change is

$$Pure\ Technical\ Efficiency\ Change\left(PTEC\right)=\frac{D_{IVRS}^{t+1}\left(x^{k',t+1},y^{k',t+1}\right)}{D_{IVRS}^{t}\left(x^{k',t},y^{k',t}\right)}$$

and hence,

$$TEC=PTEC*SEC$$

This approach was criticized and another decomposition suggested by Ray and Desli (1997) which has certain shortcomings itself. For a short discussion see Coelli et al. (2005, p. 292).

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

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

For the distance to the variable return technology we add the following constraint:

$$\sum_{k=1}^{K}z^{k,t}=1$$

Further details and the equivalent formulae for the mixed-period constant returns to scale distance functions are given in Grosskopf (1993).

The Malmquist index and its decomposition are illustrated in Figure A1 (for CRS only). 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.

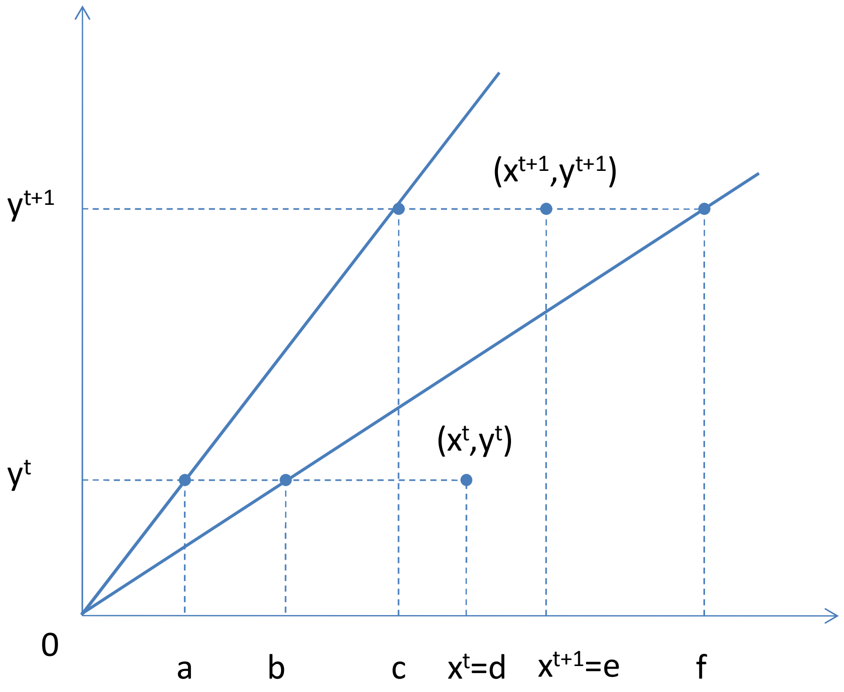


Figure A1: Illustration of a Malmquist Decomposition

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: Yearly Average Scores

Table 11: Average Malmquist Indices and their Decomposition (all relative to 1996)

	Model 1 (Totex, incl. delivery)							**Model 3 (Totex, excl. delivery)**						
Year	**TE (VRS)**	**TE (CRS)**	**TFPC**	**TEC**	**TC**	**PTEC**	**SCT**	**TE (VRS)**	**TE (CRS)**	**TFPC**	**TEC**	**TC**	**PTEC**	**SCT**
1996	0.69	0.52	1	1	1	1	1	0.59	0.45	1	1	1	1	1
1997	0.71	0.58	1.07	1.13	0.95	1.04	1.09	0.58	0.49	1.14	1.12	1.02	0.99	1.15
1998	0.69	0.61	1.09	1.22	0.91	1.01	1.21	0.60	0.50	1.21	1.18	1.03	1.04	1.19
1999	0.71	0.65	1.12	1.33	0.86	1.07	1.25	0.60	0.50	1.22	1.17	1.05	1.05	1.21
2000	0.74	0.63	1.17	1.30	0.91	1.12	1.17	0.59	0.49	1.25	1.20	1.06	1.05	1.25
2001	0.73	0.61	1.17	1.24	0.95	1.11	1.13	0.59	0.51	1.30	1.26	1.05	1.07	1.27
2002	0.73	0.63	1.25	1.33	0.96	1.13	1.19	0.60	0.52	1.44	1.32	1.11	1.09	1.34
2003	0.73	0.64	1.20	1.32	0.91	1.11	1.20	0.65	0.54	1.42	1.40	1.02	1.19	1.29
2004	0.73	0.63	1.23	1.28	0.96	1.09	1.18	0.65	0.53	1.47	1.40	1.06	1.19	1.25
growth rate p.a. (%)	-	-	2.9	3.5	-0.5	1.13	2.3	-	-	5.9	5	0.8	2.38	3.2

Table 12: Average Malmquist Indices and their Decomposition (all relative to 1996)

	Model 2 (Revenue, incl. delivery)							Model 4 (Revenue, excl. delivery)						
Year	**TE (VRS)**	**TE (CRS)**	**TFPC**	**TEC**	**TC**	**PTEC**	**SCT**	**TE (VRS)**	**TE (CRS)**	**TFPC**	**TEC**	**TC**	**PTEC**	**SCT**
1996	0.66	0.52	1	1	1	1	1	0.60	0.44	1	1	1	1	1
1997	0.67	0.54	1.10	1.04	1.05	1.03	1.01	0.61	0.48	1.18	1.11	1.06	1.04	1.07
1998	0.69	0.59	1.18	1.18	1.01	1.08	1.09	0.66	0.55	1.27	1.34	0.96	1.18	1.16
1999	0.70	0.60	1.20	1.20	1.00	1.09	1.10	0.66	0.54	1.29	1.32	0.99	1.16	1.17
2000	0.73	0.65	1.26	1.32	0.95	1.17	1.14	0.67	0.57	1.33	1.38	0.97	1.21	1.22
2001	0.71	0.60	1.26	1.21	1.05	1.14	1.07	0.64	0.54	1.37	1.32	1.06	1.17	1.23
2002	0.71	0.61	1.31	1.25	1.05	1.15	1.11	0.63	0.55	1.46	1.38	1.07	1.15	1.32
2003	0.67	0.57	1.36	1.16	1.17	1.08	1.08	0.62	0.54	1.52	1.36	1.12	1.13	1.34
2004	0.67	0.57	1.36	1.14	1.20	1.06	1.08	0.62	0.53	1.55	1.35	1.16	1.12	1.31
growth rate p.a. (%)	-	-	4.5	1.75	2.5	0.75	1			6.9	4.3	2	1.5	3.88