Incentive Regulation, Efficiency Improvements and Productivity Growth in Electricity Distribution Utility: A Norwegian Case

Livingstone Senyonga, NMBU-School of Economics and Business Phone: +47 939 72 360 E-mail: seli@nmbu.no

Abstract

Based on experience of the Norwegian electricity distribution sector, this study examines the postulation that change to incentive regulation has had a positive effect on efficiency and productivity growth in the industry. A parametric input distance stochastic frontier is used to specify the production technology. It is econometrically estimated using true fixed effects techniques by Greene (2005a, b) to generate technical efficiency scores and productivity growth rates for a panel dataset spanning 2004-2012. Malmquist productivity index is parametrically decomposed into efficiency change, technical change and changes in scale to explore the sources of changes in productivity. Compared to Miguéis et al. (2012), Edvardsen et al. (2006) and Førsund and Kittelsen (1998), this study has two advantages (1) uses a dataset that spans two regulatory regimes allowing estimates to provide a comparison of performance before and after a change in regulation (2) effects of both unobserved and observed heterogeneity in the operating environment are controlled and accounted for. The results show that overall productivity growth and technical efficiency improved significantly but scale efficiency improvements are not significant. The main conclusion is that technical change with significant embodied effects contributes more to productivity growth. Further, the results seem possibly be supported by Porter's hypothesis and regulation seems relatively more effective on lower performing distribution operators.

Key words: incentive regulation; productivity growth; input distance function; stochastic frontier analysis; electricity distribution

1.0 Introduction

The electricity industry has undergone a restructuring process the last 25 years. Competition has been introduced among generators and often also among retailers. System operation, transmission and local distribution are viewed as natural monopolies and various regulatory models and schemes have been put in place. The aims of imposing regulations on electricity distribution and regularly reviewing regulatory policies is to protect consumers from market power and ensure sustained improvements in, efficiency, productivity growth and quality of services.

Regulation of electricity distribution in Norway dates back to as early as 1990 when the energy sector was reformed (Royal Ministry of Petroleum and Energy 1990). The Norwegian Water Resources and Energy Directorate (NVE) was mandated to set policies aimed at regulating electricity distribution utilities and reviewing these policies every five years. In1991, NVE introduced rate-of-return on capital regulation, and in 1997 brought on board revenue and price cap regulation which were later anchored in efficiency benchmarking based on Data Envelope Analysis (DEA) (Førsund & Kittelsen 1998). With time, more considerations of incentives have been integrated into the regulatory framework of the distribution utility operators (DSOs) for better cost reduction without compromising the quality and security of supply. The current incentive regulation (IR) for calculation of revenue caps came into force on 1st January 2007 and was due for revision in 2011. However, NVE opted to maintain the same principles for calculating revenue caps and only revised the benchmarking model used for computing cost norms. Suffice to note that the current benchmarking model is tuned to account for data on heterogeneity in conditions under which DSOs operate like geography, climate and structural conditions.

Under the current regulation(s), the regulator does not set the price but rather puts a cap on the recoverable revenue which is determined by the computed base costs and cost norm[1](Nordic Energy Regulators 2011). The cost norm is anchored on benchmark scores and is given a greater weight in computing the revenue cap. The argument

---

[1] Revenue cap is computed by $RC_t = 0.4C_{t-1} + 0.6C_t^*$. RC_t is revenue cap determined by NVE in year t, C_{t-1} is the cost base for the DSO lagged two years and C_t^* is the cost norm -efficiency score multiplied with base cost-computed for year t considering depreciation, inflation and energy losses. Efficiency scores are computed by DEA since 1998, two stage DEA in 2010 i.e. DEA at first stage and then regression. Flexible semi-parametric methods based on panel data are being introduced (Nordic Energy Regulators 2011).

for higher weight on the cost norm is to provide incentives for effective management, utilisation and development of the network. Rungsuriyawiboon and Coelli (2004) observe that IR inspires managers to reduce costs at rates higher than allowable and recoverable rates because it allows the firm to retain any extra profit earned as a result. Hence, IR provides more incentives for cost reduction and technological innovations which reflect into efficiency improvements and productivity growth.

This study aims at finding empirical evidence in support of the hypothesis that incentive regulation methods have had positive efficiency and productivity effects on the Norwegian electricity distribution utilities. It uses panel data on 118 electricity distribution utility operators as reported to the regulation authority for 9 years 2004-2012. We are able to compare performance across regulatory regimes because data contains at least three yearly data points for which each DSO was exposed to different regulatory regimes.

We further investigate the sources of productivity growth in the Norwegian electricity distribution industry. This is done by parametrically decomposing total factor productivity change (TFPC) into change in technical efficiency (CTE), technical change (TC) and change in scale (SC). We also compare the estimates of these components across different regulatory regimes to establish if there is any evidence of the effect of incentive regulation and change in regulatory regime. In reference to environmental regulation, Porter (1991,1996) makes an interesting argument that firms benefit from tight regulation. He argues that "stronger regulation stimulates technological innovations which by enhancing productivity, increases firm's private benefits". With panel data and parametric methods, we are able to further disentangle embodied technical change from firm specific technical change which enables us to determine the productivity effects arising from innovations.

The most recent notable study about productivity growth in the Norwegian electricity distribution is by Edvardsen et al. (2006) who use index number and nonparametric DEA methods on a panel covering 8 years 1996-2003 to estimate productivity growth indices. In addition to using an up to date dataset 2004-2012 that covers the current regulatory regime, this study differs by making the following contributions:

We implement a stochastic frontier approach to benchmarking and parametrically decompose total factor productivity. Most studies conducted on the Norwegian electricity distribution relating to performance benchmarking as a basis for IR use DEA possibly because NVE has not adopted SFA in their regulatory model.

Hence, these studies suffer a disadvantage because DEA fails to separate noise from inefficiency. With SFA, we are able to easily treatment measurement errors and test hypothesis both parameters and functional form restrictions (Pantzios et al. 2011). In flexible functional forms like translog, hypothesis testing help to specify parsimonious models to find out the existence of significant scale effects, technical change, and time-varying inefficiency etc. We exploit the advantages associated with panel data which has been accumulated on regulated DSOs. As noted by Kumbhakar et al. (2012), with panel SFA we are able to account for both observed and unobserved heterogeneity and hence make consistent estimation of both technology and efficiency parameters. Our input distance SFA model is estimated by maximum likelihood estimator based on the true fixed effects model (Greene 2005a),due to its ability to account for observed and an observed heterogeneity (Greene 2004; Greene 2005b).

This study proceeds with discussion about Norwegian electricity distribution in section two, followed by model specification, methods for estimation, data description, results and conclusions in sections three, four and five respectively.

2.0 Electricity distribution Norway

Across the Nordic countries electricity grids are managed in a three level subdivision which include; distributional grid with voltage up to 22kv, regional grids handling up to between 33KV and 132KV and central grid containing high voltage lines of up to 420KV (Nordic Energy Regulators 2011). In Norway the central grid is owned and managed by TSO) called Statnett. Regional and distribution grids are managed by DSOs of differing sizes serving differing numbers and densities of customers. All DSOs are managed as private businesses much as majority are owned by municipalities or local governments. The largest DSO is Hafslund Nett AS which serves Oslo area holding 20% of the market share. It is worth noting that most of the biggest DSOs are concentrated in urban areas taking a total market share of approximately 51%, the average market share per DSO across the country is 0.8%. By 2009, 142 and 91 companies were managing distributional and regional grids respectively but a good number of companies do operate both (Nordic Energy Regulators 2011; Robles et al. 2011).

3.0 Theoretical Model

This paper models a modified cost function specified based on a multiple inputs and multiple outputs distance function with an input-orientation (Shepherd 1953;1970). Distance function specification is preferred to using a traditional cost function because of limited availability of input price data which is a common case with studies on electricity and other utility distribution (Coelli et al. 2013). Using a cost function requires making a strong assumption of cost minimization which is not always the case with utility sectors that face political influence and regulation. Distance functions are either output or input oriented distance functions and both are suitable for estimating production technologies involving multiple outputs and inputs (Coelli et al. 1998). According Saal et al. (2007), using output distance function to measure efficiency implies taking an output orientation where efficiency is improved by increasing output at a given exogenous level of inputs. On the other hand, input distance functions imply an input orientation where efficiency is improved by reducing input usage at a given exogenous level of outputs. An input orientation is suitable in modelling electricity distribution because distribution operators have more control over inputs than outputs. In practice, DSOs have limited influence over the, amount of energy they distribute, number of customers they serve, lengths of voltage line and the area operation they serve (in the short-run), all of which constitute their outputs. However, they control their expenditures in terms of capital outlay (transformers, computers, substations, vehicles, etc.), operating expenses in terms of labour (managers, technicians, engineers, accountants, etc.) and other costs like repairs and maintenance. In fact the main objective of regulating DSOs is to influence their input-mix and expenditure decisions in order to reduce costs in a manner that does not compromise the quality of their service.

