Regulated and cost efficient?

A study of the norwegian electricity distribution industry

Erik Leknesund

Norwegian University of Life Sciences
Department of economics and resource management
Master Thesis 30 credits 2012

PREFACE

This thesis was written to complete a five year master's programme in Industrial Economics at the University of Life Science in Ås. My fascination for energy economics flourished in the fall 2008 when I decided to write a term paper in introductory econometrics on spot pricing in the Norwegian electricity market. Further studies within this subject have given me wide knowledge of the energy sector, in which undoubtedly will guide me in my career choices.

I must use this opportunity to thank my supervisor, Associate Professor Olvar Bergland, for his patience and inspirational knowledge. Not only during appointed consultation hours, but also after regular working hours. It has been motivating to be guided by one of the more prominent Professors at UMB.

To all of you in TF206, thank you for making me perform both academically and socially during these four months. A special thanks to Anders Bostad for helping me when I needed it. I also want to thank Thomas Mo Willig and Julian Bell for improvements on my English.

Thanks to my lovely future wife Margrete. Even when you are far away, you are always close. I love you.

Ås, May.15.2012

Erik Leknesund

ABSTRACT

The Norwegian electricity distribution industry is a natural monopoly regulated by the Norwegian Water Resources and Energy Directorate (NVE). NVE decides the regulatory model which purpose is to stimulate the industry to produce at minimum cost in order to maximize consumer surplus. To be successful producing at minimal cost the individual company has to be efficient.

This thesis estimates a translog cost frontier using the statistical package STATA 11.1 based on the Stochastic Frontier Analysis (SFA) principles. The cost frontier is used to establish the individual firm's efficiency score during the period 2007-2010. Further analysis of the cost function shows that through asserting a second order polynomial time trend, a technological progress can be expected. However these results have been difficult to conclude as a result of the short time period. Other findings through the frontier estimation are that the individual firm did not improve their technological efficiency in the period. It is suggested that today's regulatory model fail to induce Norwegian distribution companies to perform efficiently. Important theoretic concepts related to the regulatory model and frontier analysis have been elaborated.

TABLE OF CONTENT

List of figures

List of tables

ABBREVIATIONS

DC	Distribution Company
DSO	Distribution System Operator
DWL	Dead Weight Loss
NVE	Norwegian Water Resources and Energy Directorate
OLS	Ordinary Least Squares
POLS	Pooled Ordinary Least Squares
ROR	Rate of Return
SFA	Stochastic Frontier Analysis
TI	Time-invariant model
TVD	Time-Varying Decay model

1 Introduction

The electricity distribution is a natural monopoly where different firms serve customers within their given concession area (Bogetoft & Otto 2011). A customer must buy the distribution service from the provider, referred to as the DSO (Distribution System Operator), which has concession in their area of living. This is done since distribution of electricity requires high investments in infrastructure, which should be done by a single firm and not in parallel by several firms in the same area. As the only supplier in any given area the DSO is not bounded by competitive forces yielding lowered cost, lower price and improved quality. Most countries therefore have a regulator chosen by the state to interfere with operation and the charged tariffs (Bogetoft & Otto 2011). Today's economic regulatory model in the Norwegian electricity distribution sector is defined by the Energy Act of 1990 and aims to maximize social welfare through efficiency, positive development and utilization of the distribution industry. This is one of the main mandates of the national regulator, Norwegian Water Resources and Energy Directorate (NVE). The regulatory model is decided on a five year basis, but within these five year periods minor alterations are conducted (Grammeltvedt et al. 2006).

After a period where the electricity market was dominated by vertically integrated businesses, it was decided to develop a competitive market for power trading, the day-ahead market today known as Elspot. Norway was one of the first countries in the world to do so[1] (*Deregulering av det...*). As mentioned the distribution of electricity is a natural monopoly, and therefore not effectively run unless it is subject to regulation.

According to Wangensteen (2012) there are two possible mechanisms that cause economic inefficiency in a monopoly situation:

Market inefficiency is caused when the monopolistic firm prices its services too high, because of lack of competition. X–inefficiency when the cost is larger than necessary. As long as costs can be transferred to the costumers, there will be no incentives to decrease costs.

X – Inefficiencies are caused by three different reasons:

- Scale inefficiency, wrong size of the company, – the company is either too small or large.

[1] The third country after England and Wales.

- Technical inefficiency – by using more factors of production than necessary for the level of output.
- Cost (allocative) inefficiency – wrong combination of the production factors (inputs).

Over the last two decades various regulation models have been outlined and tested. The models have raised multiple discussions on which model that serves the end user best, but also how to maintain a sustainable grid industry. NVE is searching for a model that stimulates both the distribution companies (DC's) to maximize welfare and at the same time have incentives to perform efficiently. In addition the regulation must secure the DC's a reasonable profit on their invested capital (Grammeltvedt et al. 2006).

In the period from 1993-1996 the regulation was a rate of return (ROR) regulation regime. The firm's rate of return was compared with the rate decided by NVE and adjusted ex post in cases where the two rates differed.

In the subsequent two periods, 1997–2001 and 2002-2006, the model was a revenue cap regulation model.

Today's regulation model has been used since 2007 and is not solely a revenue cap model. In a revenue cap regulation model a firm's revenue is independent of the firm's true costs. However, in the Norwegian model, which can be seen as a hybrid between an revenue cap regulation model and a ROR regulation model, the revenue is related to the cost by a cost norm (Von Der Fehr 2010). The companies' own costs (K_t) are weighted 40 percent, the business' normalized costs (K_t^*) are weighted 60 percent and the weight (ρ) is decided by NVE. With this comes two effects, firstly it reduces the firm's incentives to reduce costs and secondly it reduces the firms possibility to increase prices with increased costs (Von Der Fehr 2010).

Von Der Fehr (2010) states that the regulatory models used the last 10 – 15 years have tightened the distribution companies' economy, but claims that there is no reason to believe that quality and capacity has been significantly affected. If anything, it has been a solid economy with over investments that have been the problem, not under investments and lack of return. (Von Der Fehr 2010).

The regulatory model provides incentives to reduce costs since 40 percent of the allowed revenue is based on the firm's own costs (Von Der Fehr 2010). In order to reduce costs the firms must seek to improve their economic efficiencies. The regulator's goal is to induce the firms to increase output, decrease price, and produce at minimum cost. If this is done successfully, the optimal outcome is the alternative that serves the firm and society best under the given form of regulation (Train 1991).

1.1 Objective

The objective of this thesis is to determine whether or not there has been an improvement in the distribution industry's cost efficiency under today's regulatory model. Using Stochastic Frontier Analysis (SFA) a cost frontier based on reported data collected by NVE in the period 2007-2010 is estimated. The validity of the estimated cost frontier will be discussed and the distribution of the industry's firm specific cost efficiencies will be estimated. Further it will be discussed if the improvement (or decrease) is related to technological progress or if it is due to improvements in the firm's individual technological efficiency. Relevant theory will be presented and the regulatory model will be described.

1.2 Organisation of the thesis

This thesis is divided in four main chapters; Introduction, Theory, Results and Discussion with Conclusion.

Chapter 1 is introductory and informs the reader on content, objective and organisation of the thesis.

Chapter 2 reviews relevant econometric theory related to the stated objective. The first part of chapter 2 presents economic regulation and why this is necessary in a natural monopoly situation. The Norwegian non-linear price structure is presented theoretically in chapter 2.1.4. Further, terms related to economic efficiency is elaborated and explained graphically in chapter 2.2. Chapter 2.3 gives a historical view on the regulatory models. This chapter explains shortly important aspects of the models from 1993 until today. An introduction to the cost frontier and the yardstick principle is presented in chapter 2.4, before a more thorough explanation of the Data Envelopment Analysis (DEA) in chapter 2.5. This chapter explains how the DEA method estimates the individual efficiency scores with belonging limitations and assumptions. Chapter 2.6 introduce important aspects of the Stochastic Frontier (SFA) and two different functional forms are discussed shortly. This chapter shows how the estimated cost frontier determines the individual efficiency scores. Chapter 2 is completed by explaining the means of a time trend and the behaviour of inefficiencies in a SFA.

Chapter 3 begins with introducing the estimated model and its parameters. A simple model using Pooled Ordinary Least Squares (POLS) regression is tested with different tests such as functional form. Results from POLS regression are used as a base for further exploration and testing of the cost frontier, which is presented in chapter 3.2. Further the individual efficiencies in the period are estimated using two different assumptions.

Chapter 4 includes discussion and conclusions of the results found in chapter 3.

2 Background theory

The literature section is meant as a discussion of the main work in this thesis. For all readers to fully understand the model and its purpose it is naturally to discuss certain subjects of the economic theory mentioned. The discussion is not meant as a reference in its own, but rather a clarifying part of the regulation models true purpose of origin and anatomy. However, the reader is assumed to have some background knowledge on basic micro economics especially related to energy economics and econometrically relations. Some relevant basic econometric theory and assumptions are presented in appendix C.

2.1 Economic regulation

In this thesis the word regulation is correlated to activity controlled by the government. As explained in the introduction and extended in later chapters, regulation is necessary when it is obvious that a competitive market will not suffice. In the distribution industry the lack of competition removes incentives for efficient utilisation of society's resources. In order to prevail against inefficiencies and maximise welfare, different forms of regulation can be conducted. Before the introduction of such regulation methods, the terms maximum welfare, natural monopoly and economic efficiency will be elaborated.

2.1.1 Maximising welfare

One measure for efficiency in this context is the size of the economic surplus, which equals the sum of consumer and producer surplus. The economic surplus is at its greatest when the cost related to production of the last unit equals the consumer's marginal willingness to pay, in this case where demand (D) meets marginal cost (MC) in Figure 2-1. An economically efficient solution is therefore characterized as price equal the short term marginal cost in production which equals the consumer's willingness to pay. The market price's most important role is to clear the market in the short term. Therefore, in times with capacity shortage the price will ration such that the resources would accrue to consumers with the highest willingness to pay (Berg & Tschirhart 1988). On a longer term with free competition all producers are price takers. The price is given by the market since there are sufficiently many producers and neither of these producers are large compared to the size of their market. A single producer's supply to the market will not affect the total market supply and the price will equal the marginal cost. In this case, each producer will make a normal profit that equals the cost of capital. This is called “the first best solution” indicating that no other outcome provides a higher surplus. This is not easy to accomplish under monopoly regulation without additional instruments, such as lump sum transfers (Berg & Tschirhart 1988).

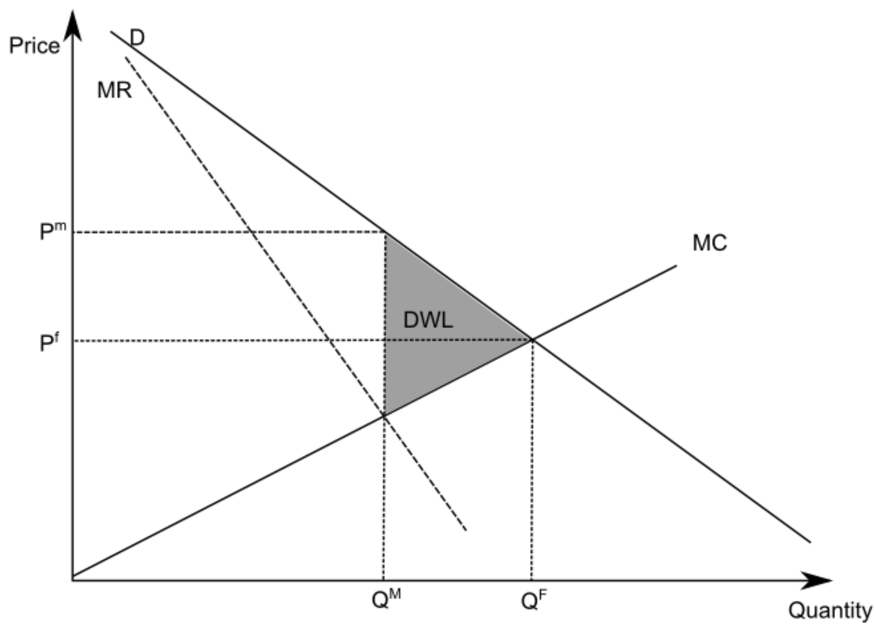

Figure 2-1: Free competition and monopolistic behavior

In a monopoly there are not sufficiently many producers rather only one producer who has the entire demand curve alone. The monopolist therefore must consider that the market price changes if offered quantum is changed. Figure 2-1 displays a monopolistic and a free competition adaption. In a free competition situation the market price (P^f) is decided from the demand (D) and marginal cost (MC). A monopolist will decrease quantum and set the price where demand equals marginal revenue (MR). The monopolistic adaption induces a dead weight loss (DWL) which is an economic loss, compared to the free competition. The DWL is illustrated by the shaded area in Figure 2-1. Production that is economically profitable is lost due to the monopolists’ market power. Profit is transferred from the consumer to the producer (Varian 2006).

2.1.2 Natural monopolies

The Norwegian electricity distribution business has a capital intensive cost structure with large initial costs (Grammeltvedt et al. 2006). Its price structure makes the industry a natural monopoly. The natural monopoly is recognized by two fundamental concepts (Berg & Tschirhart 1988). Firstly, the average total costs (ATC) are decreasing with output (Q) because of the high initial costs and a low marginal cost (MC) related to production. This means that the MC is lower than ATC over the interval considered[2] (Figure 2-2). The MC curve will intercept the ATC curve in its minimum, without exceptions. Secondly the product is delivered cheapest by a monopolist rather than having multiple producers. The cost function is said to be sub-additive[3]. The cost function must satisfy the following condition

$$C(Y) < \Sigma_i C(Y_i); \Sigma y_i = Y \; ; i = 1, 2, 3, \dots, k; k \geq 2; \qquad \text{Equation 1}$$

Which states that the single firm Y provides the service with smaller costs (C(Y)) compared to having multiple firms (ΣY_i) delivering the service at cost $\Sigma_i C(Y_i)$ (Waterson 1988).