3.1 Input Distance function

Suppose $x^t = (x_1^t, \ldots, x_K^t) \in R_+^k$ and $y^t = (y_1^t, \ldots, y_M^t) \in R_+^m$ are input and output vectors respectively at time $\mathrm{t} = 1, 2, \ldots, \mathrm{T}$. We define a multiple output production technology using a feasible input requirement set $L^t(y^t)$ which represents a set of input vectors, $x^t \in R_+^k$ which can produce output vectors, $y^t \in R_+^m$ such that,

$$L^t(y^t) = \{x^t \in R^{K+} : (y^t, x^t) \; where \; x^t \; can \; produce \; y^t \; at \; time \; t\}, \forall \, t = 1, \ldots, T \qquad (3.1)$$

As discussed in Coelli et al. (2005), $L^t(y^t)$ is assumed to satisfy strong disposability of inputs, is closed and convex for all outputs. However, the feasible technology defined in equation (3.1) is not known, it can be estimated from observed data points. Given data, we can measure the distance of each data point (x^t, y^t) relative to the estimated frontier. Hence, we can define an input distance function that relates to the input output relation $\left(\frac{x^t}{\rho}, \mathrm{y}^t\right)$ for $\rho \geq 1$ as;

$$D_I^t(x^t, y^t) = \underset{\rho}{Max}\left\{\rho \geq 1 : \frac{x^t}{\rho} \in L^t(y^t)\right\} \tag{3.2}$$

The input distance function $D_I^t\left(x^t, y^t\right)$ represents the maximum amount of a scalar ρ by which the input vector can radially be contracted to make the firm technically efficient without changing the output vectors. As shown by Färe and Primont (1990), $D_I^t\left(x^t, y^t\right)$ is concave, homogeneous of degree one, and non-decreasing in inputs. Also $D_I^t\left(x^t, y^t\right)$ is quasi-concave and non-increasing with respect to outputs. Further, $D_I^t\left(x^t, y^t\right) \geq 1$ for any feasible input-output mix or $x^t \in L^t(y^t)$ and $D_I^t\left(x^t, y^t\right) < 1$ when the input-output mix is infeasible or $x^t \notin L^t(y^t)$. Hence, technical efficiency $\left(\mathrm{TE}^t\right)$ is determined from input distance function as in equation (3.3).

$$\mathrm{TE}^t = \frac{1}{\mathrm{D}_I^t\left(\mathrm{x}^t, \mathrm{y}^t\right)} \text{such that } 0 \leq \mathrm{TE}^t \leq 1 \tag{3.3}$$

Because $D_I^t\left(x^t, y^t\right)$ is greater than or equal to a unit for any data point where the input vector lies within the feasible set, Equation (3.3) implies that $D_I^t\left(x^t, y^t\right)$ and TE^t are both equal to one when a firm produces at the frontier and TE^t tends to zero as $D_I^t\left(x^t, y^t\right)$ tends to infinity.

3.2 Malmquist Productivity index and productivity growth.

The paper is investigating the changes in efficiency and productivity growth over time given changes in the regulation governing electricity distribution. This is done by parametrically decomposing the Malmquist productivity growth index to obtain changes in technical efficiency, technical change and change in scale.

Caves et al. (1982) shows that the Malmquist (1953) index can be used to measures changes in productivity between two adjacent periods basing on a given reference technology. In practice, this reference technology is not known, it is represented by input distance function for a particular period or a combination of distance functions for both periods derived from input-output data. Färe et al. (1992), defines the nonparametric Malmquist index with an input orientation $\left(\mathrm{M}_{\mathrm{I}}^{\mathrm{CCD}}\right)$ as the geometric mean of corresponding Malmquist indices for adjacent periods t and $\mathrm{t}+1$ as shown in equation 3.4.

$$\begin{aligned} \mathrm{M}_{\mathrm{I}}^{\mathrm{CCD}}\left(x^{t}, y^{t}, x^{t+1}, y^{t+1}\right) &= \left[M_{I}^{t}\left(x^{t}, y^{t}, x^{t+1}, y^{t+1}\right) \times M_{I}^{t+1}\left(x^{t}, y^{t}, x^{t+1}, y^{t+1}\right)\right]^{0.5} \\ &= \left[\frac{D_{I}^{t}(y^{t}, x^{t})}{D_{I}^{t}(y^{t+1}, x^{t+1})} \times \frac{D_{I}^{t+1}(y^{t}, x^{t})}{D_{I}^{t+1}(y^{t+1}, x^{t+1})}\right]^{0.5} \end{aligned} \tag{3.4}$$

Where, corresponding to periods t and $\mathrm{t}+1$, $\left(x^{t}\ ,\ y^{t}\right)$ and $\left(x^{t+1}\ ,\ y^{t+1}\right)$ are vectors of inputs and outputs for the i^{th} firm, $D_{I}^{t}(\bullet)$ and $D_{I}^{t+1}(\bullet)$ are corresponding distance functions, $M_{I}^{t}(\bullet)$ and $M_{I}^{t+1}(\bullet)$ are the respective Malmquist indices.

Further, Coelli et al. (2005) and Pantzios et al. (2011) show that a change in technical efficiency between two periods is computed from input distance function and related to the Malmquist index as,

$$M_{I}^{t}(x_{i}^{t}, y_{1}^{t}, x_{i}^{t+1}, y_{i}^{t+1}) = \frac{D_{I}^{t}(y_{i}^{t}, x_{i}^{t})}{D_{I}^{t}(y_{i}^{t+1}, x_{i}^{t+1})} = \Delta TE(x_{i}^{t}, y_{1}^{t}, x_{i}^{t+1}, y_{i}^{t+1}) = \frac{TE_{i}^{t+1}}{TE_{i}^{t}} \tag{3.5}$$

Given $D_{I}^{t}\left(x^{t}, y^{t}\right) \geq 1$ for any feasible set of input-output mix, $\mathrm{M}_{\mathrm{I}}^{\mathrm{CCD}}\left(x^{t}, y^{t}, x^{t+1}, y^{t+1}\right)$ can be used to show whether productivity is declining, stagnant or improving when its calculated value is less than, equal to or greater than one, respectively. Studies have demonstrated that $\mathrm{M}_{\mathrm{I}}^{\mathrm{CCD}}$ index can be decomposed using either nonparametric (Coelli et al. 2005, pp 61-83) or parametric (Rungsuriyawiboon & Coelli 2004; Saal et al. 2007) techniques to reflect the contribution of a change in technical efficiency and technical change to the overall productivity growth. A change in technical efficiency for a given period of time measures the firm's movement towards the best practice frontier in the industry while technical change refers to the shift in the best practice frontier. Further decomposition of the index to reflect a change in productivity growth associated with scale effects is also possible using parametric (Balk 2001) or nonparametric methods (Färe et al. 1994). According to Färe et al. (1994), scale effects refer to changes in productivity resulting from a producer moving towards operating at a scale closer to the most productive scale size MPSS. As noted by Orea (2002), measuring productivity growth arising from scale effects using parametric approaches is usually challenging because it requires computing scale efficiency which is problematic for globally increasing, decreasing and constant returns-to-scale production technologies.

Therefore, to measure scale effects without scale efficiency, Orea (2002) re-defines the Malmquist parametric productivity index as the weighted index of output change less the weighted index of input change . Based on Saal et al. (2007), the weights are generated from estimates of the translog input distance function elasticities with respect to inputs and outputs as in equation (3.6). A negative sign is inserted before the output index to maintain it positive.

$$\ln M_I = -\frac{1}{2}\sum_{m=1}^{M}\left(\left(\varepsilon_m^{t+1} + \varepsilon_m^t\right)\left(\ln y_m^{t+1} - \ln y_m^t\right)\right) - \frac{1}{2}\sum_{k=1}^{K}\left(\left(\varepsilon_k^{t+1} + \varepsilon_k^t\right)\left(\ln x_k^{t+1} - \ln x_k^t\right)\right) \tag{3.6}$$

Where, for periods t and $t+1$ respectively, $\varepsilon_m^t = \frac{\partial \ln D_I^t}{\partial \ln y_m}$, $\varepsilon_m^{t+1} = \frac{\partial \ln D_I^{t+1}}{\partial \ln y_m}$ and, $\varepsilon_k^t = \frac{\partial \ln D_I^t}{\partial \ln x_k}$, $\varepsilon_k^{t+1} = \frac{\partial \ln D_I^{t+1}}{\partial \ln x_k}$ are distance elasticities of m outputs and k inputs obtained from derivatives of the input distance function with respect to outputs and inputs evaluated with data points at time t and $t+1$. Because the input distance function homogenous of degree one in inputs, the sum of input change weights or input elasticities is equal to a unit but the sum of output change weight does not which violates the proportionality property. The negative inverse sum of output weights -output elasticities- measures the scale elasticity RTS^t from the distance function for the technology in use (Atsbeha et al. 2012). That is,

$$\mathrm{RTS}^t = \left(-1 \Big/ \sum_{m=1}^{M} \varepsilon_m^t\right) \tag{3.7}$$

To recover the proportionality property stated by Orea (2002), equation(3.6) is transformed into a generalised form with both input and output weight equal to a unit.

$$\begin{aligned}\ln G_I = &-\frac{1}{2}\sum_{m=1}^{M}\left(\left(\varepsilon_m^{t+1} \Big/ \sum_{m=1}^{M}\varepsilon_m^{t+1}\right) + \left(\varepsilon_m^t \Big/ \sum_{m=1}^{M}\varepsilon_m^t\right)\right)\left(\ln y_m^{t+1} - \ln y_m^t\right) \\ &-\frac{1}{2}\sum_{k=1}^{K}\left(\left(\varepsilon_k^{t+1} + \varepsilon_k^t\right)\left(\ln x_k^{t+1} - \ln x_k^t\right)\right)\end{aligned} \tag{3.8}$$

Saal et al. (2007), defines $SFI_I^t = \left(\left(\sum_{m=1}^{M}\varepsilon_m^t + 1\right) \Big/ \sum_{m=1}^{M}\varepsilon_m^t\right) = 1 - RTS^t$ as the input distance scale factor relative to constant scale elasticity at time t such that, $SFI^t = 0$ when $RTS^t = 1$(constant RTS), $SFI^t > 0$ when $RTS^t < 1$(decreasing RTS) and $SFI^t < 0$ when $RTS^t > 1$ (increasing RTS) and transforms the generalised productivity index into 3.9.

$$\ln G_I = \ln M_I + \frac{1}{2}\sum_{m=1}^{M}\left(\left(\varepsilon_m^{t+1}\ SFI_i^{t+1}\right) + \left(\varepsilon_m^t\ SFI_i^t\right)\right)\left(\ln y_m^{t+1} - \ln y_m^t\right) \tag{3.9}$$

To decompose the generalised productivity index into its components, Orea (2002) used the Diewert (1976)[2] quadratic lemma which states that; if $z \in R^n$ is a vector of N arguments for quadratic function $F(z)$ evaluated at two points t and $t+1$ then, $F(z^{t+1}) - F(z^t) = \sum_{r=1}^{R}\frac{1}{2}\left[\nabla F(z^{t+1}) + \nabla F(z^t)\right]\left[z^{t+1} - z^t\right]$ where $\nabla F = \dfrac{\partial F}{\partial z}$. If $D_I^t(x^t, y^t)$ is quadratic with arguments $x^t \in R_+^k$, $y^t \in R_+^m$ and t, the difference of the input distance function estimates evaluated at data points t and $t+1$ is presented as,

$$\begin{aligned} -\left[\ln D_I^{t+1} - \ln D_I^t\right] \equiv & -\frac{1}{2}\sum_{m=1}^{M}\left(\left(\varepsilon_m^{t+1} + \varepsilon_m^t\right)\left(\ln y_m^{t+1} - \ln y_m^t\right)\right) \\ & -\frac{1}{2}\sum_{k=1}^{K}\left(\left(\varepsilon_k^{t+1} + \varepsilon_k^t\right)\left(\ln x_k^{t+1} - \ln x_k^t\right)\right) \\ & -\frac{1}{2}\left[\left(\partial \ln D_I^{t+1}/\partial t\right) + \left(\partial \ln D_I^t/\partial t\right)\right] \end{aligned} \tag{3.10}$$

If we substitute for $-\ln D_I^t = \ln TE^t$ - the Farrell's inverse technical efficiency (equation 3.3) into equation (3.10) and combine it with 3.9 and 3.6, we obtain the final decomposition of the generalised total factor productivity change (Saal et al. 2007) as,

$$\begin{aligned} \ln G_I = & \left[\ln TE_i^{t+1} - \ln TE_i^t\right] \\ & + \frac{1}{2}\left[\left(\partial \ln D_I^{t+1}/\partial t\right) + \left(\partial \ln D_I^t/\partial t\right)\right] \\ & + \frac{1}{2}\sum_{m=1}^{M}\left(\left(\varepsilon_m^{t+1} SFI_i^{t+1}\right) + \left(\varepsilon_m^t SFI_i^t\right)\right)\left(\ln y_m^{t+1} - \ln y_m^t\right) \end{aligned} \tag{3.11}$$

[2] The quadratic identity lemma states that *"the difference in quadratic function of N variables evaluated at two points is exactly equal to the sum of the arithmetic average of the first order partial derivatives of the function evaluated at the two points times the differences in the independent variables"(Diewert 2002).*

Equation (3.11) represents firm level parametric decomposition of change in total factor productivity ($TFPC_{t,t+1}$)[3] into change in technical efficiency ($CTE_{t,t+1}$)[4], technical change ($TC_{t,t+1}$)[5] and change in scale ($SC_{t,t+1}$)[6] respectively between period's t and t+1 whose relationship is shown in equation (3.16)

$$\mathrm{TFPC}_{t,t+1} = \mathrm{CTE}_{t,t+1} + \mathrm{TC}_{t,t+1} + \mathrm{SC}_{t,t+1} \tag{3.12}$$

As Orea (2002) observes – the scale effect component in equation (3.11) can accurately account for the impact of change in scale on productivity growth without imposing constant returns to scale technology. Thus, estimates in 3.11 are computed from stochastic frontier estimates for each firm i at specific time t evaluated with input-output data at adjacent periods.

3.3 The Empirical model

We have defined the t period's production for the i^{th} firm which uses k inputs x_{ik}^t, m outputs y_{im}^t and faces technical change as an input distance function - $D_I^t\left(x^t, y^t\right) = F[(x_{i1}^t,...,x_{iK}^t, y_{i1}^t,...,y_{iK}^t), t]$. As required by Färe and Primont (1990), we normalise with respect to the r^{th} input $x_{ir}^t \in x_{ik}^t$ to make the input distance function homogeneous of degree one in inputs.

$$\frac{1}{x_{ir}^t} = F\left[\left(y_{i1}^t, \ldots, y_{iM}^t, \frac{x_{i1}^t}{x_{ir}^t}, \ldots, 1, \ldots, \frac{x_{iK}^t}{x_{ir}^t}\right), t\right] - D_I^t\left(x^t, y^t\right) \tag{3.13}$$

For empirical purpose, we impose a translog function form on $F(\bullet)$ because it is, simple to derive its quadratic form in arguments (x_{ik}^t, y_{im}^t and t), linear, flexible enough to impose homogeneity and test hypotheses on parameters (Christensen et al. 1973). Taking natural logarithms on both sides of model (3.13) and introduce the translog transformation for $F(\bullet)$, and $g(Z_{iq}^t)$ -a linear function of observable exogenous variables, we obtain a translog input distance function with parameter vector θ as,

---

[3] $\ln G_I = TFPC_i^{t,t+1} = (\ln TFP_i^{t+1} - \ln TFP_i^t)$

[4] $CTE_{t,t+1} = \ln TE_i^{t+1} - \ln TE_i^t$

[5] $TC_{t,t+1} = \frac{1}{2}\left[\left(\partial \ln D_I^{t+1}/\partial t\right) + \left(\partial \ln D_I^t/\partial t\right)\right]$

[6] $SC_{t,t+1} = \frac{1}{2}\sum_{m=1}^{M}\left(\left(\varepsilon_m^{t+1}\ SFI_i^{t+1}\right) + \left(\varepsilon_m^t\ SFI_i^t\right)\right)\left(\ln y_m^{t+1} - \ln y_m^t\right)$

$$-\ln x_{ij}^{t} = TL\left(y_{im}^{t}, x_{ik}^{*t}, t, \theta\right) + g(Z_{iq}^{t}) - u_{it} \text{ , where } x_{ir}^{*t} = x_{ik}^{t} \big/ x_{ir}^{t} \text{ and } u_{it} = \ln D_{i}^{t}(x^{t}, y^{t}) \tag{3.14}$$

The time-varying one sided error term u_{it} captures firm-level inefficiencies that change over time. Introducing v_{it} the stochastic noise term, the translog input distance stochastic frontier is,

$$\begin{aligned}
-\ln x_{ir}^{t} = {} & \alpha_0 + \sum_{m=1}^{M} \alpha_m \ln y_{im}^{t} + \frac{1}{2}\sum_{m=1}^{M}\sum_{n=1}^{M} \alpha_{mn} \ln y_{im}^{t} \ln y_{in}^{t} \\
& + \sum_{k=1}^{K-1} \beta_k \ln x_{ik}^{*t} + \frac{1}{2}\sum_{k=1}^{K-1}\sum_{j=1}^{K-1} \beta_{kj} \ln x_{ik}^{*t} \ln x_{ij}^{*t} \\
& + \sum_{k=1}^{K-1}\sum_{m=1}^{M} \phi_{mk} \ln y_{im}^{t} \ln x_{ik}^{*t} + \gamma_1 t + \frac{1}{2}\gamma_2 t^2 \\
& + \sum_{m=1}^{M} \eta_m t \ln y_{im}^{t} + \sum_{k=1}^{K-1} \psi_k t \ln x_{ik}^{*t} \\
& + \sum_{q=1}^{Q} \xi_q Z_{iq}^{t} - u_{it} + v_{it}
\end{aligned} \tag{3.15}$$

The regularity condition requires the input distance function to be homogeneous of degree $+1$ in input quantities and satisfy Young's symmetry requirement for second order parameters, which imply imposing restrictions (3.16) and (3.17) respectively.

$$\sum_{k=1}^{K} \beta_k = 1, \sum_{k=1}^{K} \beta_{kj} = 0, \sum_{k=1}^{K} \phi_{km} = 0 \text{ and } \sum_{k=1}^{K} \psi_k = 0 \tag{3.16}$$

$$\alpha_{nm} = \alpha_{mn} \;\forall\, m, n = 1, 2, \ldots, M \text{ and } \beta_{kj} = \beta_{jk} \;\forall\, k, j = 1, \ldots, K \tag{3.17}$$

3.4 Econometric specification

The distance functions (3.15) can be estimated by either nonparametric (e.g. data envelope analysis –DEA and nonlinear programming-NLP) or parametric methods like stochastic frontier analysis (SFA). Parametric methods are preferred because they, (1) allow for separation of statistical noise from inefficiency, (2) make treatment of measurement errors easy, (3) permit hypothesis testing on both parameters and functional form restrictions (Pantzios et al. 2011). In flexible functional form like translog, hypothesis testing is very important to specify parsimonious models and test whether; (a) significant technical change exists, (b) inefficiency is constant, (c) technology exhibits a particular form of returns-to-scale. The True Fixed Effects (TFE) panel estimator- by Green (2004, 2005a, b) is used because it, permits time varying inefficiency, controls for observed and unobserved heterogeneity, and separates inefficiency from unobserved heterogeneity. Exogenous environmental variables Z_{iq}^{t} are included to account for observable factors that contribute to inefficiency beyond managerial issues. The intercept coefficient α_i is allowed to vary across DSOs to control for the unobserved heterogeneity. Unlike other panel frontier estimators, TFE is also preferred for its maximum likelihood dummy variable estimator (MLDV) which consistently estimates fixed effect models without dropping time-invariant variables –a common case with geographical variables.

Given that we use three outputs, three inputs and four environmental variables, we maintain restrictions (3.16) and (3.17), and fix $m = 1, 2, 3$, $k = 1, 2$, and $q = 1, 2, 3, 4$ to econometrically estimate model (3.15). We make additional assumptions u_{it} and v_{it} as per the following discussion.

The overall error term is $\varepsilon_{it} = v_{it} - u_{it}$ where $i = 1, 2, \ldots, N$ are DSO's and $t = 1, \ldots, T$ are time periods. The random error term $\mathrm{v}_{\mathrm{it}} = iid\ N(0, \delta_v^2)$ is assumed to be independent and identically distributed following a normal density with mean zero and variance δ_v^2. Also, $u_{it} = N^+(\mu_{it}, \delta_u^2)$ is assumed to follow a strictly positive truncated normal distribution[7] (Stevenson 1980) with mean μ_{it} and variance δ_u^2. The mean inefficiency is estimated as $\mu_{it} = \alpha_i + \xi' Z_{iq}^t$ where Z_{iq}^t are observed exogenous factors whose parameters ξ_q are estimated and α_i are firm specific fixed effects capturing unobserved heterogeneity. The overall estimated variance is $\delta_{\varepsilon_{it}}^2 = \delta^2 = \delta_u^2 + \delta_v^2$. Two additional parameters $\lambda = \delta_u^2 / \delta_v^2$ and $\gamma = \delta_u^2 / (\delta_u^2 + \delta_v^2)$ are estimated purposely to test hypothesis and compare the proportion of inefficiency to noise variations in the variance of the estimated model. As noted by Wang and Schmidt (2002) and Peter (2011), we do maximum likelihood estimation in one step for both technology and exogenous variables to avoid biases associated with the two-step approach. Estimate of technical efficiency are recovered via Battese and Coelli (1988) as, $TE_{it} = E[\exp(-u_{it} \mid v_{it} - u_{it})]$.