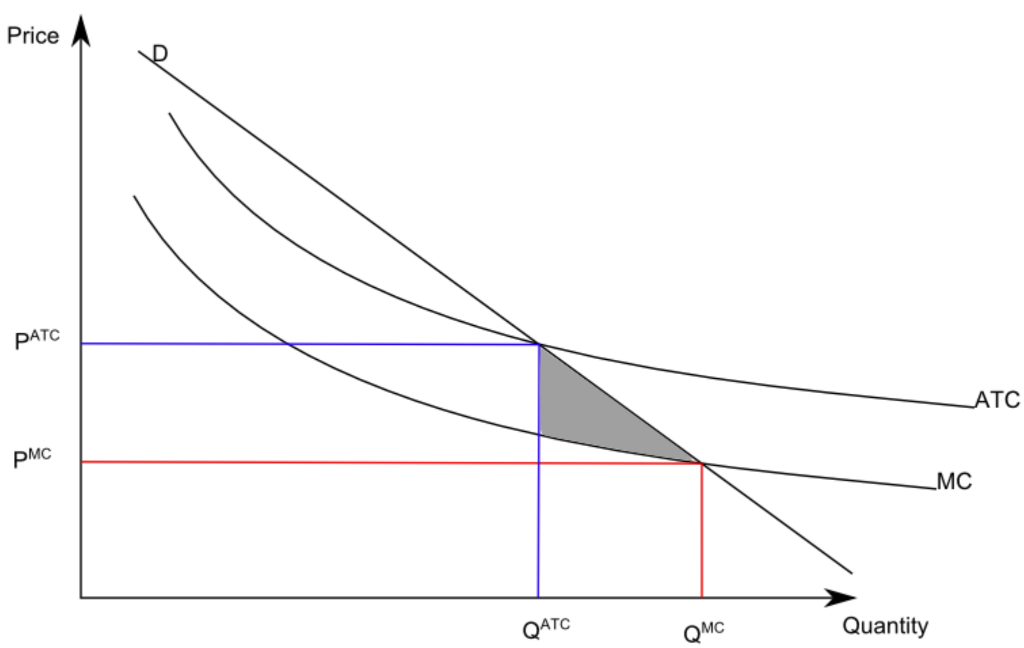

Figure 2-2: Natural monopoly.

[2] Interval considered is within the relevant production interval. As production approaches maximum capacity, marginal cost increase intercepting ATC in its minimum and dragging ATC up.

[3] Subadditivity is recognized by lower average costs as a result from three possible advantages; economies of scale, scale advantages and/or co-production rewards.

Figure 2-2 displays why demand (D) equal marginal cost (MC) cannot determine the price[4]. Pricing at this level will not cover all costs, since ATC > MC, so the deficit must be covered somehow, for instance by the community through taxes. This type of cost recovery is an unrealistic solution in most cases (Wangensteen 2012). Another solution to this problem would be to increase the price to equal the ATC. In this case the pricing will induce a loss in the social surplus, equal to the shaded area, an unwanted outcome. The effect is somewhat exaggerated in the figure. In fact the demand curve is much steeper in reality, i.e. the price elasticity in this business is smaller than illustrated.

2.1.3 Pricing in natural monopolies

This section will introduce the Norwegian industry price structure that seeks to solve the mentioned issue with DWL. By introducing the term non linear pricing and second best pricing it will be shown how the DWL is attempting to be removed.

2.1.4 The non-linear price structure

The Norwegian distribution industry solves the pricing challenges with a tariff with a non-linear price structure (Equation 2). The tariff is paid by the customers to the local distribution company to enter the electricity market. The customer can then choose from whom to buy their electricity in a competitive market. The tariff is established by the individual distribution company to recoup costs and must not exceed the yearly given revenue cap, which is specified by NVE (Grammeltvedt et al. 2006). The tariffs differ since the price is set by the individual distribution companies as a reflection of their costs. These differences are related to differences in organisational structure and costs. If the tariff does not reflect the costs, the distribution companies (DC's) are subject to adjust the next year's tariff, either up or down. This is controlled by NVE on a yearly basis (Grammeltvedt et al. 2006).

The non linear pricing rule is a two part tariff and involves a per unit price (price per kWh in this case) below average cost, plus a lump sum fee, which is of sufficient size to ensure break even (Waterson 1988). The tariff (P) is divided in two parts; one fixed amount (K), and one that varies

[4] The rectangular area formed by the axis and red lines is smaller than the area formed by the axis and the blue lines. This effect will only increase when apply a more correct demand curve, i.e. with smaller elasticity.

with consumption (q) with a fixed price (P_0) per kWh. In this case P_0 is equal to MC. Equation 2 shows the structure.

$$P = K + P_0 q \qquad \textbf{Equation 2}$$

In order to price the variable part of the tariff at marginal price one could set the price K equal or larger than the difference between average total cost and marginal cost, as shown in Equation 3.

$$K \geq (ATC - MC)Q^Y \qquad \textbf{Equation 3}$$

This way there is no DWL, but a transfer of surplus from customer to producer. This is illustrated in Figure 2-3.

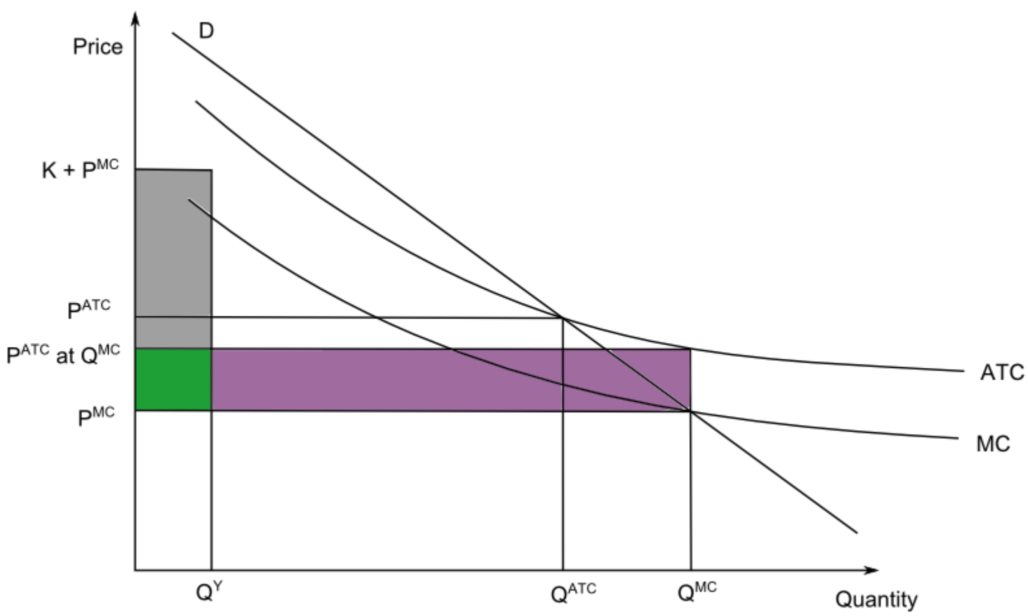

Figure 2-3: Non-linear pricing

The purple + green areas are losses from marginal cost pricing. The size of K is the sum of the green and grey area. When the losses from marginal cost pricing equals to the area of K, profit equals zero. Therefore, the non linear pricing, in this case known as second best pricing, maximizes the consumer surplus. If this is done successfully, the regulator has induced the firms to increase output, decrease price and produce at minimum cost, hence the alternative that serves the firm and society best under the given form of regulation as indicated in the introduction.

2.2 Efficiency measures

This section will outline the efficiency measures mentioned in the introduction. The real evaluation of performance differs from the standard microeconomic perspective as information on indifference curves and isoquants are missing. None of these elements are known initially, therefore in order to describe actual behavior of the companies this information must be estimated by collecting data and estimate the relation between inputs and outputs. How this can be done is evaluated in later chapters. One commonly used approach is known as the Farrell efficiency. In 1957 Farrell suggested that the economic efficiency of a firm consisted of technical efficiency and allocated efficiency (Coelli et al. 2005). Combining these two measures of efficiency yields the overall cost efficiency measure (Bjørndal et al. 2010).

2.2.1 Technical efficiency

A firm is technically efficient when it cannot produce more output with the given inputs or when it is impossible to produce the same amount of output by decreasing the inputs (Coelli et al. 2005).

To demonstrate how the technical efficiency can be measured a simple illustration is provided (Figure 2-4). The figure shows how the technical efficiency is calculated using an input-oriented method. Like Farrell, the illustration shows an example involving firms that use two inputs (x_1 and x_2), under the assumption of constant returns to scale. The input oriented model defines an isoquant, the curve (PP'), i.e. all the firms on this line produce the same output with different combination. This curve represents the efficient firms. In practice, knowledge of this curve must be estimated (Coelli et al. 2005).

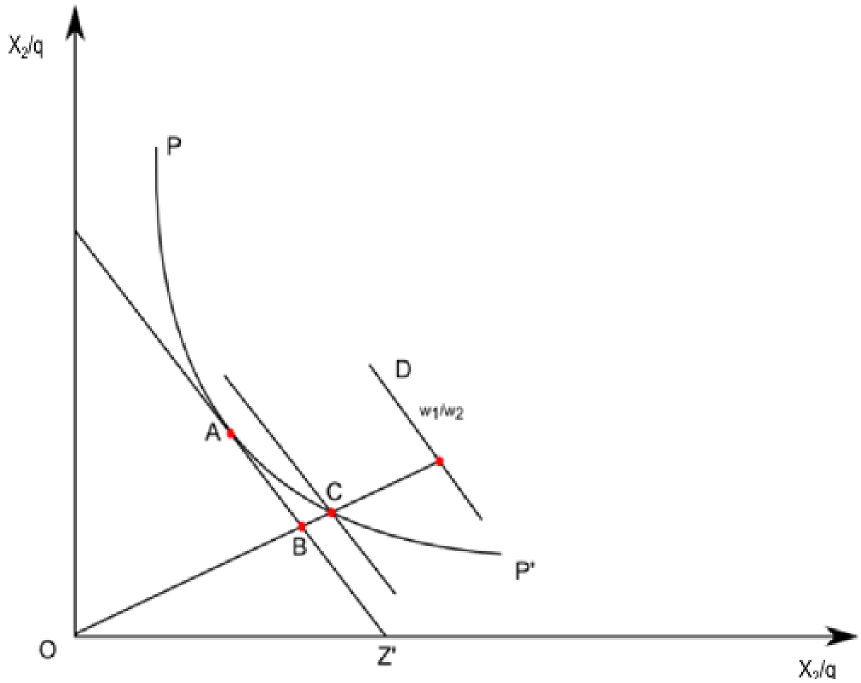

Figure 2-4: Technical and allocative efficiencies from an input orientation

The technical efficiency for the inefficient (the inefficiency is represented by the distance C-D) company D is measured by the ratio:

$$TE = OD/OC$$ **Equation 4**

TE < 1 implies that a firm is technically inefficient. Company A is the best performer in this setting because it produces at the technically and allocatively most efficient point (Coelli et al. 2005).

2.2.2 Allocative efficiency

If input price data is available, both allocative efficiency and cost efficiency are possible to measure. The isocost curve is recognized by the ratio on the input prices w_1 and w_2, i.e. the slope of the isocost curve.

A firm's capability of utilising the inputs in an optimal mix is known as allocative efficiency (Coelli et al. 2005). It measures the additional cost reduction by improving the input combination at given prices. For instance; the distance between point B and firm C in Figure 2-4 represents the production costs if the production was done similar to firm A. Firm C is technically efficient, but allocatively inefficient. Hence, firm C can produce the output at a lower cost with a different mix of inputs. Company D is neither technically nor allocative efficient. Point B is not a firm rather a point used for comparison. The allocative efficiency for firm C is calculated according to (Coelli et al. 2005):

$$AE = OB/OC$$ **Equation 5**

2.2.3 Cost efficiency

Firm A is cost efficient because it is able to choose the right mix of inputs and use them in a technically efficient manner. The firm must choose the right resources and use them correctly in terms of efficiency. The cost efficiency is expressed as the product of technical and allocative efficiency (Bjørndal et al. 2010):

$$CE = TE * AE = \frac{OD}{OC} * \frac{OB}{OC}$$ **Equation 6**

2.2.4 Scale efficiency

A firm may be both allocatively efficient and technically efficient but the scale of production could be improved. The production technology can have variable return to scale (VRS), increasing or decreasing, or constant returns to scale (CRS). Assuming VRS technology, a firm that is too large for its scale of operation may operate with decreasing return to scale and on the contrary if the firm is too small for its scale of operation it may experience increasing returns to scale. The scale efficiency can be improved if the firm succeeds in changing their scale of operation; hence it needs to change the size of operation keeping the same input mix (Coelli et al. 2005). The scale efficiency concept is illustrated in Figure 2-5.