[7] The SFPANEL module developed by Belotti et al. (2013a), implemented in STATA 13 restrict parametric SFA models with environmental variables only to be specified following a truncated normal distributions.

4.0 Data and variables

This study uses a balanced panel dataset containing 118 Norwegian electricity distribution operators for a period of 9 years 2004-2011[2]. We use three outputs, three inputs and a number of environmental (nondiscretionary variables) in the input distance SFA model. Studies for efficiency analysis in utility distribution sectors require extra caution in the choice and measurement of inputs and output variables. Given that the major objectives of regulation is to enforce cost efficiency on DSO's, knowledge of the sector's key cost drivers is a prerequisite to choose what may constitute inputs, outputs and exogenous nondiscretionary environmental variables. Our choice of variables is guided by insights gained from previous studies carried out on the Nordic electricity distribution by Førsund and Kittelsen (1998); Edvardsen et al. (2006); Growitsch et al. (2010); and Growitsch et al. (2012). Further guidance is derived from various studies on efficiency in electricity distribution in European countries like, France by Coelli et al. (2013), United Kingdom by Jamasb and Pollitt (2000) and cross-country studies by Jamasb and Pollitt (2003) and Jamasb et al. (2012).

4.1 Output Variables

The three outputs adopted are number of customers served by a given electricity distribution utility, amount of units of energy delivered to customers in megawatt-hours (MWh) and the lengths of high voltage line in kilometres (KM) operated by the DSO. As noted by Førsund and Kittelsen (1998), Jamasb and Pollitt (2000) and Robles et al. (2011), these output variables reflect the joint services which electricity distribution offers and the derived demand electricity enjoys (Neuberg 1977). As argued by Coelli et al. (2013), amount of energy delivered is the most important output because the main aim of any DSO is to deliver electricity to customers when they need it and in the right amount. Given that distribution network operator cannot decide the amount of energy to deliver, what DSOs can do is to ensure that the infrastructure they operate has the right load capacity to meet customer's needs at all times. The amount of energy delivered can be viewed to reflect the load capacity of the network. Some studies have adopted to use gross energy delivered by including network losses. For this study we use the energy delivered through the network net of losses because we view energy losses as an input - a proxy for network voltage quality and reliability.

The number of customers is another important output variable that is inseparable from amount of energy distributed. Much as most of the activities a distribution operator does like customer metering, new connections, billing, emergence calling, are closely related to the number of customers, the operator has less control and discretion over their number. Ignoring to include the number of customers alongside the amount of energy delivered may cause a bias where those that sell to small scale consumers would seem inefficient despite serving many customers.

The intensity of other activities such as vegetation clearing, routine line and substation maintenance, among others, done by a distribution operator are directly related to the size of area served or the lengths of high voltage, medium or low voltage lines operated. In addition, customer density, a combination of length of voltage line and number of customers, is directly related to the costs and efficiency of the distribution company and hence influence its ability to deliver on security of supply.

It is these three output variables in addition to number of transformers that NVE regulation uses in its DEA model for benchmarking electricity distribution utility companies (Nordic Energy Regulators 2011; Robles et al. 2011).

4.2 Input Variables

The study uses three inputs -capital expenditures and total operating expenses are measured in monetary value while the amount of energy losses into the network is measured in physical MW/h. A number of inputs may form capital expenditures for example underground cables, overhead voltage lines, transformers substations, computers and machinery whose values may be provided in monetary value or not. The two sources of investment capital include; capital contributions from new customers and capital investment contributions made by the distribution operator. We include customer's investment contributions because the output resulting from its investment will form part of the operator's output (Edvardsen et al. 2006). For this study, we chose to aggregate these sources into total capital expenses based on book value as reported in the regulatory accounts for a given year. Following Coelli et al. (2013), we opt to ignore depreciation of total capital because using net capital would bias results. Using depreciated reflects DSO's who have made a lot of recent investments as being more inefficient because their net capital is higher compare with those that made relatively equal investments earlier.

For the second monetary input, we consider non-capital expenses made by the DSO to meet its day-to-day activities. Included are expenditures on labour – managers, technicians, engineers, accountants, legal teams etc., operating costs – repairs, connection services, public relations costs, metering and billing costs, etc. We chose to aggregate these expenses with penalties arising from operator's failure to supply electricity to willing customers (cost of energy not supplied -CENS) to form one monetary variable called total operating expenses (OPEX). Inclusion of CENS is to adjust OPEX for quality of supply and reflect it as a social-economic cost recoverable from society. For Norwegian DSO's, CENS is computed as a product of the value of energy not supplied and the estimated customer's willingness to pay for uninterrupted electricity supply- details are available in Ajodhia and Hakvoort (2005) and Langset et al. (2001).

The third input is energy losses in MWhs. Based on the laws of physics, part of the energy transmitted through the network is lost as either heat or system leakages. Physical units of system losses –a bad input- acts as an imperfect substitute to the other two monetary inputs in the model. This is essential to accommodate the trade-off faced by distribution operators in making expenditure decisions (Coelli & Perelman 1999; Edvardsen et al. 2006).

4.3 Environmental factors

We have so far attempted to capture much of the observed heterogeneity between DSO's using the three inputs and outputs. However, there are other factors –observable and unobservable– outside the operator's control that may affect its efficiency. In a big country like Norway with a sparse population and diverse geography, environmental factors like location, forest cover, storms, climatic conditions and sea conditions are likely to cause significant variations in performance of DSO's. The nondiscretionary variables assumed to be exogenous and used in this study are constructed as follows:

Geography:- when benchmarking DSO's, NVE-the regulator considers a number of factors including geography and weather variables to be nondiscretionary. Ignoring the effects these factors have on DSO's performance would falsely imply they are affected equally despite the heterogeneous nature of Norway-DSO's do not have any control over these factor yet they affect their efficiency. The two weighted composite variables used **(Geo-1 and Geo-2)** are computed using principal component analysis and factor analysis from factors such as; proportion of deciduous or coniferous forests, small hydro generators, wind speed and number of coastal islands, snow cover, coastal climate and landscape[8] (Growitsch et al. 2010; Nordic Energy Regulators 2011).

Portion of underground cable: - the ratio of underground cable to the total length of voltage line operated by the DSO. Underground voltage lines are likely to significantly reduce the DSO's OPEX (reduce maintenance costs), outages and system energy losses but require high capital investments to construct. Hence, this variable is a proxy to capture heterogeneity in network quality across DSO's that may not be captured by inputs and outputs.

Growth in the number of customers:- $\mathrm{Gcus}_t = \frac{\mathrm{Cust}_t - Cust_{t-1}}{\mathrm{Cust}_t} + 1$ is the annual growth in the number of customers served by a given DSO, relative to customer numbers in 2004 and hence equal to 1 in 2004.

Time trend: - this is the time interval variable $(1 = 2004)$ which is included in the model in order to capture the effects of technical change.

4.4 Descriptive summary of data

Table 4.1 gives the summary statistics of the variables used in the model. Table 4.2 we observe that the annual average amount of energy delivered to the customer increases between 2004 and 2012.

Table 4.1: Descriptive statistics

Variables	Units	N	Mean	Std. Dev	Minimum	Maximum
Output Variables						
Energy delivered (y_1)	MWh	1062	587,198.80	1,616,877	17,825.00	16,800,000.00
Number of customers (y_2)	N	1062	22,222.86	57,911.40	947.00	562,501.00
High Voltage line (y_3)	KM	1062	789.25	1,319.11	50.00	8,744.00
Area served	KM^2	1062	79,627.24	173,611.4	4,310.97	1,516,552.00
Input Variables						
Capital (x_1)	$(€'s000)$	1062	275,328.00	554,651.70	12,978.00	4,605,669.00
Operational Expenditure (x_3)	$(€'s000)$	1062	47,001.49	95,528.41	3,358.00	955,845.00
Grid Losses (x_2)	MWh	1062	31,948.26	79,505.65	349.00	898,381.00
Environmental variables						
Underground cable	$Ratio$	1062	0.31	0.18	0.03	0.86
Customer's growth	$Ratio$	1062	1.01	0.02	0.83	1.24
Trend t $(1 = 2004)$		1062			1.00	9.00
Geographical variable 1		1062		1.48	-2.06	4.86
Geographical variable 2		1062		1.51	-0.63	11.83

[8] The two variables Geo-1 and Geo-2 used are computed by NVE and used in this study as they are provided in NVE 2011 & 2012 dataset.

A slight decline was recorded for 2006 and 2011; and the highest amount was distributed during the year 2010. Likewise, on average number of customers and the lengths voltage lines have been slowly increasing. This implies an increase customer density because the number of customers and the lengths high voltage line operated has been growing over a constant area of operation. On the other hand the observed higher increase in energy over a slowly growing voltage line signals an increase load density over time. This growth in customer density and load density reflects a growing need for services of distribution operators, and hence higher need for capital investments and operating expenses.

Further, annual average level of capital expenditures has also continued to rise while expenses on operations of the distribution utilities are fluctuating. Our expectations were that higher capital expenses would match with grid expansion but we instead observe small growth in the voltage. The possible explanation for the steadily growing capital expenditure would be grid upgrades by installing modern expensive machinery and cables. The annual average level of grid losses is falling between 2004 and 2006 but increases thereafter with the highest recorded in 2010. A puzzling scenario is why grid losses remain relatively high despite a steady increase in capital and operating expenses that would be meant to mitigate this problem. We would expect higher capital and operating expenditures to translate into higher investments in infrastructure like installing of transformers, laying more underground cables and maintenance of lines and poles which would lower energy losses.

Table 4.2: Annual Averages

Variable	**Units**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
Energy Del	GWH	530.84	553.59	542.89	558.23	566.71	578.94	618.12	569.65	596.56
Voltage line	KM	753.75	766.86	766.15	766.68	767.57	775.97	781.62	785.62	791.90
Total capital	10^6 €'s	234.18	237.21	243.56	249.81	260.00	272.11	284.35	298.21	314.25
Grid Losses	GWH	31.88	30.95	29.52	30.70	30.79	31.69	36.41	25.76	32.02
Total OPEX	10^6 €'s	54.26	52.41	54.13	55.35	61.16	60.99	63.80	68.96	57.945
Customers	N	20511	20993	21022	21318	21567	21742	21949	22204	22517

5.0 Estimation and Results

All variable are normalised by their arithmetic mean before they are converted into logarithms. The aim is to interpret the first order input distance function estimates as elasticities at sample average. Parameters in equation (3.17) are estimated using maximum likelihood dummy variable estimation programed in the SFPANEL module (Belotti et al. 2013a; Belotti et al. 2013b) implemented in STATA 13. We conducted restriction tests not only for parsimony but also to reduce multicollinearity common in most flexible functional forms like translog. As shown in Table 5.1, at 1% critical level we reject the Cobb-Douglas technology, Hicks neutral technology, constant inefficiency and no unobserved heterogeneity restrictions while the restriction for no observed heterogeneity is rejected at 5% critical level. We decided to ignore the restriction for scale neutral output technology despite failure to reject it at 10% critical level. This is because of its conflict with the joint restriction of both inputs and outputs being Hicks neutral which is rejected at 5% critical level. Likewise we ignore imposing constant returns to scale despite failing to reject it. This is done to avoid biasing the estimates for scale efficiency[9]. The unrestricted translog input distance function estimated is non-decreasing (positive) in inputs and non-increasing in outputs (negative) thus satisfying the desired conditions to avoid biased scale effects estimates as noted by Orea (2002).

Table 5.1: Properties of Norwegian electricity distribution technology

Restriction	Hypothesis	Wald test Statistic	P-Value	Decision
Cobb-Douglas technology	H0:All interaction terms are equal to zero	105.28	0.000	Reject
Input Hicks neutral technology	H0: $\psi_1 = \psi_2 = 0$	9.45	0.009	Reject
Output Hicks neutral technology	H0: $\eta_1 = \eta_2 = \eta_3 = 0$	6.02	0.111	Accept
Input and output Hicks neutral	H0: $\psi_1 = \psi_2 = \eta_1 = \eta_2 = \eta_3 = 0$	11.93	0.036	Reject
Constant Returns to Scale (RTS =1.109)	H0: $\alpha_1 + \alpha_2 + \alpha_3 = -1$	0.22	0.638	Accept
No unobserved heterogeneity	H0: $Var(u_{it}) = 0$	1717.60	0.000	Reject
No observed heterogeneity	H0: $\xi_1 = \xi_2 = \xi_2 = \xi_3 = \xi_4 = 0$	12.08	0.017	Reject
Inefficiency is constant	H0: eta $= 0$	977.24	0.000	Rejected

Table 5.2 presents the estimated SFA input distance function. The scores of time varying technical efficiency are obtained from the conditional composite error term via the approach by Battese and Coelli (1988). The estimated parameters indicate the estimated portion of variance due to inefficiency –lambda $(\lambda = 1.172)$ – is greater than one and statistically significant. Gamma (γ) =0.540 implies that the error is mainly associated with inefficiency than statistical noise.

All parameters ξ_q for the four environmental variables in the model are statistically significant with the expected signs. Note that, a negative coefficient for a z-variable indicates an increase (decrease) in technical efficiency (inefficiency) respectively and the reverse is true. The coefficient for the proportion of underground cable to the

[9] Curvature and monotonicity conditions for input distance functions require unrestricted first order input and output parameters to be positive and negative respectively.

entire voltage line ξ_1 is negative and statistically significant indicating a positive effect on technical efficiency. This result is very reasonable given the fact that underground cables are less prone not only to energy losses but also disturbances due to bad weather, vegetation cover, snow cover, etc. and require less effort and costs to maintenance. The parameters ξ_2 and ξ_3 for composite geographical variables[10] are statistically significant and positive reflecting variations in inefficiency resulting from diversity in forest cover, coastal climate, storms etc. This result emphasizes the importance of considering location differences while implementing regulatory policies. Likewise, the coefficient for annual growth in the number of customers ξ_4 is also statistically significant thereby underscoring the negative effect of the size customer base on the efficiency level of the distribution utility. Most of the activities included in the day-to-day operations of a distribution utility (metering, repairs, call service, etc.) are linked to the number of customers. Hence a higher growth in the number of customers is more likely to attract higher share of operating expenses in input requirement lest the company becomes inefficient.

Recall that data for inputs and outputs was normalised around their mean values using computed sample averages. The aim is to interpret the respective estimated first order parameters α_{m} and β_{k} for inputs and outputs as input elasticities and output elasticities for a hypothetical distribution utility operating at sample average (all variables are transformed into logarithms except trend and Z-variables).

Likewise the trend variable was normalised by taking a deviation from its mean value. Therefore, we interpret the trend parameter γ_1 as the rate of technical change the average distribution utility achieves by the mid-year of the sample (2008). The second order trend parameter γ_2 measure the annual rate of technical change experienced by the hypothesised DSO operating at sample average. Our model includes interactions of trend variables with inputs and outputs, hence parameters ψ_{k} and η_m capture the annual rate of change in input and output elasticities for the sample average DSO.

[10] Composite geographical variables are derived from weather and geographical factors using factor analysis and principal component analysis.

Table 5.2: Parameter Estimates of the input distance function 2004-2012 (N = 1062, 118 Groups)

Variable	Parameter	Coefficient	T-Stat
Amount of energy delivered (lny_1)	α_1	-0.286***	-3.00
		(0.095)	
Number of customers (lny_2)	α_2	-0.138	-0.66
		(0.209)	
Lengths of voltage line (lny_3)	α_3	-0.477** *	-2.92
		(0.163)	
Capital expenditure (lnx_1)	β_1	0.707***	29.05
		(0.024)	
System energy losses (lnx_2)	β_2	0.160***	8.13
		(0.020)	
Total operating expenses ($-lnx_3$)	Dep_var	**0.133**	
$(lny_1)^2$	α_{11}	0.854***	9.93
		(0.086)	
$(lny_2)^2$	α_{22}	1.053***	3.44
		(0.306)	
$(lny_3)^2$	α_{33}	0.123	0.69
		(0.177)	
$(lnx_1)^2$	β_{11}	0.136*	1.63
		(0.083)	
$(lnx_2)^2$	β_{22}	0.193***	3.22
		(0.060)	
$(lny_1)(lny_2)$	$\emptyset_{12}$	-0.891***	-5.27
		(0.169)	
$(lny_1)(lny_3)$	$\emptyset_{13}$	0.008	0.05
		(0.157)	
$(lny_1)(lnx_1)$	$\beth_{11}$	0.258***	3.22
		(0.080)	
$(lny_1)(lnx_2)$	$\beth_{12}$	-0.185***	-3.86
		(0.048)	
$(lny_2)(lny_3)$	$\emptyset_{23}$	-0.140	-0.74
		(0.188)	
$(lny_2)(lnx_1)$	$\beth_{21}$	-0.108	-1.15
		(0.094)	
$(lny_2)(lnx_2)$	$\beth_{22}$	0.109**	2.19
		(0.050)	
$(lny_3)(lnx_1)$	$\beth_{31}$	-0.156***	-3.11
		(0.050)	

Variable	Parameter	Coefficient	T-Stat
$(lny_3)(lnx_2)$	$\beth_{32}$	0.107**	2.26
		(0.047)	
$(lnx_1)(lnx_2)$	$\beth_{22}$	-0.069*	-1.60
		(0.043)	
t	γ_1	-0.022***	-6.64
		(0.003)	
t^2	γ_2	-0.002**	-2.36
		(0.001)	
$t(lny_1)$	η_1	-0.013**	2.17
		(0.006)	
$t(lny_2)$	η_2	0.018**	2.43
		(0.007)	
$t(lny_3)$	η_3	-0.004	-0.89
		(0.005)	
$t(lnx_1)$	ψ_1	-0.017***	-2.96
		(0.006)	
$t(lnx_2)$	ψ_2	0.012*	1.83
		(0.006)	
Environmental variables			
$Constant$	ξ_0	-0.035	-0.17
		(0.205)	
$\%\ of\ under\ ground\ cable$	ξ_1	-0.457**	-2.57
		(0.178)	
$Geographical\ variable\ 1$	ξ_2	0.109**	2.25
		(0.049)	
$Geographical\ variable\ 2$	ξ_3	0.054***	3.17
		(0.017)	
$Growth\ in\ number\ of\ customes$	ξ_4	0.342*	1.87
		(0.184)	
$Sigma(u)$: ∂_u^2		0.056***	5.91
		(0.009)	
$Sigma(v)$: ∂_v^2		0.048***	15.43
		(0.003)	
$Lambda$: $\lambda = \partial_u^2/\partial_v^2$		1.172***	106.46
		(0.011)	
$Gamma$: $\gamma = \partial_u^2/(\partial_u^2 + \partial_v^2)$		0.540	
$Log\ likelihood\ function$			1567.15

***, ** and * indicate significant at 1%, 5% and 10% critical levels respectively Reported in parentheses are standard errors.

Focusing on output elasticities, all parameter estimates α_m are negative as required, $\alpha_1 = -0.286$ and $\alpha_3 = -0.477$ the output elasticities for energy delivered and lengths of voltage line a DSO operates are statistically different from zero but $\alpha_2 = -0.138$ the elasticity of number of customers[11] is not significant. The estimates reveal the returns to scale for a distribution utility operating at sample average to be 1.109. As explained before, we are not able to reject the hypothesis that the estimated technology faces constant returns to scale despite the returns to scale being greater than one. The magnitude of α_1 and α_3 reveal physical energy delivered and the grid infrastructure respectively to be the major cost drivers for DSOs.