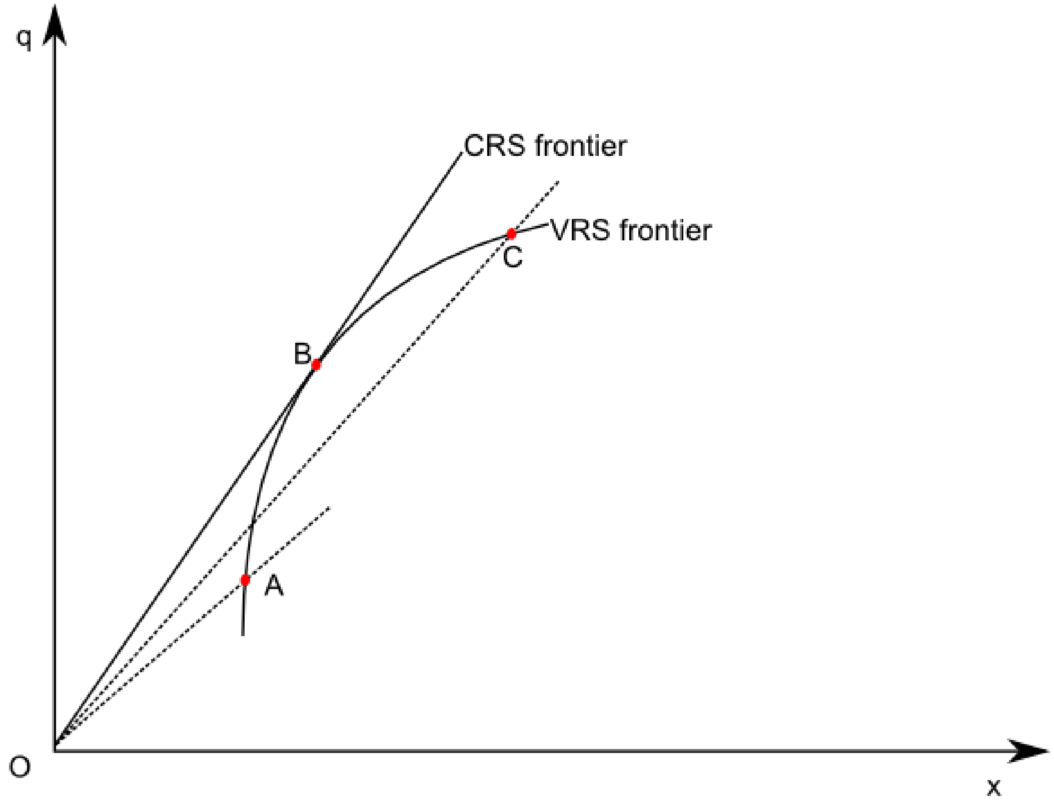

Figure 2-5: Scale of productivity

Firm A, B and C are all technically efficient as they are all represented on the production frontier. However, their productivity varies, determined by the slope of the dotted lines. Firm A could be more productive if it increased its production, i.e. firm A operates in the increasing returns to scale portion of the production frontier. Firm C is operating in the decreasing portion and could be more productive by reducing its scale of operation (Coelli et al. 2005).

2.3 Regulation regimes

This chapter will briefly describe the regulatory models used in the Norwegian distribution industry during the period of regulation. As introduced earlier the distribution industry is characterized as a natural monopoly with the need for regulation. However, the regulator lacks information on the industry's technology and costs. Therefore the regulatory problem can be seen as a game between the regulator and the regulated firms (Bogetoft & Otto 2011). One instrument used in the Norwegian regulatory history is benchmarking. Regulation and benchmarking have traditionally been separated, but the last 15 years they have been attempted integrated (Bogetoft & Otto 2011). More on this topic in Bogetoft and Otto (2011).

2.3.1 Rate of return regulation 1993 – 1996

In the first period of regulation the model used was based on the rate of return principle. For each firm a rate of return was determined on the basis of their financial statements. NVE performed revision every year to check if the companies stayed within the allowed limit. If the return turned out to be larger than the limit, adjustments were made with the next year's tariff, hence ex ante. If the return was smaller than the allowed and wanted rate, this could be corrected for by charging a higher tariff the year after. However, this regulation model was only used while searching and developing a better strategy. From economic theory it is known that firms under ROR regulation have a tendency to favour the use of capital since the use of this input can increase profits alone, giving small or no incentives for efficient production (Train 1991). This is known as the Averch Johnson effect (A-J effect).

Averch and Johnsen's work from 1962, "Behavoir of the regulated firm", shows that a regulation model based on Rate of Return (ROR) has two important conclusions (Train 1991 p.40):

Firstly,

- *"The regulated firm uses more capital than the unregulated firm"*

And secondly,

- *"The capital/labor ratio of the regulated firm is inefficiently high for its level of output. That is, the output that the regulated firm produces could be more cheaply produced with less capital and more labor than the regulated firm chooses".*

To conclude, in this period the regulation method used was temporarily and induced inefficiency amongst the distribution industries utilisation of inputs. The inputs where utilised effectively but the input combination was not. During 1996 a new model was developed based on revenue regulation.

2.3.2 Revenue cap regulation, 1997 – 2001

The new model in this five year period used reported values from the financial statements in the years 1994 and 1995 to set the allowed revenue. The revenue cap was annually adjusted for inflation, changes in energy prices (related to power loss) and increase in delivered energy. In addition a general and individual efficiency requirement was introduced. The efficiency scores were based on results from a DEA analysis. The general efficiency improvement requirement was 1.5 percent. In 1998 an individual efficiency improvement of 0-3% was included. In 2001 a quality mechanism was included in the regulation. “In order to avoid deteriorations of the quality *Compensation for Energy Not Supplied (CENS)* has been introduced in the Norwegian grid regulation” (Wangensteen 2012, p.330). CENS was established to improve incentives to minimise customer’s costs related to grid failure as a result from under investments and cost reductions (Grammeltvedt et al. 2006).

2.3.3 Revenue cap regulation, 2002 – 2006

This period’s model had minor adjustments made compared to the earlier model. The starting point for the revenue cap was updated with accounting values from 1996 to 1999. The reference rent was changed from fix to floating. Like in the previous period, adjustments for inflation and electricity price were made. The general requirement for efficiency improvements of 1.5 percent was withheld, but the individual efficiency requirement increased, with a maximum at 5.2 percent. The efficiency analysis was based on the DEA analysis, as in last period. It was concluded that the CENS solution did not fulfil its purpose as intended; the delivery reliability was valued too much. This problem had to be sorted out with the next model, introduced in 2007 (Grammeltvedt et al. 2006).

2.3.4 Revenue cap regulation 2007 – 2011(12) (in more detail)

Today's regulation model, introduced January 1. 2007, is a continuation of the earlier revenue cap. However, some major revisions have been implemented. The time from new investments associated depreciation and interest were accounted for in the total cost was too long (Bjørndal et al. 2010). On the one hand the time lag gave strong incentives for efficiency improvements, but on the other hand there are losses related to the time from the investments affecting the net present value (Grammeltvedt et al. 2006). Therefore the revenue cap was annually updated with the latest available financial year accounts. This means that the revenue cap for 2012 is decided from accounted values in 2010 (Grammeltvedt et al. 2006).

Unlike earlier models today's model is determined from a weighted average of the firm's true costs and a benchmark's normalized costs. The true costs are weighed 40% and the normalised costs are weighted 60%. The term normalised costs will be defined later. Which in turn give a revenue cap that is more consistent with the costs of the year the revenue cap is prevailing compared to earlier.

Equation 7 shows today's revenue cap (RC_t):

$$RC_t = (1-\rho)C_{t-2} + \rho C^*_{t-2}$$ **Equation 7**

Where:

C_t = the firms reported costs in eRapp, further investigated below.

C_t^* = normalized costs, cost of the firm(s) that form the frontier.

ρ = level of normalised costs included in the revenue cap, sat by NVE to 0.6 in 2009, earlier models had a ρ =0.5.

A more restrictive regulation would have ρ closer to 1. As ρ approach 1, the regulator ignores the firm's own costs and allows the firms revenue equal no more than the costs of the most efficient company (Bergland 2011c).

The normalized costs (C_{t-2}^*), also referred to as the cost norm, are calculated from a comparison analysis done by NVE to identify the distribution business' true costs. The DEA analysis is further investigated in chapter 2.5. Since the revenue cap is given ex-ante and the costs are not possible to know up front, NVE use costs lagged 2 years in order to set the normalised costs (Grammeltvedt et al. 2006).

2.3.5 Criticism from the industry of today's model

NVE has a close dialog with the distribution companies concerning changes in the regulation model. In November 2011 NVE sent the firms a letter with possible changes in the cost norm and received 11 answers (Lundteigen 2011). I will not present all critics raised, but a review of the most important points.

- Correction for age

  One of the main criticisms that is pointed out relates to the fact that there is no correction for the age of the grid. This has been discussed with earlier models as well and NVE suggests that instead of using total costs as a base for the cost norm, one solution could be book keeping value. The industry does not agree and argues that the effect will not be sufficient. The work on finding a solution to this problem must therefore continue (Lundteigen 2011). The cost structure of the industry is capital demanding, therefore the quality of the efficiency analysis correlated to how the capital expenses are measured and included. According to Bjørndal et al. (2010) using book values may lead to a negative bias in the efficiency score. A new grid would have a higher book value and depreciations than an old grid would. Two companies with an equal amount of capital assets, identical maintenance and operation cost get a different DEA efficiency score (Lundteigen 2011). More on this topic can be found in Bjørndal et al. (2010) and (Lundteigen 2011).

- Black box

  The DEA model is referred to as a black box that is filled with some data and out comes the efficiency measures. Since it is difficult to understand how the efficiency score is decided, it is difficult to understand how a company has become ineffective. There are no obvious signals that tell how one can improve the efficiency score (Lundteigen 2011).

- Slack

  The DEA model gives different investment incentives for firms with and without slack, what is meant with slack is explained in chapter 2.5.2 (Lundteigen 2011).

- Small companies

  Companies that are special and best within one of the outputs way get an efficiency score equal 100 % just because of this one output (Lundteigen 2011).

- Mergers

  Companies that merge will in most cases come worse out with a lower efficiency score than without merging. This could lead to that social welfare enhancing mergers are left undone (Lundteigen 2011).

2.4 Cost frontier

By considering the cost efficiency as the product of two components, technical efficiency and allocative efficiency, theory on how to estimate the cost frontier will be presented in this chapter. The cost frontier can be estimated using two different methods (amongst others), a non parametric method and a parametric method.

Parametric models estimate parameters from prior data. The frontier is estimated with econometric methods assuming a functional form for the relationship between the inputs and an output. When the functional form is chosen, the parameters can be estimated using econometric techniques (Coelli et al. 2005).

Non parametric models have fewer restrictions a priori, often simply a fundamental mathematical assumption taken in order to estimate the production activity. One such fundamental assumption in the Norwegian DEA model is the assumption on whether to use constant returns to scale or variable returns to scale (Econ 2008). This is further explained below in chapter 2.5.

2.4.1 Yardstick regulation

Yardstick regulation (competition) uses a benchmark to compare the distributors, in this case in terms of cost efficiency. The main purpose of yardstick competition is to create competition between distributors. Competition is created by measuring firms (dealing in the same market) performances relative to their competitors, given the same geographic area and time (Shleifer 1985). The yardstick competition method is attractive since the firms are measured against other firms, rather than their own past cost (Bogetoft & Otto 2011).

The simulated competition should create incentives for firms to increase their efficiency. The Norwegian regulator use the best performers cost as a benchmark and compare all the firms relative to this. This way the regulated firms could increase their efficiency by making decisions making them better performers.

In spite of the “game” between the regulated firms and the regulator the Norwegian regulator has a close dialog with the industry. NVE takes into account the firms insights on the regulation model when introducing changes, through written submissions (Lundteigen 2011). According to Shleifer (1985) the regulator must commit himself not to pay attention to the firms’ complaints and to be prepared to let the firms go bankrupt if they choose inefficient cost levels. As of today's regulation this is not the case since the industry is protected by a minimum profit of at least 2% (Grammeltvedt et al. 2006).

The yardstick competition will be expanded through the yardstick based Data Envelopment Analysis (DEA) in chapter 2.5. There it will be shown that there are some challenges related to this way of benchmarking. One challenge is related to comparing the different companies. The regulator can in no way make simple calculations to compare the competitor's costs. Therefore it is necessary to obtain useful tools to calculate the competitor`s costs (Grammeltvedt et al. 2006).

2.5 Data Envelopment Analysis (DEA)

Data Envelopment analysis (DEA) is a non parametric performance method for comparing the relative performance of different utilities with more dimensions[5] (Coelli et al. 2005). DEA use linear programming for solving an optimisation problem. From this problem an efficient frontier is estimated and an efficiency measure is calculated relative to the front. Since Farell introduced the method in 1957 it has been extended and developed and it was first in 1978 that the term, data envelopment was first used (Coelli et al. 2005).

The DEA-algorithms can be quite complex, but the underlying theory of the method is easy to understand. The following chapters present an example using two dimensions. A model with more than two dimensions needs advanced mathematics. A two dimension method can be transferred to a more complex model.

Solving these kinds of problems is done using different models. Here there are two methods, one output oriented model where inputs are held fixed and one input oriented, where output is held fixed.

In the models used by NVE inputs are measured in terms of total costs. By using total costs as the dependent variable, differences amongst the firms in input mix and input prices are ignored (Grammeltvedt et al. 2006).

An assumption regarding return to scale must be chosen before making the calculations. The two methods above calculate the same result for inefficiency using CRS, but with different results using VRS. The reason for this will be explained below. An input oriented model assuming constant returns to scale (CRS) and one with variable returns to scale (VRS) will be presented and then compared.

2.5.1 Solving differences in scale, input oriented model

In order to analyse the costs and to determine which companies that are cost efficient, a cost function is estimated. The way the optimisation problem in DEA is specified, as with basic cost function properties, the costs will always increase with the increasing input (in special cases they can remain constant). This implies that the larger the company is, the larger the compared

[5] Dimensions: More than one output and or input at the same time.

reference costs. Which reference point to be compared too is decided on the individual firms input combinations. The efficiency score is crucially dependent on which scale properties are chosen when analysing (Grammeltvedt et al. 2006).

Under the assumption of CRS solving the comparison problem is done by setting up a simple linear programming problem like Equation 8.

$$\begin{aligned} &\min_{\theta,\lambda} \theta \\ &st. -\boldsymbol{q}_i + \boldsymbol{Q\lambda} \geq 0, \\ &\theta \boldsymbol{x_i} - \boldsymbol{X\lambda} \geq 0 \\ &\boldsymbol{\lambda} \geq 0 \end{aligned}$$

Equation 8

Where $\mathbf{x_i}$ and $\mathbf{q_i}$ are column vectors for the i-th firm's inputs and outputs, respectively **X** is the NxI input matrix and **Q** is the MxI output matrix, representing the data from the observed companies. θ is a scalar and the obtained efficiency score of firm i. If a company obtains a value of θ equal to 1, this firm is on the efficient frontier, assuming slack is ignored. The term slack is defined in chapter 2.5.2. **λ** is a **Ix1** vector of weights, the weights tells us how much of each company that is included in the cost frontier (Grammeltvedt et al. 2006). The linear problem must be solved as many times as there are firms for all the firms to obtain their individual efficiency score θ (Coelli et al. 2005).