The technical change parameters γ_1 and γ_2 are significantly different from zero and negative. This indicates that a DSO assumed to be operating at sample average will achieve positive annual rate of technical change of 2.2 percent by the 2008-the middle year of the sample. Likewise, γ_2 estimates technical change to be increasing at a rate 0.2 percent per annum. A joint test on η_m concluded that technical change has not had significant scale effects. Jointly, $\eta_1 + \eta_2 + \eta_3 = 0.009$ showing that due to technical change the elasticity of these outputs with respect to overall output activities of electricity distribution decrease at an annual rate of 0.9 percent. Parameters η_1 and η_3 are negative while η_2 is positive, η_3 -the annual rate of change of elasticity length of the voltage line is not significant. For a sample average DSO, the elasticity energy distributed increases at an annual rate of 1.3 percent $(\eta_1 = -0.013)$ while the elasticity of customers decreases at a rate of 1.8 percent $(\eta_2 = 0.018)$ per annum. Results also reveal technical change to be input augmenting –the hypothesis of technology being hicks neutral is rejected. Estimates of ψ_k reveal that technical change significantly increases the elasticity share of capital and significant decreases in the elasticity share of energy losses. $\psi_1 = -0.017$ shows that technical change has been enhancing the productivity of capital at a rates of 1.7 percent per year and reducing the elasticity share of system losses at

[11] Estimates of the same model using Battese and Coelli (1995) gives statistically significant estimates for all outputs including number of customers.

an annual rate of 1.2 percent. Combining both input elasticity shares as $\psi_3 = -(\psi_1 + \psi_2) = -0.005$ reveals an increase in the elasticity share of expenditures on operations by 0.5 percent per annum.

From Table 5.2, all first order input distance function parameters for inputs β_k estimated at sample average are positive and statistically significant. This is consistent with theory which requires input distance functions to be non-decreasing in inputs. The estimated input elasticities are 0.707, 0.160 and 0.133[12] for capital, system energy losses and operating expenses respectively. These input elasticities indicate capital as having a higher input share (70%) which clearly reflects electricity distribution as a capital intensive industry. We also observe that an estimate of 16% of the input requirement goes into system losses as either mitigation expenses or direct losses.

5.1 Technical efficiency and incentive regulation

We seek to find empirical evidence in support of the claim that DSO's are performing better under incentive regulation (2007-2012) compared to the previous regime of regulation (2001-2006). Two potential gains have been investigated; efficiency gains and gains in productivity.

Table 5.3: Estimated Technical Efficiency Scores by year 2004 - 2012

Year	2004	2005	2006	2007	2008	2009	2010	2011	2012
Minimum	0.639	0.653	0.629	0.630	0.657	0.681	0.677	0.677	0.685
Lower Quater	0.929	0.910	0.931	0.931	0.946	0.956	0.959	0.949	0.957
Median	0.975	0.969	0.973	0.975	0.983	0.984	0.984	0.983	0.983
Average	0.934	0.930	0.932	0.936	0.943	0.947	0.948	0.951	0.950
Maximum	0.998	0.997	0.997	0.997	0.997	0.997	0.998	0.997	0.997

Focusing on technical efficiency, results indicate an increase in efficiency scores after a change to the in regulatory regime. As illustrated in Fig 5.1 and Table 5.3 the minimum, first quartile, mean and median all improve from lower to higher efficiency score levels after 2007. Observing these scores into two period categories 2004-2006 and 2008-2012, considering 2007 as a transition period to a new regulation. Results indicate higher efficiency improvements for lower performing distribution utilities. The minimum efficiency level in the entire industry increases from an average of 0.641 to 0.675 and the first quartile

[12] Input distance functions is estimated under homogeneous of degree one. Hence, the elasticity share of operating expenses is recovered from $\beta_3 = 1 - \beta_1 - \beta_2$.

increases from 0.923 to 0.953, an increase of 3.5 percent and 3.0 percent respectively. The median efficiency score remains above the mean efficiency score and improves by 1.1 percent (0.973 to 0.983) while mean efficiency score increases by 1.6 percent (0.932 to 0.948).

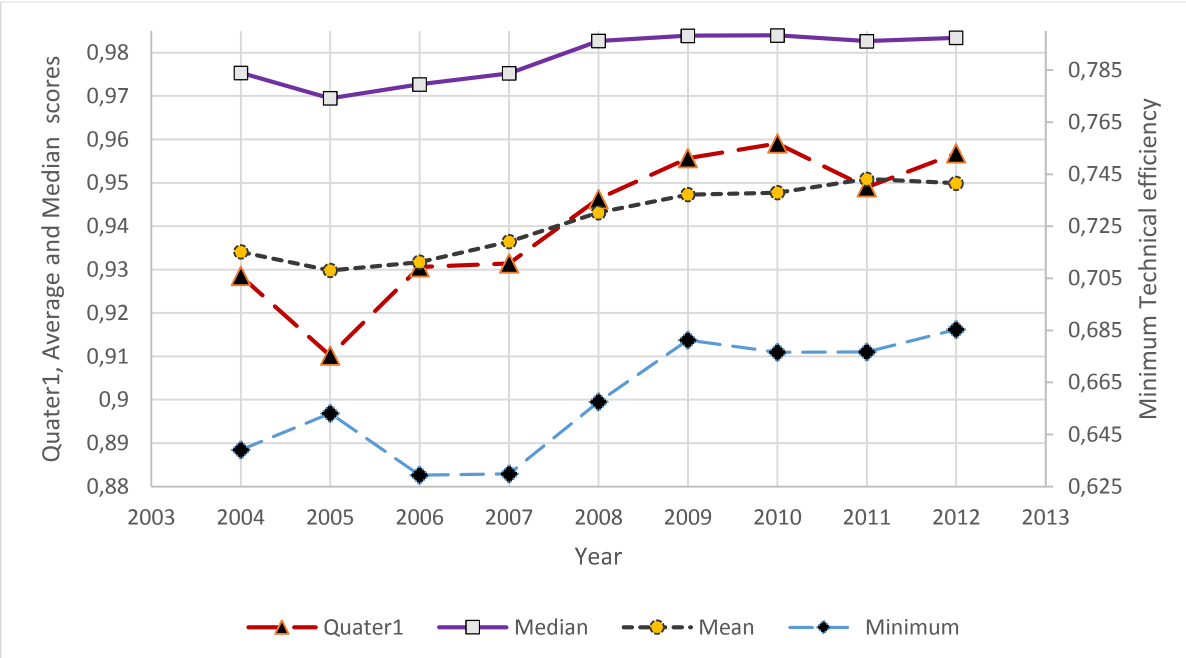

Fig 5.1: Technical efficiency scores over the period 2004 – 2012.

An interesting result is the increased negative skewness (-1.758 to -2.057) in efficiency scores -the first lower performing quarter (25%) improves from below average to above average (despite an increase in the mean) after the introduction of incentive regulation. This implies that more than 75 percent of DSO’s currently perform above average from 2008-2011. The upward shift of efficiency levels including the minimum could imply that revenue caps have had an effect on relatively inefficient DSO’s. An upward shift in efficiency is a sign that firms adjust their cost efficiency to cope with the new regulation that hinges more of their profitability on their efficiency. However, towards the end of the regulatory period we observed a slight drop in efficiency levels. This is typical of the ratchet effect because the new regulation commonly uses the costs incurred in the final year as a basis for the revenue cap of first year in the new regulatory regime which gives DSO’s incentive to over spend.

We test a number of hypotheses regarding the degree to which incentive regulation has improved efficiency and productivity across these periods.

Hypothesis 1:

The first hypothesis we consider is whether there are significant differences in average efficiency scores when DSO's were under a new regime of incentive regulation. The null hypothesis is means of TE are equal across the two periods. Results in Table 5.5 (appendix) indicate that at 1%level of significance (p=0.006) we reject the null hypothesis in favour of the alternative that the mean TE is greater in the second period.

5.2 Productivity growth and incentive regulation

Based on the estimated input distance function, we computed annual parametric productivity change indices as described in equations (3.15) and (3.16). We decomposed total factor productivity change (TFPC) into change in technical efficiency (CTE), change in technical change (CTC) and scale efficiency change (SEC). In Figure 5.2, we observe that CTE is largely positive throughout the period of the new regulation showing that efficiency improvements are sustained by majority of DSO's. The highest average efficiency improvement is observed immediately after the change –a CTE of 0.754 and 0.494 percent between 2007/8 and 2008/9 respectively., in the last period 2011/12 CTE is negative. The cumulative average CTE in only three years immediately after the change in regulation is 1.306 percent compare to 0.259 percent in the years before. Borrowing and idea from environmental regulation, this type of behaviour is consistent with our expectations that DSO's are likely to reorganise to gain competitiveness in order to cope with a new relative tighter regulation and hence perform better (Porter 1991).

Changes in TFP resulting from SEC are positive for the entire sample period except for 2010/2011 when SEC is negative 0.21percent .The estimated returns to scale is 1.109 and we could not reject the null of constant return to scale technology for a hypothetical DSO operating at sample average. This justifies the observed small values for scale efficiency change. The change seems to have little effect on scale adjustments –the average SEC is 0.141 with a cumulative of 0.687 percent in the first three years of IR while on average SEC is 0.141 with a cumulative of 0.424 for the prior period.

As discussed before, a hypothetical DSO operating at sample average is faced with a statistically significant rate of technical change γ_1 of 2.2 percent by the middle year of the sample and it is increasing at a significant rate γ_2 of 0.2 percent per year. Extending this estimate across time for all DSO's in the industry which gives an increasing trend representing time specific -pure technical change (PCTC)[13]. Test on whether technology is input and output augmenting are able to reject the null hypotheses of technology being, Hicks neutral and combined Hicks & Scale neutral. As discussed parameters η_m and ψ_k are statistically significant indicating that technical change has a significant firm-specific effects on the productivity of DSO's. Deducing from equation (3.15) we compute firm-level specific technical change (CTC)[14]. As seen in Fig 5.2, the computed annual average percentages of CTC and PCTC are both increasing and different which confirms the embodied nature of technical change.

To find out whether this embodied technical change is significant and differs across with change to the new incentive regulation, we test two hypotheses 2 and 3 respectively.

Hypothesis 2:

We test whether technical change gives substantial evidence of embodied technology. At 1% critical level, we reject the null hypothesis that the different between means of CTC and PCTC is equal to zero in favour of the alternative that difference in means is greater than zero. This confirms a significant embodied component of technical change which is line with the finding of technical change being input and output augmenting.

The level of embodied technical change is positive throughout but higher after introducing the new regulation in 2007. The difference increases from an annual average of 0.113 percent during 2004-2007 to 0.189 percent after 2007. To investigate whether this observable difference is statistically significant, we test hypothesis 3.

[13] Time specific technical change across the industry is calculated by $CPTC_{t,t+1} = -\left(\gamma_1 + \gamma_2(t+0.5)\right)$ hence $TC_{t,t+1} = \exp\left[-\left(\gamma_1 + \gamma_2(t+0.5)\right)\right]$

[14] Technical change experienced by each specific DSO is calculated (Saal et al. 2007)

$$CTC_{t,t+1} = PCTC_{t,t+1} + -0.5\left[\left(\sum_{k=1}^{m=2}\psi_k\left(\ln x_{k,i}^{t} + \ln x_{ki}^{t+1}\right)\right) + \left(\sum_{m=1}^{m=3}\eta_m\left(\ln y_{m,i}^{t} + \ln y_{mi}^{t+1}\right)\right)\right]$$

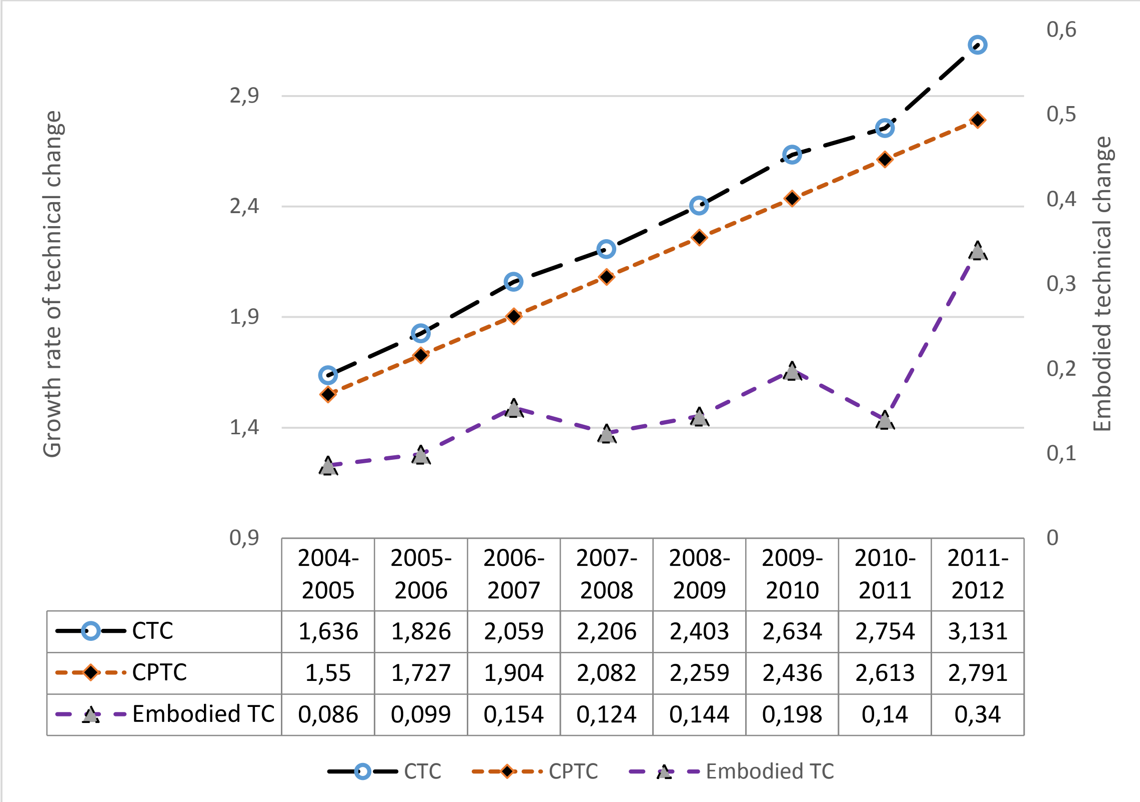

Fig 5.2: Average annual percentage change in technical change (CTC & PCTC)

Hypothesis 3.

We test whether embodied technical change increases after the introduction of IR. At 1% critical level, we reject the null hypothesis that the mean of embodied technical change after 2007 is equal to the mean before in favour of the alternative that the mean of embodied technical change is higher during 2008-2011

Table 5.4: Average total factor productivity 2004-2012

Year	**2004/05**	**2005/06**	**2006/07**	**2007/08**	**2008/09**	**2009/10**	**2010/11**	**2011/12**
Efficiency change	-0.489	0.184	0.564	0.754	0.494	0.058	0.411	-0.126
Technical change	1.636	1.826	2.059	2.206	2.403	2.634	2.754	3.131
Scale efficiency	0.192	0.026	0.206	0.174	0.145	0.368	-0.210	0.250
Productivity growth	1.339	2.035	2.829	3.134	3.042	3.060	2.955	3.256
Share decomposition of average productivity rate								
Efficiency change	-0.365	0.090	0.199	0.241	0.162	0.019	0.139	-0.039
Technical change	1.222	0.897	0.728	0.704	0.790	0.861	0.932	0.962
Scale efficiency	0.143	0.013	0.073	0.056	0.048	0.120	-0.071	0.077
Productivity growth	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

From Figure 5.3 and Table 5.4, we observe maximum annual average productivity growth in industry is registered during the 2007/8 and 2011/12. We note that total factor growth rates in the industry are largely driven by technical change. Table 5.4 shows the distribution of percentage factor shares as 86.2, 8.5 and 5.3 for technical change, efficiency change and scale efficiency change respectively. Figure 5.3 and 5.4

shows during incentive regulation a significant sustained improvement in productivity growth in the industry is registered. Average annual factor productivity growth rate increases from 2.068 to 3.078 percent after 2007 - a significant difference at 1% critical level as shown ***hypothesis 4*** in Table 5.4 (the null hypothesis is: difference in mean annual TFPC before and after 2007 is zero versus the alternative that mean TFPC is greater after 2007).

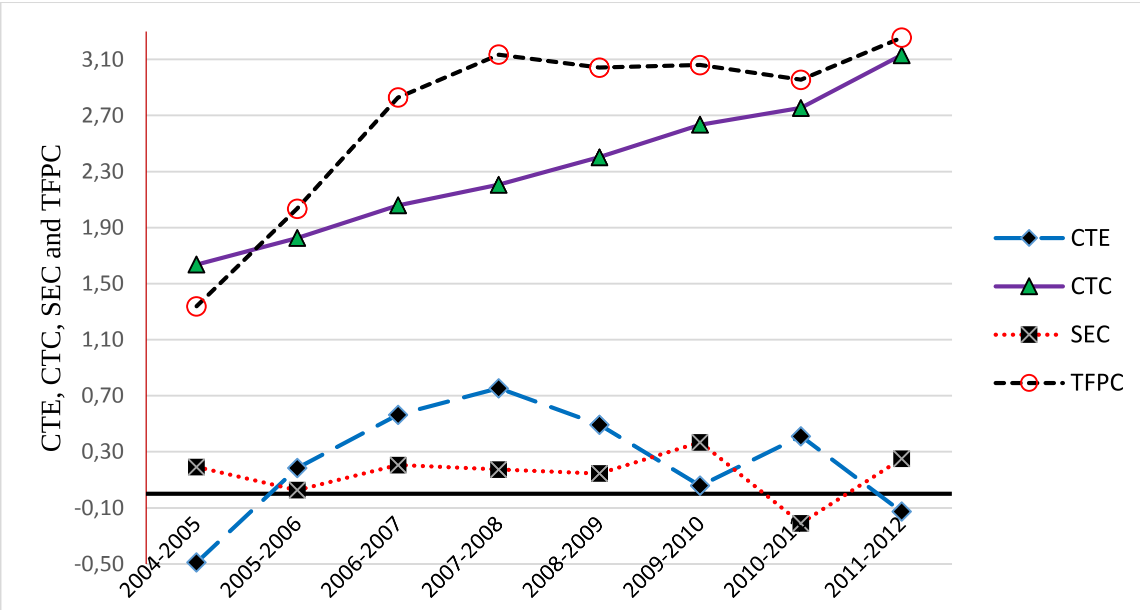

Figure 5.3: Productivity growth indices 2004/05 to 2011/12

Are these changes in efficiency and productivity a results of change in regulation?

It may require analysis outside the scope of this paper to completely disentangle the effect of policy change to incentive regulation on the performance of DSO's. However, possibly argue that changes coincide with the timing of a change in regulation. Porter (1991,996) argues that firms perform better when they are subjected to stringent environmental regulation because it stimulates innovation of technological change and improved firm competiveness. As to whether the new incentive regulation is tighter than the previous regulation may be debatable. But the new regulation is inclined towards yardstick completion than to rate-of-return regulation (Shleifer 1985) that would guarantee the survival of inefficient firm. We see that the industry experiences relatively higher annual average levels of technical change. In hypotheses 2 and 3 we confirm that the industry faces relatively higher levels of embodied technical change that is both input and output augmented which is synonymous to increased level of innovation On the other hand, results indicate significant efficiency improvements in the industry. In the first hypothesis we concluded that DSO's increase their efficiency levels after the change in regulation most especially lower performing DSOs. Saal et al. (2007) and Rungsuriyawiboon and Coelli (2004) indicate this to the behaviour of firms subjected to revenue and price caps. We note that improvements in cost efficiency is necessary for a firm to remain competitive (earn profits) and survive under yardstick completion. Therefore, there seems to be link between change to incentive regulation and the observed trend in efficiency and productivity improvements.

6.0 Conclusion

This paper has analysed the impact of incentive regulation on efficiency and productivity of companies engaged in distribution of electricity in Norway. A sample of 118 DSOs observed for a period of 9-years 2004-2012 was used to estimate an input distance function stochastic frontier. True fixed effects model is estimated with three inputs, three outputs, environmental variables and unobserved heterogeneity. We estimated technology parameters, technical efficiency scores and parametrically decomposed total factor productivity into efficiency change, technical change and scale efficiency change. Hence, the paper updates the literature on productivity growth in Norwegian electricity distribution sector and demonstrates the applicability of parametric approaches to decomposing total factor productivity.