The problem can be transformed into a model assuming VRS by adding the assumption on convexity to Equation 8, as shown in Equation 9 (Coelli et al. 2005).

$$\begin{aligned} &\min_{\theta,\lambda} \theta \\ &st. -\boldsymbol{q_i} + \boldsymbol{Q\lambda} \geq 0, \\ &\theta \boldsymbol{x}_i - \boldsymbol{X\lambda} \geq 0 \\ &\boldsymbol{I1}'\boldsymbol{\lambda} = 1 \\ &\boldsymbol{\lambda} \geq \boldsymbol{0} \end{aligned}$$

Equation 9

Where **I1** is an Ix1 vector of ones, “this constraint ensures that an inefficient firm is only benchmarked against a firm of a similar size” (Coelli et al. 2005, p.172). This restriction makes sure that all firms are benchmarked against other firms which are smaller and larger regarded their dimensions[6]. The individual firm's reference point is a weighted average by the closest effective companies is larger than itself, λ is the weights in the weighted average (Grammeltvedt et al. 2006).

Figure 2-6 shows the input oriented model with input, total costs, and the output, km of grid. When assuming CRS the frontier is found by drawing a line from the origin and through the company with the lowest unit cost. In this case this is firm 2. Assuming VRS firm D1 – D3 form the frontier by drawing a straight line between the three. D6 is smaller than D3 but larger than D4, therefore shape D3 and D4 a reference companies for D6, the imaginary company D6*. This shows that D6 is compared to the companies that have the closest cost structure as itself. The cost represented in D6* are the costs that D6 could have if it was efficient. The efficiency is measured as the ratio between the two lines OD6-D6* and OD6-D6.

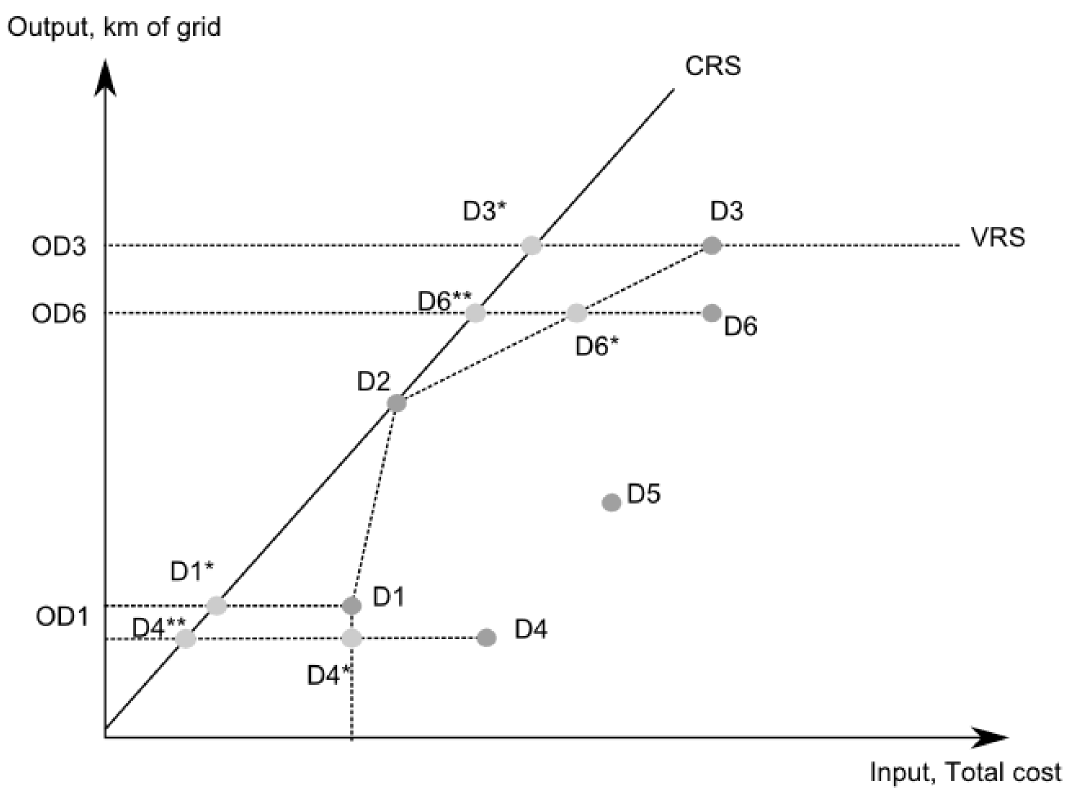

Figure 2-6: Input oriented DEA analysis with CRS and VRS.

[6] Except where the firms itself is largest in a dimension.

D4's reference company D4* is on the frontier where slack arise. D4 can D4* has the same amount of km grid as D1 but at lower cost. Therefore D1 must be more efficient than D4*. This means that D4 can increase its output without decreasing its cost efficiency, which is the background for why companies which are small in one output does it very well in a VRS model. On the other side, large companies are measured as efficient because of their size. D3 is only efficient because of its size. If D3 was removed, D6 would be efficient simply because there is no observed larger company. The scale effects are not present when assuming CRS since this method use the same unit cost for comparison, in this case the unit cost of D2. The inefficient companies under VRS are still inefficient under CRS. The general efficiency score obtained under VRS would always be larger or equal the one obtained under CRS. It is important to note that this is not the same as saying that the company is more effective, it is simply a different measure (Grammeltvedt et al. 2006).

2.5.2 Solving structural differences, output oriented model

Technical inefficiency can also be calculated as an increase in output holding the input constant. This model is used in industries where the amount of inputs is fixed and the production goal is to maximise output with these given inputs. The maximizing problem, shown in Equation 10 is quite similar the one used in an input oriented model, but instead of minimising inputs one wishes to maximise output.

$$\begin{aligned} &\max_{\phi,\lambda} \phi \\ &st. -\phi \boldsymbol{q_i} + \boldsymbol{Q\lambda} \geq 0, \\ &\boldsymbol{I1'\lambda} = 1 \\ &\boldsymbol{\lambda} \geq \boldsymbol{0} \end{aligned}$$

Equation 10

Where: $1/\varphi$ is the technical efficiency score, a scalar between 0 and 1.

Figure 2-7 shows the output oriented method with two outputs, km of grid and number of customers. D1 – D6 represent different distribution companies with unequal combinations of customers and km of grid lines. The companies D1 – D3 are cost efficient companies that the other companies are compared with. Between the cost efficient companies a line is drawn making the efficient frontier.

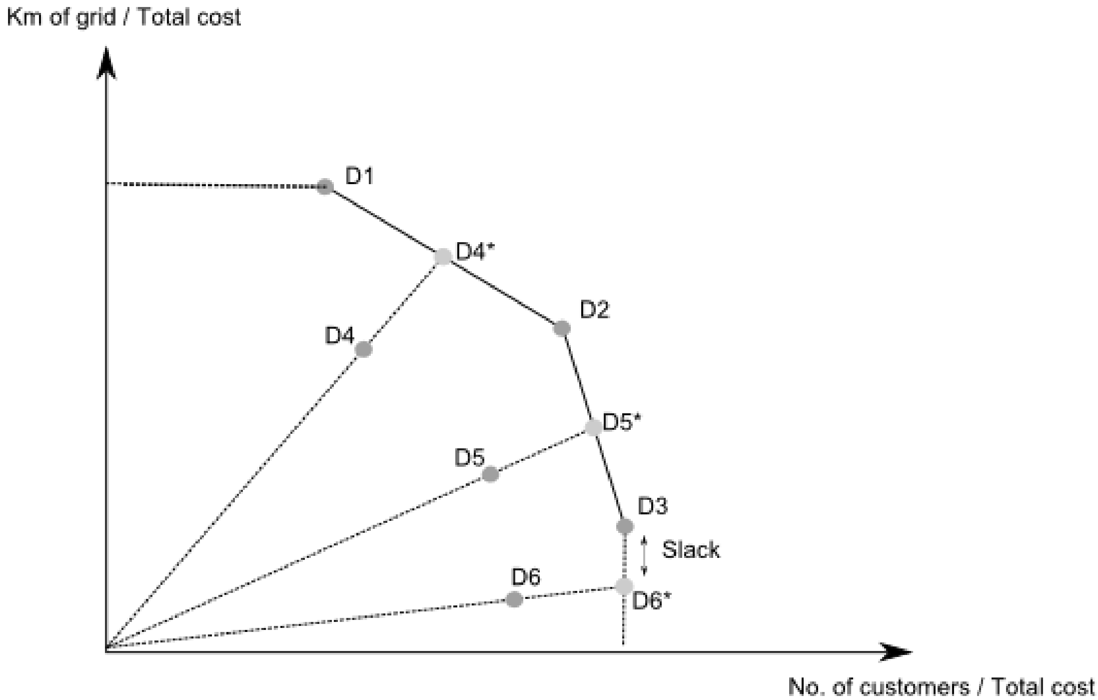

Figure 2-7: Output oriented DEA model

Company D1 is a company with few customers per km grid and can illustrate a company in a rural part of Norway, whilst company D3 has many customers per km grid can illustrate a company serving a city. Company D4 – D6 are all inefficient. By drawing a straight line from the origin to the frontier through the company one finds the reference point for each company, as done in Figure 2-7. Company D1 and D2 are reference companies for firm D4, because, as seen from Figure 2-7, D4 is placed between these two companies. D1 is a company with fewer customers per km grid and D2 is a company with more customers per grid making these the companies that D4 should be compared with. Likewise, are company D2 and D3 reference firms for D5 (Grammeltvedt et al. 2006).

Overall this shows that companies with different customer density have different reference points and that these reference points are made up by companies that have the output combination that is most similar to their own, all independent of the size of the company. Under the given assumptions the cost efficiency score is the measured ratio between the lines OD4 and OD4* as shown in Equation 11.

$$Efficinecy = \overline{OD4}/\overline{OD4^*}$$ **Equation 11**

For company D6 the measure is somewhat different. Even though company D6 had improved its efficiency to be D6*(on the frontier) it would still have the potential to improve its performance (this is called slack). Increasing the output km of a grid (holding the number of customers constant) would not make the efficiency score any better. It would still be on the frontier and hence improving one output would not relate in an improved efficiency score. This slack arises because the efficiency score is incorrectly measured in the first time, it is measured too large (Grammeltvedt et al. 2006).

NVE discuss different possibilities to cope with this problem but these solutions arises new challenges. As far as I can see, slack is solely a discussed theme and not handled(ignored) in today’s regulation model.

2.5.3 Super efficiency

With the above methods all the efficient companies making the frontier gets an efficiency score equal 1. To evaluate the efficient firms against each other and to prove super efficiency (efficiency score above 1) NVE removes the efficient firms to make a new frontier (Grammeltvedt et al. 2006). The superefficient score is set relative to this new frontier, as illustrated in Figure 2-8.

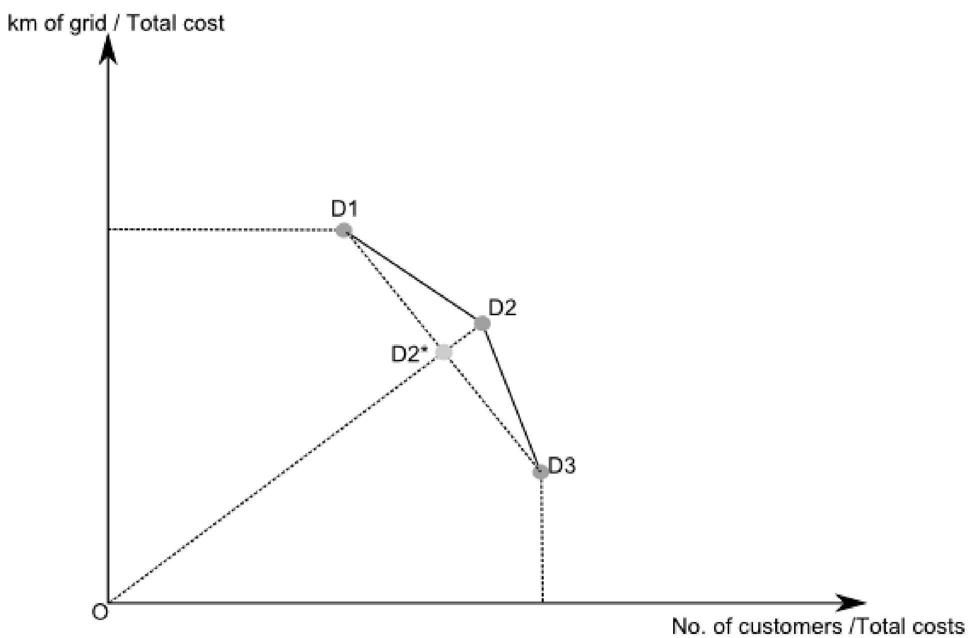

Figure 2-8: DEA output oriented model, super efficiency.

The distribution companies D1 – D3 form the CRS efficient frontier in the output oriented model above. By removing D2 a new front appears and D2* is the reference point for D2. The super efficiency score is calculated by taking the line segment O-D2 divided by O-D2*, which obviously is greater than 1.

This method may give companies a too a high score based on false conditions, like extreme results in one dimension, and must be handled with care. NVE does not use the super efficiency score without corrections. The score may be too large simply related to lack of comparable firms. Therefore NVE decided to compare the super efficient firms with previous year’s observation.

With this correction super efficient firms are rewarded if they improve compared to last year's measure (Grammeltvedt et al. 2006).

2.5.4 Environmental factors in DEA

Environmental factors are factors not controllable by the manager that influence the efficiency score (Coelli et al. 2005). These factors are related to costs, but not directly observable. Costs related to wind, snow and forest are examples of factors in the DEA model (Grammeltvedt et al. 2006).

According to Coelli et al. (2005) there are a number of different methods used to include such factors in an efficiency analysis. NVE suggests two of these methods for solving these challenges. The first includes the environmental factor as any other parameter, directly in the model. The second method estimates the efficiency score without the environmental factor and then analyse how much of the inefficiency is related to the factors. Additional details are available in standard textbooks such as Coelli et al. (2005).

2.6 Stochastic frontier analysis (SFA)

Stochastic frontier analysis is a parametric method for estimating efficiency. The estimation method is underpinned the same assumptions as mentioned in relation to POLS in appendix C. This makes it possible to assume a stochastic relationship between the used inputs and produced outputs. One of the main differences between DEA and SFA is that the SFA regression model distinguishes between statistical noise and technical inefficiency. This is done by estimating a function with two random variables, one to account for the statistical noise and the other for technical inefficiency, shown in Equation 12. Statistical noise arises if relevant variables are omitted as well as measurement errors as well as errors connected to choice of functional form (Coelli et al. 2005).

Treating the total costs (C) as the only input (as in the output oriented DEA model), a function of the produced quantity (**x**) is illustrated in Equation 12.

$$lnC = \boldsymbol{x}_i\boldsymbol{\beta} + \mathrm{v_i} + \mathrm{u_i} \qquad \textbf{Equation 12}$$

Where $\mathrm{v_i}$ is the variable associated with statistical noise and $\mathrm{u_i}$ is a non negative random variable associated with the technical inefficiency. In order to estimate the parameter's ($\boldsymbol{\beta}$) of the cost function in SFA one first needs to make an assumption on the functional form. Two widely used methods are the translog and the Cobb-Douglas functional forms. These functional forms are presented in Table 2-1.

Table 2-1: Cobb-Douglas and translog, functional forms (Coelli et al. 2005).

Cobb-Douglas	$y = \beta_0 \prod_{n=1}^{N} x_n^{\beta_n}$
Translog	$y = \exp\,(\beta_0 + \sum_{n=1}^{N} \beta_n ln x_n + \frac{1}{2} \sum_{n=1}^{N} \sum_{m=1}^{N} \beta_{nm} ln x_n ln x_m)$

According to Coelli et al. (2005, p.211-212) does the preferred models hold some of the following characteristics.

- ***Flexible***. *"A functional form is said to be first order flexible if it has enough parameters to provide a first-order differential approximation to an arbitrary function at a single point*[7]*. A second order flexible form has enough parameters to provide a second order approximation. The Cobb-Douglas form is first order flexible, while the translog functional form is second order flexible. All other things being equal, we usually prefer functional forms that are second-order flexible. However, increased flexibility comes with a cost – there are more parameters to estimate, and this may give rise to econometric difficulties (eg., multicollinearity*)" The issue is further discussed in chapter 3.1 on model specification.

- **Linear in the parameters**. Both translog and the Cobb-Douglas are linear in the parameters. This is necessary for estimation using the linear regression. "*At first glanc, the Cobb-Douglas and translog functions appear not to satisfy this property. However, taking the logarithms of both sides of these functions yields linearity*".

- **Parsimonious**. *"The principle of parsimony says we should choose the simplest functional form that "gets the job done adequately". Sometimes we can assess the adequacy of a functional form prior to estimation. For example, the Cobb-Douglas function is inadequate in situations where elasticities may vary across data points, and both the Cobb-Douglas and translog functions are problematic when the data contain zeros because this makes it impossible to construct the logarithms of the variables. However, model adequacy is often determined after estimation by conducting a residual analysis (i.e. assessing whether residuals exhibit any systematic patterns that are indicative of poorly chosen function), hypothesis testing, calculating measures of goodness-of-fit and assessing predictive performance".*

[7] The phrase n-th order differential approximation to an arbitrary function at a single point means it is possible to choose values of the parameters so that the value of the approximating function and all its derivatives up to order n are equal to those of the arbitrary function at that point.

SFA utilises observations from the different firms to estimate the cost function. From this estimated function, the efficiency measures are calculated. Hence, the unknown parameters of Equation 13 are estimated using actual observations. One method for finding these estimates is the maximum likelihood principle. This method estimates β's that explain the actual observations as likely as possible (Bogetoft & Otto 2011). More on the maximum likelihood method can be found in Coelli et al. (2005).

The statistical noise can arise from effects as weather, strikes, luck etc. on the value of the output variable. "However, these effects have less to do with our statistical models than with the risky environment in which production takes place" (Coelli et al. 2005, p.243). Methods dealing with risk are not handled in this thesis, more on this subject is found in Coelli et al. (2005). The random error v_i can be positive or negative as illustrated in Figure 2-9. This illustration use, as indicated in Equation 13, total costs as the dependent variable and one output, the actual model has more outputs, but this is not easily illustrated. If functional form is assumed to be a Cobb-Douglas stochastic frontier model it would take the form in Equation 13.

$$C_i = \exp\,(\beta_0 + \beta_1 lnx_i + \ v_i + u_i) \qquad \textbf{Equation 13}$$

Where C_i is the output, total cost, $\exp\,(\beta_0+\beta_1 \ln x_i)$ is the deterministic component forming the frontier, $\exp(v_i)$ is noise and $\exp(u_i)$ is the inefficiency term[8]. The noise can be both positive and negative.

Figure 2-9 shows the plotted inputs and outputs of two different firms, A and B indicated with grey dots. At the cost level C_A, firm A has an output level X_A and likewise for firm B, at cost level C_B follows output level X_B. If there were no inefficiency effects, hence $u_A=0$ and $u_B=0$ the output would only include noise indicated by C_A^* and C_A^*, also indicated in Equation 14 . The plotted values for firm A and B with no inefficiency are indicated with red dots.

$$C_A^* = \exp(\beta_0 + \ \beta_1 lnx_A + v_A)$$

Equation 14

$$C_B^* = \exp\,(\beta_0 + \beta_1 lnx_B + \ v_B)$$

[8] Exp= Exponential.

By comparing the individual firms two plots (e.g. C_A and C^*_A) the technical efficiency score is calculated, as of Equation 15. As illustrated in Figure 2-9, firm A has a positive noise effect and firm B a negative noise effect. One could say that B has had more influential episodes affecting their cost than firm A.

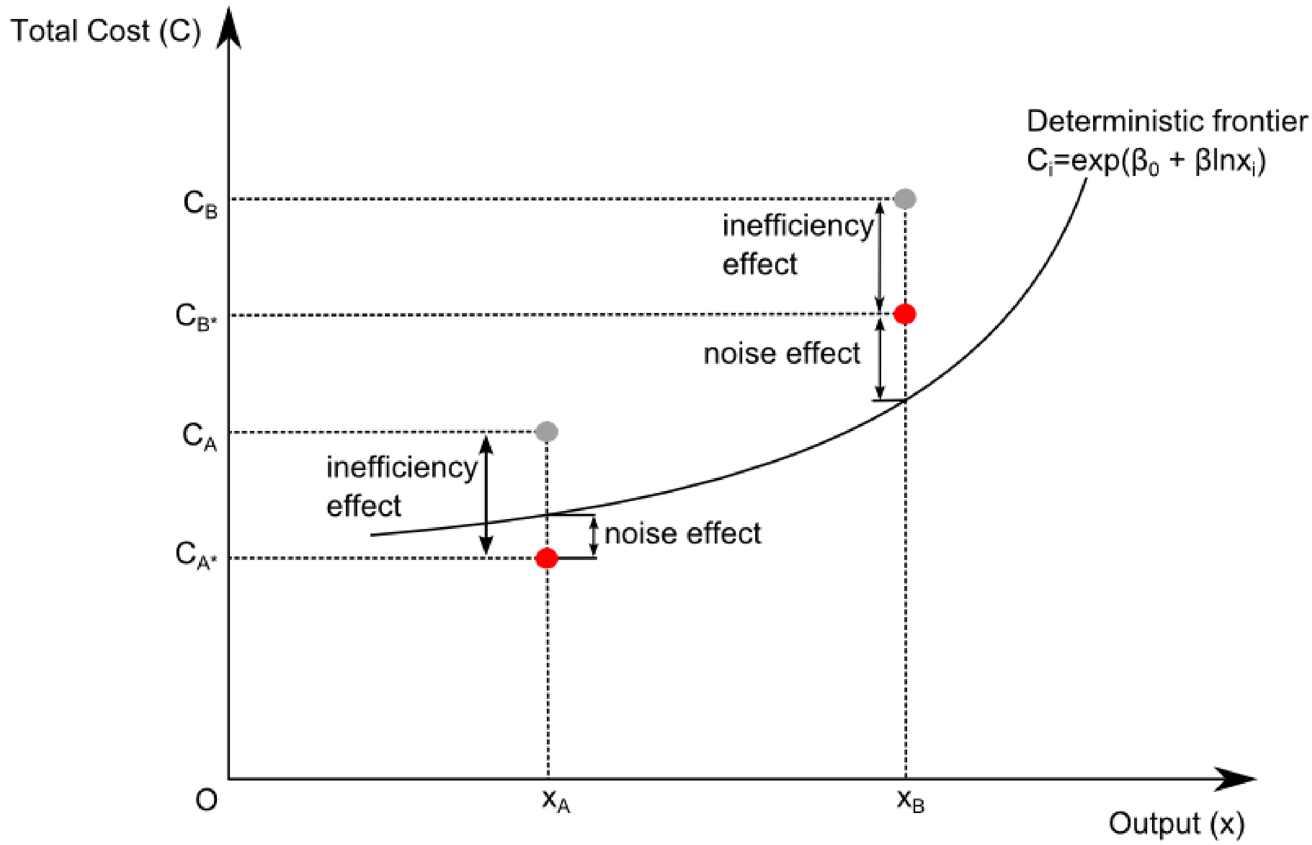

Figure 2-9: The Stochastic Cost Frontier

$$TE_i = \frac{Q_i}{\exp(\boldsymbol{x_i\beta} + v_i)} = \frac{\exp(\boldsymbol{x_i\beta} + v_i - u_i)}{\exp(\boldsymbol{x_i\beta} + v_i)} = \exp(-u_i) \qquad \textbf{Equation 15}$$

TE is the (i:th) individual firm's technical efficiency scores, a value between 0 and 1. Obviously the first step to determine the efficiency measure is by solving Equation 13.

2.6.1 Estimating the parameters

As with pooled ordinary least squared (POLS) regression the stochastic frontier estimation is underpinned by some assumptions. These assumptions are outlined in Appendix C in relation with the maximum likelihood method. The regression of the stochastic frontier is more complicated than a POLS, due to the fact that there are two random terms to estimate, the noise and the inefficiency. Both the noise and inefficiency components are assumed to have identical properties to the noise in a classical linear regression model. However, the inefficiency is said to be a half normal model and assumed to have a non-zero mean. This is because the inefficiency is always larger or equal to zero (Coelli et al. 2005).

2.6.2 The half normal model

The statistical noise, $v_{it,}$ is assumed to have a symmetric distribution, $v_{it} \sim iidN(0,\sigma^2_v)$ the inefficiency, u_{it}, is assumed to have a strictly non-negative distribution, $u_{it} \sim iidN^+(0,\sigma^2_u)$. Each u_i is determined by a probability density function (pdf). Figure 2-10 illustrates three examples of what this pdf could look like.

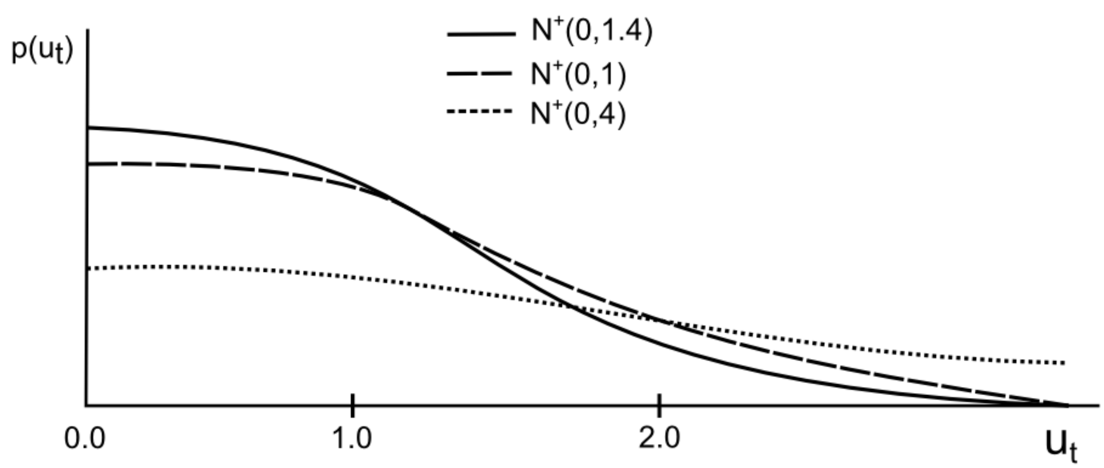

Figure 2-10: Half-Normal distributions

In order to understand how the two variables are determined it is necessary to know how their variances. Assume $\varepsilon = u+v$, hence ε is the total residual. By determining σ_ε^2 (variance of the residual) one can determine if the distribution is a normal distribution or a truncated normal distribution. If the distribution of ε looks like the distribution of u, the distribution of u dominates v and the other way around, if the distribution ε looks like the distribution of v, the distribution of u dominates v (Bogetoft & Otto 2011).