From efficiency scores calculated, we found that the industry experiences significant improvements in technical efficiency after the introduction of IR. Interestingly, the new regulation is found to have more pronounced efficiency effects in relatively more inefficient DSO's;after the introduction of IR, over 75% of DSO's perform above a higher average efficiency score. Likewise, we found that the industry experiences sustained productivity growth mainly driven by embodied technical change. Productivity growth increases by an annual average of 3.089 percent after the introduction of IR compared to a rate of 2.068 percent before.

Using empirical results, we tested four hypotheses regarding the impact of change in regulatory policy on efficiency, technical change and productivity. Conclusions from hypothesis testing indicated that during incentive regulation the industry had significant improvements in technical efficiency, technical change and productivity growth. Consistent with Porter's hypothesis, we find that during incentive regulation the industry experiences significant improvement in embodied technology and overall factor productivity growth. Hence, the policy seems to achieve its intended economic objective but a sign that a review is needed is seen –a negative growth in technical efficiency is observed for the period 2011-2012.

List of References

Ajodhia, V. & Hakvoort, R. (2005). Economic regulation of quality in electricity distribution networks. *Utilities Policy*, 13 (3): 211-221.

Atsbeha, D. M., Kristofersson, D. & Rickertsen, K. (2012). Animal Breeding and Productivity Growth of Dairy Farms. *American Journal of Agricultural Economics*, 94 (4): 996-1012.

Balk, B. M. (2001). Scale efficiency and productivity change. *Journal of Productivity Analysis*, 15 (3): 159-183.

Battese, G. E. & Coelli, T. J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of econometrics*, 38 (3): 387-399.

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

Belotti, F., Daidone, S., Atella, V. & Ilardi, G. (2013a). SFPANEL: Stata module for panel data stochastic frontier models estimation. *Statistical Software Components*.

Belotti, F., Daidone, S., Ilardi, G. & Atella, V. (2013b). Stochastic frontier analysis using Stata. *STATA Journal*, 13 (4): 719-758.

Caves, D. W., Christensen, L. R. & Diewert, W. E. (1982). The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica*, 50 (6): 1393-1414.

Christensen, L. R., Jorgenson, D. W. & Lau, L. J. (1973). Transcendental logarithmic production frontiers. *Review of economics and statistics*, 55 (1): 28-45.

Coelli, T., Rao, D. & Battese, G. (1998). *An Introduction to Efficiency and Productivity Analysis*. Second ed. Kulwar Academic Publishers, Massachusetts, USA: Springer US.

Coelli, T. & Perelman, S. (1999). A comparison of parametric and non-parametric distance functions: With application to European railways. *European journal of operational research*, 117 (2): 326-339.

Coelli, T. J., Rao, D. P., O'Donnell, C. J. & Battese, G. E. (2005). *An introduction to efficiency and productivity analysis*. Second ed. USA: Springer Science and Business Media, LLC.

Coelli, T. J., Gautier, A., Perelman, S. & Saplacan-Pop, R. (2013). Estimating the cost of improving quality in electricity distribution: A parametric distance function approach. *Energy Policy*, 53 (0): 287-297.

Diewert, W. E. (1976). Exact and superlative index numbers. *Journal of econometrics*, 4 (2): 115-145.

Diewert, W. E. (2002). The quadratic approximation lemma and decompositions of superlative indexes. *Journal of Economic and Social Measurement*, 28 (1): 63-88.

Edvardsen, D. F., Førsund, F. R., Hansen, W., Kittelsen, S. A. & Neurauter, T. (2006). Productivity and regulatory reform of Norwegian electricity distribution utilities. In Coelli, T. & Lawrence, D. (eds) *Performance Measurement and Regulation of Network Utilities*, pp. 97-131. Cheltenham, UK . Northampton, MA, USA: Edward Elgar Publishing Limited.

Färe, R. & Primont, D. (1990). A distance function approach to multioutput technologies. *Southern Economic Journal*, 56 (4): 879-891.

Färe, R., Grosskopf, S., Lindgren, B. & Roos, P. (1992). Productivity Changes in Swedish Pharamacies 1980–1989: A Non-Parametric Malmquist Approach. In Gulledge, T., Jr. & Lovell, C. A. K. (eds) *International Applications of Productivity and Efficiency Analysis*, pp. 81-97: Springer Netherlands.

Färe, R., Grosskopf, S., Norris, M. & Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries. *The American Economic Review*, 84 (1): 66-83.

Førsund, F. R. & Kittelsen, S. A. (1998). Productivity development of Norwegian electricity distribution utilities. *Resource and Energy Economics*, 20 (3): 207-224.

Greene, W. (2004). Distinguishing between heterogeneity and inefficiency: Stochastic frontier analysis of the World Health Organization's panel data on national health care systems. *Health Economics*, 13 (10): 959-980.

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

Greene, W. (2005b). Reconsidering heterogeneity in panel data estimators of the stochastic frontier model. *Journal of Econometrics*, 126 (2): 269-303.

Growitsch, C., Jamasb, T. & Wetzel, H. (2010). Efficiency effects of quality of service and environmental factors: experience from Norwegian electricity distribution: EWI Working Paper.

Growitsch, C., Jamasb, T. & Wetzel, H. (2012). Efficiency effects of observed and unobserved heterogeneity: Evidence from Norwegian electricity distribution networks. *Energy Economics*, 34 (2): 542-548.

Jamasb, T. & Pollitt, M. (2000). Benchmarking and regulation: international electricity experience. *Utilities Policy*, 9 (3): 107-130.

Jamasb, T. & Pollitt, M. (2003). International benchmarking and regulation: an application to European electricity distribution utilities. *Energy Policy*, 31 (15): 1609-1622.

Jamasb, T., Orea, L. & Pollitt, M. (2012). Estimating the marginal cost of quality improvements: The case of the UK electricity distribution companies. *Energy Economics*, 34 (5): 1498-1506.

Kumbhakar, S. C., Lien, G. & Hardaker, J. B. (2012). Technical efficiency in competing panel data models: a study of Norwegian grain farming. *Journal of Productivity Analysis*, 41: 321-337.

Langset, T., Trengedered, F., Samdal, K. & Heggset, J. (2001). Quality dependent revenue caps - a model for quality of supply regulation. *IET Conference Proceedings*: v6:4-v6:4. Available at: http://digital-library.theiet.org/content/conferences/10.1049/cp_20010906.

Malmquist, S. (1953). Index numbers and indifference surfaces. *Trabajos de Estadística y de Investigación Operativa*, 4 (2): 209-242.

Miguéis, V. L., Camanho, A. S., Bjørndal, E. & Bjørndal, M. (2012). Productivity change and innovation in Norwegian electricity distribution companies. *Journal of the Operational Research Society*, 63 (7): 982-990.

Neuberg, L. G. (1977). Two issues in the municipal ownership of electric power distribution systems. *The Bell Journal of Economics*, 8 (1): 303-323.

Nordic Energy Regulators. (2011). Economic regulation of electricity grids in Nordic countries. Copenhagen, Denmark: Nordic Energy Regulators -NordREG. 116 pp.

Orea, L. (2002). Parametric decomposition of a generalized Malmquist productivity index. *Journal of Productivity Analysis*, 18 (1): 5-22.

Pantzios, C., Karagiannis, G. & Tzouvelekas, V. (2011). Parametric decomposition of the input-oriented Malmquist productivity index: with an application to Greek aquaculture. *Journal of Productivity Analysis*, 36 (1): 21-31.

Peter, S. (2011). One-step and two-step estimation in SFA models . *Journal of Productivity Analysis* 36: 201-203.

Porter, M. (1991). America's green strategy. *Scientific American*, 264 (4): 168.

Porter, M. (1996). America's green strategy. In Welford, R. & Starkey, R. (eds) *Business and the Environment*, pp. 33-35. London , UK: Earthscan Publication Limited.

Robles, H., Ilonen, M. & Syvertsen, S. (2011). Economic regulation of electricity grids in Nordic countries. Copenhagen, Denmark: Nordic Energy Regulators -NordREG. 119 pp.

Royal Ministry of Petroleum and Energy. (1990). *Act relating to the generation, conversion, transmission, trading, distribution and use of energy etc. (The Energy Act)*, Act No.50 of 29 June 1990. Norway. pp. 43-58.

Rungsuriyawiboon, S. & Coelli, T. (2004). Regulatory reform and economic performance in US electricity generation. In Coelli, T. & Lawrence, d. (eds) *Perfomance Measurement and Regulation of Network Utilities*, pp. 267-296. Centre for Efficiency and Productivity Analysis, School of Economics , University of Queensland: Edward Elgar Publishing Limited.

Saal, D. S., Parker, D. & Weyman-Jones, T. (2007). Determining the contribution of technical change, efficiency change and scale change to productivity growth in the privatized English and Welsh water and sewerage industry: 1985–2000. *Journal of Productivity Analysis*, 28 (1-2): 127-139.

Shepherd, R. W. (1953;1970). *Cost and Production Functions*. Princeton University Press: Princeton, New Jersey.

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

Stevenson, R. E. (1980). Likelihood functions for generalized stochastic frontier estimation. *Journal of econometrics*, 13 (1): 57-66.

Wang, H.-J. & Schmidt, P. (2002). One-step and two-step estimation of the effects of exogenous variables on technical efficiency levels. *Journal of Productivity Analysis*, 18 (2): 129-144.

7.0 Appendices

Table 5.5: Hypotheses tested regarding the effect of IR

Hypothesis	Mean		T-value	P-value	Decision
	2004-2006	2008-2012			
Hypothesis (1)	*H0: means of technical efficiency before and after 2007 are equal*				
	0.9318	0.9478	2.756***	0.006	Reject
Hypothesis (3)	*H0: means of embodied technical change before and after 2007 are equal*				
	0.1130	0.1895	2.800***	0.005	Reject
Hypothesis (4)	*H0: means of productivity growth before and after 2007 are equal*				
	2.0680	3.0893	6.801***	0.000	Reject
Hypothesis (2)	*H0: difference in between CTC and PCTC is equal to zero*				
	CTC =2.170	PCTC =2.331	6.692***	0.000	Reject

8.0