2.6.3 Technical change

Observations over time usually include a time trend to account for technological change (Coelli et al. 2005). The functional form chosen decides the nature of this periods technology change. In a Cobb-Douglas function this change is assumed to be constant and convex, in a translog function this trend can increase or decrease with time. The time trend should be included to allow some of the slope coefficients (**β**) to change over time and reflect the industry's knowledge about the technology behavior. In a translog cost function this done by including the t^2 (as opposed to a C-D function that only include t) in the model (Equation 16).

$$lnC = \theta_1 t + \theta_2 t^2 + \boldsymbol{x}_i \boldsymbol{\beta} + \mathrm{v}_\mathrm{i} + \mathrm{u}_\mathrm{i}$$ **Equation 16**

θ_1 and θ_2 are the unknown parameter to estimate. The percentage change is given by the first order derivative of lnC with respect to t, indicated in Equation 17.

$$\frac{\partial lnC}{\partial t} = \theta_1 + 2\theta_2 t$$ **Equation 17**

θ_1 and θ_2 tell whether or not there has been a technological improvement over the time period looked at (Coelli et al. 2005).

2.6.4 Technical efficiency change

Panel data provides the opportunity to calculate estimates of technological efficiencies (Coelli et al. 2005). Over time hopefully the inefficient companies will improve their efficiency level and the efficient firms stay efficient, all other equal. In order to decide if this is the case, some structure on the inefficiency must be introduced (Coelli et al. 2005). One such parameterization is a time invariant model where the inefficiency is assumed to have a truncated-normal distribution. The other is a time variant model. The time variant model is assumed to have a truncated-normal distribution multiplied by a specific function of time (*xt-frontier - Stochastic frontier models for panel data* 2012).

One example of a time varying model assumes that the technical inefficiency develops according to a function is the one developed by Battese and Coelli (Coelli et al. 2005).

The inefficiency term u_{it} can follow the function in Equation 21.

$$u_{it} = f(t)u_i \quad \textbf{Equation 18}$$

Where f(t) is the function that describes the variation in the technological inefficiency over time. The function f(t) is modeled as in Equation 19.

$$f(t) = exp[\eta(t - T)] \quad \textbf{Equation 19}$$

Eta (η) is the inefficiency parameter to estimate. The sign of η tells us if the inefficiency increases or decreases. Figure 2-11 have replicated possible functions for the efficiency development, (Coelli et al. 2005, p.278). Either eta is negative or positive, but always constant and convex.

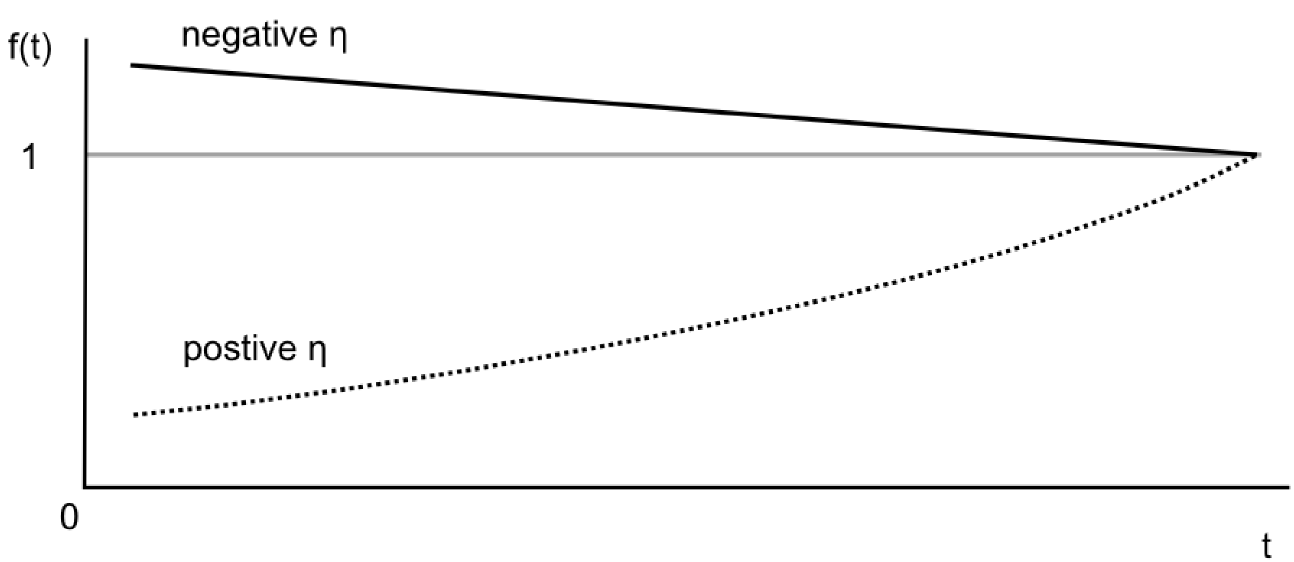

Figure 2-11: Functions for time-varying efficiency models

3 Results

This chapter presents an analysis of cost efficiency by estimating the respective frontier. As mentioned in chapter 2.3.5, criticism has been raised towards NVE's method and the interpretation of the analysis' results. By using data reported by the Norwegian distribution firms collected by NVE in the years 2007-2010 an alternative method to the DEA is presented. The alternative method is a parametric method using econometric theory to establish the cost frontier. The frontier is estimated using the statistical package STATA 11.1, accompanied by Microsoft Excel 2007. In the process of establishing such a frontier it is necessary to decide which outputs to use. As opposed to the theoretic one-input one-output models in chapter 2.6, there are several dimensions in both inputs and outputs. Therefore the frontier is thought of as a multidimensional plane rather than a line (Wangensteen 2012). The cost frontier is estimated using total costs as the dependent variable and three different outputs as the explanatory variables, all of which are reported to NVE by the distributing companies on a yearly basis. The data is strongly balanced, i.e. with observations for every firm each year.

3.1 Model specification

Outputs treated in this model, as suggested by Wangensteen (2012) are:

- Energy distributed (kWh)
- Total number of customers served
- Extension of the grid (km)

As NVE suggests in their output oriented DEA model, the analysis presented here assumes that all companies experience the same input prices. This makes it possible to exclusively look at total cost as the dependent variable and concentrate on the quantity of the explanatory variables (Grammeltvedt et al. 2006). In order to ascertain that the above outputs explain the variations in total costs, a regression analysis on my model is performed before making the frontier analysis. The total costs have been adjusted for the general price increase using the consumer price index provided by Statistics Norway[9]. Other adjustments have been made, as removing companies with an atypical grid. 9 companies (27 observations) were removed because of their small amount of

[9] Statistisk Sentral Byrå, SSB.

customers. All the removed companies have fewer than 100 customers. These companies are large industrial firms with short high voltage lines and a large yearly consumption compared to number of customers. Examples of such companies are Hydro Aluminum AS and Yara Norge AS Glomfjord. There is a leap in number of customers from 90 to 340, depending on which year considered. Therefore, the companies left for the analysis have 340 customers or more. After removing these observations, 130 companies are left for the analysis giving a total of 520 observations over the 4 year time period.

3.1.1 Functional form

The first step in estimating the parameters of a regression model is to specify functional form. As mentioned in chapter 2.6 two appropriate choices are the Cobb-Douglas and translog forms. The following will provide evidence on which model that is applicable in estimating the cost frontier. Starting with a translog function illustrated in Equation 20.

$$\begin{aligned} lnC_i &= \beta_0 + \theta_1 t + \theta_2 t^2 + \beta_1 lnx_{1i} + \beta_2 lnx_{2i} + \beta_3 lnx_{3i} \\ &+ \frac{1}{2}\beta_{11}(lnx_{1i})^2 + \beta_{12}(lnx_{1i}lnx_{2i}) + \beta_{13}(lnx_{1i}lnx_{3i}) \\ &+ \frac{1}{2}\beta_{22}(lnx_{2i})^2 + \beta_{23}(lnx_{2i}lnx_{3i}) \\ &+ \frac{1}{2}(\beta_{33}(lnx_{3i})^2 + \epsilon_i \end{aligned}$$

Equation 20

Where, C_i is the dependent variable, total costs. The total costs are calculated as illustrated in chapter 2.3.4. The explanatory variables $x_{1,}$ $x_{2,}$ $x_{3,}$ are km of high voltage lines, total number of customers, and delivered energy, respectively. Table 3-1 shows the results from a Pooled Ordinary Least Square (POLS) with robust standard errors and clustered sample[10] (Equation 20). The model includes a time trend (t and t^2) with a polynomial of second degree as introduced in chapter 2.6.3. All tests presented assumes a 5% significance level, if not anything else is specified. The insignificant estimates are labelled red.

[10] Cluster is a sample of the individual firm decided from id number of the companies.

Table 3-1:Pooled OLS (POLS) regression with robust standard errors and clustered sample

Estimated variables	Coef.	Robust Std. Err.	t-value
R-squared	0.9853		
β_1 (hv_lines)	0.348	0.035	9.820
β_2 (cust_tot)	0.489	0.094	5.200
β_3 (del_energy)	0.093	0.085	1.100
β_{11}	0.008	0.089	0.080
β_{12}	-0.333	0.132	-2.530
β_{13}	0.265	0.123	2.150
β_{22}	0.643	0.284	2.270
β_{23}	-0.297	0.184	-1.610
β_{33}	0.062	0.130	0.470
θ_1 (time trend t)	0.053	0.008	6.530
θ_2 (time trend t^2)	-0.011	0.002	-4.390
β_0	11.508	0.020	588.040

Indicated in Table 3-1, not all the estimated parameters have expected signs. Neither are all statistical significant. One would expect positive signs on all the estimates. It is a reasonable expectation that costs increase as either of the parameters increase. There does not seem to be a connection between which parameters that is insignificant and which that has a negative sign. The high R^2 indicates that the model explains a large portion of the differences among the firms. Some of the parameters that are included in a translog function and not a Cobb-Douglas (β_{12}, β_{13}, β_{22}, β_{23}) are significant. As noted in chapter 2.5, NVE assumes constant returns to scale (CRS) in their calculation of efficiency scores. This is not applicable in the translog model, but with a Cobb-Douglas functional form. However, after testing if the parameters β_{11} β_{12} β_{13} β_{22} β_{23} β_{33} are mutually equal to zero, the null hypothesis must be rejected (p-value=0.000), hence it is decided to keep the translog model. Therefore, no test on CRS has been conducted. However, testing a Cobb-Douglas function shows that there are sufficient evidences to reject CRS, but this will not be further investigated.

Further, leverage against residual squared plot was conducted[11]. Leverage measures how far a firm is from the industry mean. A company in the upper left corner has high leverage, and could influence on the regressed estimates in Table 3-1. The leverage against residuals squared plot is shown in Figure 3-1.

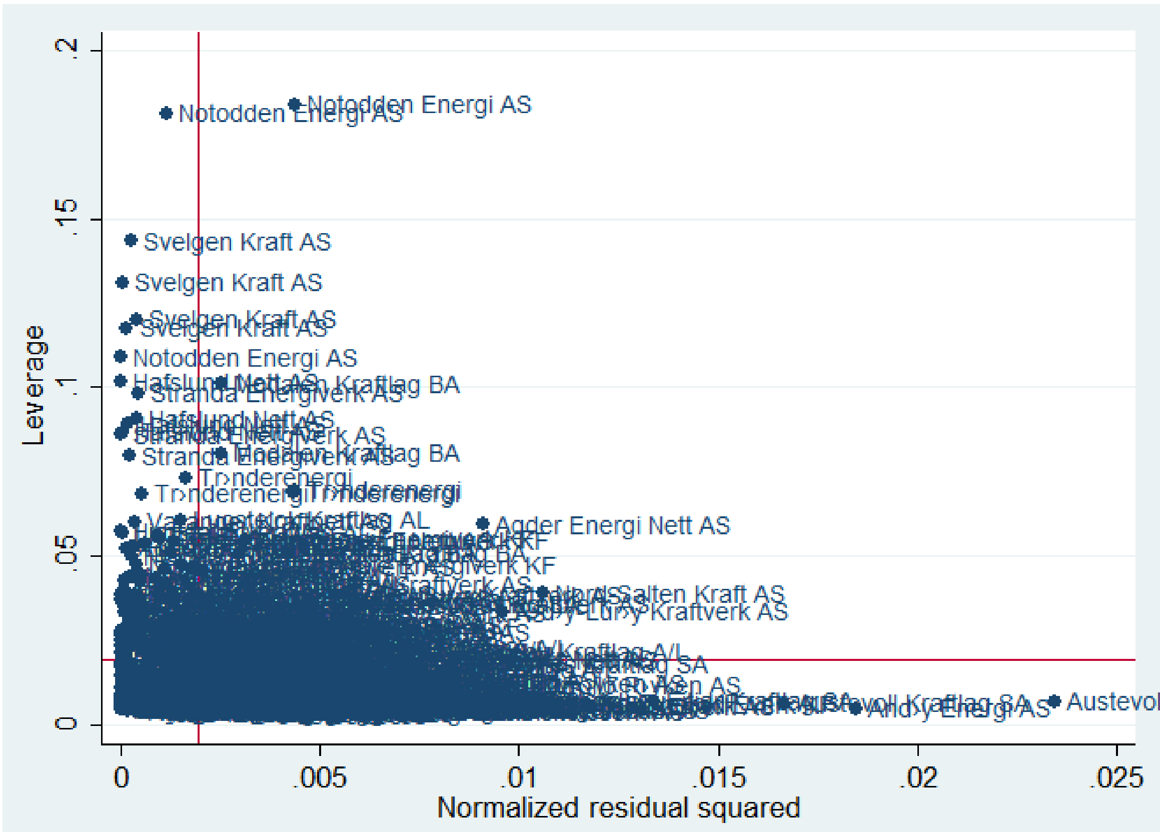

Figure 3-1: Leverage and residual squared plot, translog 2007-2010

Investigation of the companies with either high leverage or large residuals showed that some of these companies delivered a large amount of energy per customers, some more than twice the average (e.g. Notodden Energi AS). Others had an atypical length of their grid (Svegen with only 15 km, average = 713 km). As expected, Hafslund is represented with a high leverage, related to their size. Despite the findings on their size in either customer or length of grid, neither of these was removed from the sample.

[11] In fact the leverage plot was done before the POLS regression with robust standard errors. The leverage plot is not available after introducing robust standard errors.

For further exploration and verification of the model it was tested for functional form using the Ramsey Reset test. The Ramsey Reset test null hypothesis states H_0: Misspecification of functional form.

There is enough evidence provided to reject the null hypothesis (p-value=0.000). Therefore, the result from the Ramsey Reset test provides evidence that supports model misspecification.

3.2 Estimating the cost efficiency frontier

In order to analyse the cost efficiency of the 130 distribution companies a cost frontier was estimated. As with the POLS regression the functional form is a translog function. However, in order to predict such a frontier, SFA introduced in chapter 2.6 is used. The explanatory variables are length of high voltage lines, delivered energy and total number of customers (as before). The parameters are estimated as a linear model with a disturbance term with two components, as described in chapter 2.6.

To account for technological improvement the model includes a time trend. The time trend in a translog function is a second degree polynomial as illustrated in Equation 21 ($\theta_1 t + \theta_2 t^2$) and further explained in chapter 2.6.3[12].

$$
\begin{aligned}
lnC_i &= \beta_0 + \theta_1 t + \theta_2 t^2 + \beta_1 lnx_{1i} + \beta_2 lnx_{2i} + \beta_3 lnx_{3i} \\
&+ \frac{1}{2}\beta_{11}(lnx_{1i})^2 + \beta_{12}(lnx_{1i}lnx_{2i}) + \beta_{13}(lnx_{1i}lnx_{3i}) \\
&+ \frac{1}{2}\beta_{22}(lnx_{2i})^2 + \beta_{23}(lnx_{2i}lnx_{3i}) \\
&+ \frac{1}{2}(\beta_{33}(lnx_{3i})^2 + u_i + v_i
\end{aligned}
\qquad \textbf{Equation 21}
$$

Estimating the frontiers was done assuming different parameterisation, one time invariant and one time varying, (chapter 2.6.3). These estimates are presented in Table 3-2 and Table 3-3, respectively.

[12] t = (year-2007)

Table 3-2: Time invariant estimates, translog paneldata 2007-2010

Estimated variables	Coef.	Std. Err.	z
β_1 (hv_lines)	0.339	0.035	9.730
β_2 (cust_tot)	0.543	0.084	6.430
β_3 (del_energy)	0.046	0.077	0.600
β_{11}	-0.006	0.087	-0.070
β_{12}	-0.415	0.130	-3.190
β_{13}	0.349	0.103	3.380
β_{22}	0.586	0.258	2.270
β_{23}	-0.169	0.169	-1.000
β_{33}	-0.137	0.125	-1.100
θ_1 (time trend t)	0.053	0.009	5.890
θ_2 (time trend t^2)	-0.011	0.003	-3.660
β_0	11.220	0.052	213.940
σ^2 (sigma²)	0.020	0.003	
σ_u^2 (sigma_u²)	0.016	0.003	
σ_v^2 (sigma_v²)	0.004	0.000	
Log likelihood	506.443		

Table 3-3: Time varying estimates, translog panel data 2007-2010

Estimated variables	Coef.	Std. Err.	z
β_1 (hv_lines)	0.339	0.035	9.690
β_2 (cust_tot)	0.544	0.084	6.440
β_3 (del_energy)	0.045	0.077	0.590
β_{11}	-0.007	0.087	-0.080
β_{12}	-0.414	0.130	-3.180
β_{13}	0.349	0.103	3.380
β_{22}	0.585	0.258	2.270
β_{23}	-0.168	0.169	-1.000
β_{33}	-0.138	0.125	-1.100
θ_1 (time trend t)	0.052	0.011	4.640
θ_2 (time trend t^2)	-0.011	0.003	-3.660
β_0	11.221	0.053	210.910
η (eta)	**-0.003**	**0.022**	**-0.120**
σ^2 (sigma²)	0.020	0.003	
σ_u^2 (sigma_u²)	0.016	0.003	
σ_v^2 (sigma_v²)	0.004	0.000	
Log likelihood	506.449		

Here as in the POLS regression, testing if β_{11} β_{12} β_{13} β_{22} β_{23} β_{33} are mutually equal to zero was rejected. Testing the frontier estimates for specified functional form was done manually by adding polynomials of second and third order and testing if these were significant. The F-test provided a p-value = 0.0526 (results shown in Figure 5-3, Appendix B). Since a significant F-statistic suggests some sort of problem with functional form, the provided p-value suggests the opposite. The polynomials are insignificant, and incorrect functional form can be rejected.

The estimated values in the time-varying model have hardly changed compared to the time-invariant model. There is no change in which estimated variables that are insignificant comparing the time invariant parameterisation and the time-varying parameterisation. Both time trend parameters are significant, but Eta is not. This determines the nature of the efficiency improvements (more on this section in the discussion part of the thesis). Both time trend coefficients are significant under both parameterisations. For comparison and to ease the interpretation of a time trend, a frontier with a linear time trend was estimated. These results are found in Figure 5-5 in Appendix B, the results are mentioned in the discussion in chapter 4.

Figure 3-2 shows the density of the individual cost efficiencies estimated. The efficiency score is reported as the inverse of the efficiency score calculated using the method discussed in chapter 2.6. Doing this gives an efficiency score with maximum of 1, making the measure more intuitive. Reporting the inverse cost efficiency is also done by NVE.

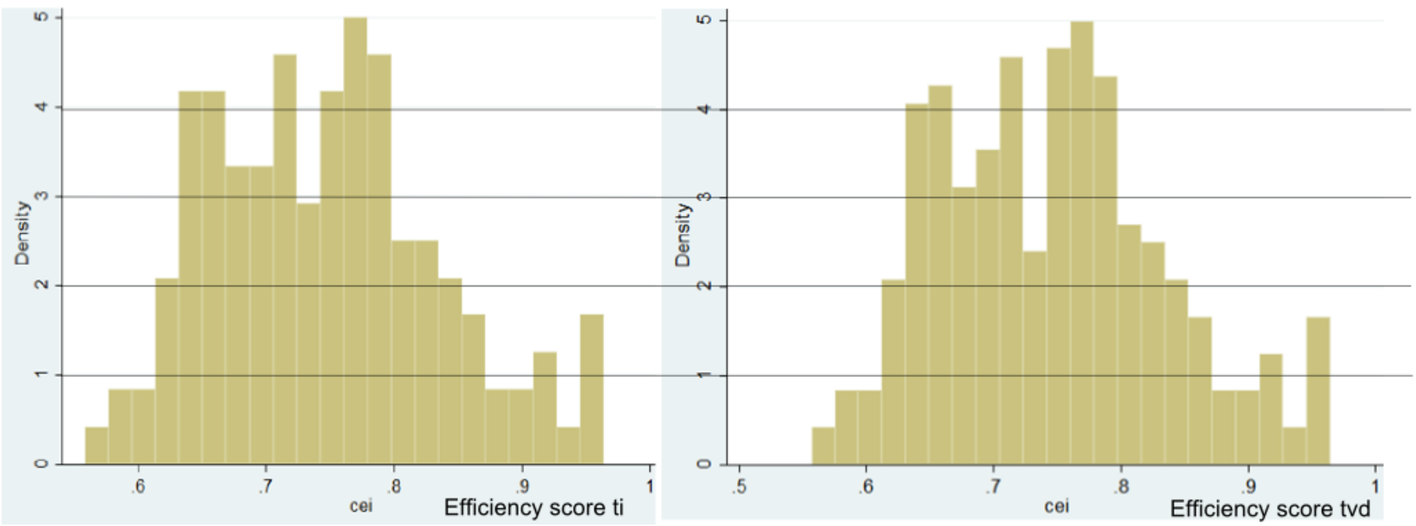

Figure 3-2: Histogram of the inverse cost efficiency score from a translog time-invariant (ti) and a time-varying decay (tvd) frontier model. Panel data 2007-2010

The efficiency scores from the two parameterizations are almost similarly distributed. A calculation of the differences between the two efficiency scores show that these do not differ more than ± 0.002, hence they are effectively equal.

4 Discussion and conclusion

The presented theory and results in the previous two chapters includes many arguments and findings. An extended discussion and conclusion is presented in this chapter.

Even though the initial POLS regression model does not fulfil the functional form test, testing the higher order **β** shows that including these cannot be rejected. Translog has therefore been tested as the functional form in estimation of the cost frontier. However, the estimated cost frontier passes the functional form tests provided. This makes the earlier results less important as the objective of this thesis is to estimate a cost frontier and not a linear function of best fit.

The model includes three outputs and does not correct for geographical or environmental factors. One could argue that the exclusion of such factors influences the frontier estimates. However, tests show that the frontier has a significant correct functional form. In addition, the POLS regression yields $R^2 = 0.985$, indicating that the model explains a large part of the variation amongst the firms. Therefore, there are reasons to believe that the model explains the cost level to a very high degree, not leaving much for the geographical factors to explain. Another argument for not including the environmental factors is that there are ongoing discussions on which method to use when including the geographical factors. NVE changed their practice in this area in setting the revenue cap for 2009 (Øvergaard 2010). Methods and results on testing how the inclusion on how geographical factors influence the model can be found in Øvergaard (2010). However, concluding that the model should not include significant geographical factors would be a bold conclusion. "If they are both statistically significant and economic relevant, it would be hard to argue against their inclusion" (Bergland 2012).

Further testing for individual effects was performed. The Fixed Effects (FE) estimator gave the best estimates. However, the objective of the thesis is to estimate a cost frontier and therefore this is not further investigated in this discussion, the results can be found in Appendix A.

The estimated coefficients of the residuals ($\boldsymbol{\sigma^2 = \sigma_u^2 + \sigma_v^2}$) in Table 3-2 and Table 3-3 indicate that the total residual is highly influenced by the inefficiency parameter. This suggests that the residuals follow a truncated normal distribution, as outlined in chapter 2.6.2. Hence, the noise effect related to uncontrollable effects is small and the inefficiencies are related to how the firm operates.

It is worth noticing that a large portion of the variation in costs is related to the variation in length of the grid, but that the grid squared is insignificant. One percent increase in the number of customers gives an increase of about 0.35 % in total costs. However, there might be potential problems with the inclusion of length of grid as an explanatory factor. In the long run it is a reasonable assumption that this variable is closely correlated to the use of capital (which mainly means grid investments) making it not an exogenous variable[13]. "An investment made in grid extension will lead to an increase in the output, whether or not the investment is needed" (Wangensteen 2012). Hence, the level of endogeneity is controlled by the companies, where making bad investments could increase their level of inefficiency (endogeneity). However, the cost function is estimated over a four year period and must be seen as a short term cost function, not including capital costs. Therefore, despite the potential long run endogeneity problem, both hv_lines and hv_lines squared are kept as explanatory variables in the model.

Comparing the estimates provided by the time invariant model in and the time-varying model shows that these are practically equal. Eta (η) in the time-varying model (Table 3-3) is negative indicating a decrease in technical efficiency, however this effect is insignificant (z-value = -0.120) making the function (f(t)) from Equation 19 in chapter 2.6.3 equal 1. This means that there are not enough evidence provided to suggest an improvement of the individual firm's technical efficiency, either positive or negative, with the Battese and Coelli function from chapter 2.6.3. However, several more flexible models are found in the economic literature on efficiency. The Battese and Coelli model does not allow for a change in the rank ordering of the firms over time meaning that the firm ranked as; e.g. the third most efficient firm is always ranked as the third most efficient firm in the period explored. Other models, such as the model introduced by Cornwell, Schmidt and Sickles in 1990 assume a different function for $u_{it.}$ By setting u_{it} as in Equation 22

$$u_{i_t} = \beta_{0_t} - \theta_{i_1} - \theta_{i_2} t - \theta_{i_3} t^2 \qquad \textbf{Equation 22}$$

the rank ordering of the firms is allowed to change over time (Coelli et al. 2005). Other functions for explaining the nature of the technical efficiencies are also available, but neither are discussed any further as they are outside of the scope this thesis.

---

[13] Any variable that is uncorrelated with the error term in the model of interest (Wooldridge 2009).

The time trend included in the translog function (Equation 21) is significant and the second order polynomial allows the technological effect to change with time according to Chapter 2.6.3. Table 3-3 provides the estimated values of θ_1 and θ_2 hence the technological change effect is determined. $\theta_1 = 0.052$ and $\theta_2 = -0.011$, and according to the negative sign of θ_2 determines that the technological change is decreasing the costs (Coelli et al. 2005). However the interpretation of this model is difficult. There are two parameters to estimate and it has a non-linear effect over time. The short period makes an interpretation of a linear time trend more feasible. In Appendix B Figure 5-6 shows a estimated model with a significant linear time trend. This time trend is however positive. θ_1 is estimated to be 0.0209 indicating an increase in cost related to technological progress. This result does not agree with the expected in development of technology regarding costs, one would expect technology to decrease cost not increase.

Summing the above, the estimated parameters indicates that the simplest form of time trend is significant, however concluding that a cost reduction during the time period is related to a change in the technology is inappropriate.

Conclusively, the findings of this thesis suggest that the regulatory model so far has failed to induce firms to operate at an efficient level. It is particularly striking that the inefficiency in the industry has not declined over time. This could lead to a loss of the consumer surplus related to higher prices and smaller quanta, than if correctly regulated in regards to efficiency. One possible explanation could be that the model suffers from a low ρ, the weighting of normalized costs. The size of the weight was adjusted in 2009 from 0.5 to 0.6. A further increase would make the model more restrictive. In addition, the guaranteed profit minimum of 2% could be an argument for not improve efficiency, however this thesis is not accusing any of the firms of speculate in this. This has not been investigated and changing ρ would probably have other side effects, none in which discussed here. However, the findings suggest that there is a technological progress varying with time when including a second polynomial time trend, but that including a linear time trend stated the opposite. Therefore, no conclusion is made about the technological development, a result that probably has to do with the short time period.

5 References

Berg, S. V. & Tschirhart, J. (1988). *Natural monopoly regulation: principles and practice*. Cambridge, Cambridge University Press. xii, 564 s. p.

Bergland, O. (2011a). *Econometric Methods*. 02 .11.2011 ed. Lecture ECN301. Ås, Norway, University of Life Science, Department of economics and resource management.

Bergland, O. (2011b). *Econometric Methods.* 01 .11.2011 ed. Lecture ECN301. Ås, Norway, University of Life Science, Department of economics and resource management.

Bergland, O. (2011c). *Energy Markets and Regulation*. 31.10.2011 ed. Lecture ECN380. Ås, Norway, University of Life Science, Department of economics and resource management.

Bergland, O. (2012). *Geographical factors* (14.5.2012).

Bjørndal, E., Bjørndal, M., Pardalos, P. M. & Rönnqvist, M. (2010). *Energy, Natural Resources and Environmental Economics*. Berlin, Heidelberg, Springer-Verlag Berlin Heidelberg.

Bogetoft, P. & Otto, L. (2011). *Benchmarking with DEA, SFA, and R*. New York, NY, Springer New York.

Coelli, T. J., Rao, D. S. P. & Battese, G. E. (2005). *An Introduction to efficiency and productivity analysis*. Boston, Mass., Springer Science+Business Media. 342 s. p.

Deregulering av det norske el-markedet. Energi Norge Available at: http://www.energifakta.no/documents/Oekonomi/Kraftmarked/deregulering.htm (accessed 4.3.2012).

Econ, P. (2008). *Benchmarking modeller og incentiver*. Available at: http://www.nve.no/Global/Kraftmarked/%C3%B8konomisk%20regulering%20av%20nettselskap/Rapport%20benchmarking%20og%20incentiver.pdf?epslanguage=no (accessed: 14.3.2011).

Grammeltvedt, T. E., Hansen, W., Haugland, C. P., Karlsen, E. N., Langset, T., Neurauter, T. M., Tjeldflåt, A. & Wiull, S. O. (2006). Modell for fastsettelse av kostnadsnorm

Lundteigen, M. F. (2011). Oppsummering av høringsuttalelser. Available at: http://www.nve.no/PageFiles/12683/Oppsummering%20av%20h%C3%B8ringsuttalelser.pdf.

Shleifer, A. (1985). *A Theory of Yardstick Competition* 10.

Train, K. E. (1991). *Optimal regulation: the economic theory of natural monopoly*. Cambridge, Mass., MIT Press. xiv, 338 s. p.

Varian, H. R. (2006). *Intermediate microeconomics: a modern approach*. New York, W.W. Norton & Co. XXIV, 715, 39 s. p.

Von Der Fehr, N. H. M. (2010). *Den økonomiske reguleringen av strømnettet*. Available at: http://www.regjeringen.no/upload/OED/2010-1018_nettregulering.pdf (accessed: 16.01.2012).

Wangensteen, I. (2012). *Power system economics: the Nordic electricity market*. Trondheim, Tapir Academic Press. 365 s. p.

Waterson, M. (1988). *Regulation of the firm and natural monopoly*. Oxford, Blackwell. 164 s. p.

Wooldridge, J. M. (2009). *Introductory econometrics: a modern approach*. Mason, Ohio, South-Western Cengage Learning. XX, 865 s. p.

xt-frontier - Stochastic frontier models for panel data. (2012). Available at: http://www.stata.com/features/stochastic-frontier/xtfrontier.pdf (accessed 23.3.).

Øvergaard, H. (2010). *Rammevilkårenes betydning i en effektivitetsanalyse*. Available at: http://brage.bibsys.no/umb/handle/URN:NBN:no-bibsys_brage_15644 (accessed: 6.1.2012).

5.1 Appendices

A. Individual effects, fixed effects versus random effects

Testing for Random Effects (RE) versus Fixed Effects (FE) was performed for further investigation of unobserved effects. First a Breusch and Pagan Lagrangian multiplier test was done to check for RE effects. The null hypothesis of Var(ε) = 0 was rejected at a 1 % significance level (Figure 5-1). Secondly a Hausmann test was performed to test if the RE and FE estimates differed. The null hypothesis states: H_0: The differences in the coefficients are not systematic (Bergland 2011a). With a p-value = 0.0011 (chi^2 = 27.70), the null hypothesis was rejected, making FE the better estimator. Testing for serial correlation in the FE estimates was (as earlier) done with the inclusion of lagged residuals. The lagged residual were tested if equal to zero. The t-test provided not enough evidence to reject H_0: no autocorrelation, (p-value = 0.3233), hence no serial correlation suspected in the FE model.

Testing for unobserved effects using lagged residuals was performed. According to the null hypothesis of significant lagged residual (H_0: l.ε. = 0) must be rejected, indicating unobserved effects. This means that the lagged residual affect the dependent variable, significantly. Since there are proven unobserved effects further investigation was conducted looking for random or fixed effects. The key requirement for RE or FE to produce reliable estimates is that the explanatory variables are uncorrelated with the unobserved cluster effect (Wooldridge 2009). In the process of deciding between RE and FE, the Breusch-Pagan and the Hausman tests provided evidence which rejected RE as the correct estimator and confirmed that FE is a good estimator for the model. Further testing showed that there is no serial correlation in the FE estimates.

```
. xttest0

Breusch and Pagan Lagrangian multiplier test for random effects

        ltc_deai[id,t] = Xb + u[id] + e[id,t]

        Estimated results:
                         |       Var     sd = sqrt(Var)
                ---------+-----------------------------
                ltc_deai |   1.295666       1.138273
                       e |   .0048307       .0695032
                       u |   .0158327       .1258282

        Test:   Var(u) = 0
                             chi2(1) =   423.18
                         Prob > chi2 =     0.0000
```

Figure 5-1: Breusch and Pagan Lagrangian multiplier test for ransom effects

```
hausman hfe hre

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |      hfe          hre         Difference          S.E.
-------------+----------------------------------------------------------------
         lw1 |    .0988083     .3480558       -.2492475        .2500546
         lw2 |    1.322945     .4100008        .9129442        .3980538
         lw3 |    .1718898      .166569        .0053208        .0975974
        lw11 |   -.2293284    -.0284576       -.2008708        .3088754
        lw12 |   -.3938228    -.3950232        .0012004        .2722596
        lw13 |     .502578      .356889         .145689         .137792
        lw22 |    .5172537     .4284222        .0888315        .4255611
        lw23 |     .061381    -.0574856        .1188666        .1843802
        lw33 |   -.4791567    -.2312892       -.2478675        .1177089
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                  chi2(9) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                          =       27.70
                Prob>chi2 =      0.0011
```

Figure 5-2: Hausman test for differences in FE and RE estimations.

```
Log likelihood  =  514.49569                    Prob > chi2        =    0.0000

------------------------------------------------------------------------------
    ltc_deai |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         lw1 |   2.226002   .7754406     2.87   0.004     .7061664    3.745838
         lw2 |   3.269164   1.140513     2.87   0.004       1.0338    5.504527
         lw3 |   .4422988   .1739447     2.54   0.011     .1013734    .7832242
        lw11 |   .1610836   .1036965     1.55   0.120    -.0421578     .364325
        lw12 |  -2.225913   .7796039    -2.86   0.004    -3.753909   -.6979177
        lw13 |   1.916821   .6721371     2.85   0.004     .5994569    3.234186
        lw22 |   2.894972   1.034517     2.80   0.005     .8673565    4.922587
        lw23 |  -.5822678   .2577794    -2.26   0.024    -1.087506   -.0770295
        lw33 |  -1.067023   .3870755    -2.76   0.006    -1.825677    -.308369
           t |   .2789186   .0970077     2.88   0.004      .088787    .4690503
          t2 |  -.0664674   .0231587    -2.87   0.004    -.1118576   -.0210773
          y2 |    -.36414   .1537531    -2.37   0.018    -.6654904   -.0627895
          y3 |   .0073165   .0038402     1.91   0.057    -.0002102    .0148432
       _cons |   46.79732    14.6762     3.19   0.001     18.03248    75.56215
-------------+----------------------------------------------------------------
         /mu |   .3028831   .0569025     5.32   0.000     .1913562    .4144101
        /eta |  -.0211401   .0237005    -0.89   0.372    -.0675921     .025312
   /lnsigma2 |  -3.949144   .1354731   -29.15   0.000    -4.214667   -3.683622
  /ilgtgamma |   1.305181   .1898791     6.87   0.000     .9330247    1.677337
-------------+----------------------------------------------------------------
      sigma2 |   .0192712   .0026107                      .0147772    .0251318
       gamma |   .7867056   .0318617                      .7176885    .8425516
    sigma_u2 |   .0151607   .0026087                      .0100477    .0202738
    sigma_v2 |   .0041104   .0002965                      .0035293    .0046915
------------------------------------------------------------------------------

. test y2 y3

 ( 1)  [ltc_deai]y2 = 0
 ( 2)  [ltc_deai]y3 = 0

           chi2(  2) =    5.89
         Prob > chi2 =    0.0526
```

Figure 5-3:Testing for functional form cost frontier

B. Additional Results

Correlation between the explanatory variables is presented in Table 5-1.

Table 5-1: Correlation explanatory variables

	ldel_energy	lhv_line	lcust_tot
ldel_energy	1		
lhv_lines	0.9098	1	
lcust_tot	0.984	0.9268	1

As expected, the explanatory variables are nearly perfectly correlated. An increase in length of the lines is correlated with number of costumers and delivered energy.

Figure 5-4 displays the estimated efficiency scores assuming time-invariant and one with time-varying decay model.

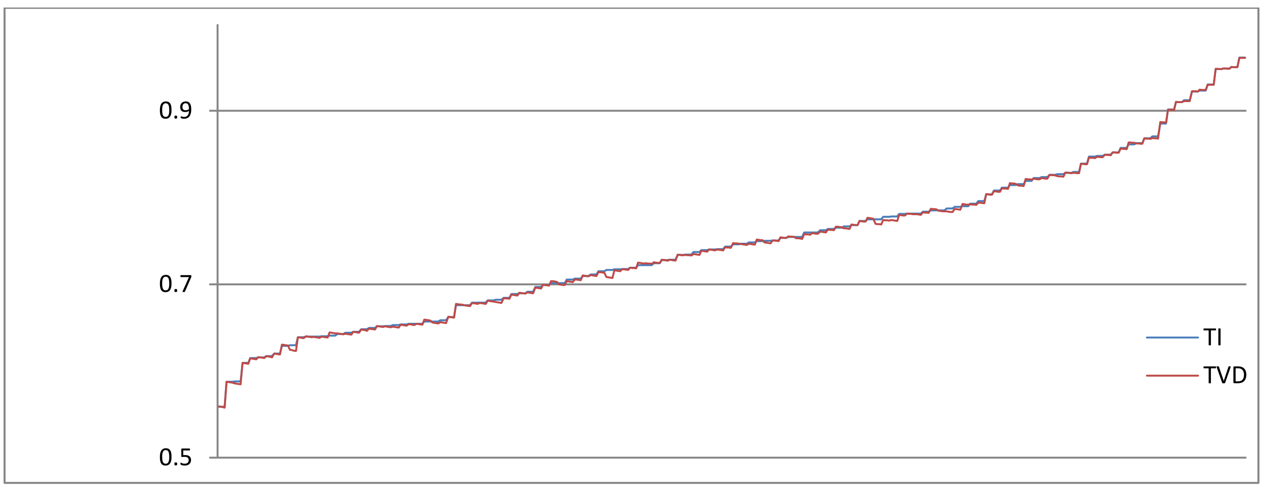

Figure 5-4: The above show the tvd and the ti efficiency score. As showed in histogram (Figure 4 2) these are basically the same scores. Sorted from smallest value to the largest value

Figure 5-5 shows estimated coefficients of the cost frontier estimated with a linear time trend.

```
Time-varying decay inefficiency model           Number of obs      =       520
Group variable: id                              Number of groups   =       130

Time variable: year                             Obs per group: min =         4
                                                               avg =         4
                                                               max =         4

                                                Wald chi2(10)      =  11100.88
Log likelihood  =  499.85157                    Prob > chi2        =    0.0000
```

ltc_deai	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
lw1	.3372415	.0350037	9.63	0.000	.2686354	.4058475
lw2	.6029911	.0838736	7.19	0.000	.4386019	.7673803
lw3	-.0113064	.0760292	-0.15	0.882	-.1603209	.1377081
lw11	-.0046415	.0873708	-0.05	0.958	-.1758852	.1666022
lw12	-.4218401	.1308062	-3.22	0.001	-.6782155	-.1654647
lw13	.3551157	.1038947	3.42	0.001	.1514858	.5587457
lw22	.6917329	.2588252	2.67	0.008	.1844449	1.199021
lw23	-.2478632	.1690838	-1.47	0.143	-.5792614	.083535
lw33	-.0834594	.1253817	-0.67	0.506	-.3292031	.1622843
t	.0208727	.0068968	3.03	0.002	.0073553	.0343901
_cons	11.22489	.0517612	216.86	0.000	11.12344	11.32634
/mu	.2998218	.0518005	5.79	0.000	.1982947	.4013488
/eta	-.0026689	.0208129	-0.13	0.898	-.0434613	.0381236
/lnsigma2	-3.886307	.1358874	-28.60	0.000	-4.152642	-3.619973
/ilgtgamma	1.306846	.1905221	6.86	0.000	.9334293	1.680262
sigma2	.020521	.0027885			.0157228	.0267834
gamma	.7869849	.0319391			.7177705	.8429393
sigma_u2	.0161497	.0027877			.0106859	.0216135
sigma_v2	.0043713	.0003143			.0037552	.0049873

Figure 5-5: Time-varying cost frontier with a linear time trend, translog 2007-2010.

Figure 5-6 shows the distribution of the estimated efficiency scores with a linear time trend.

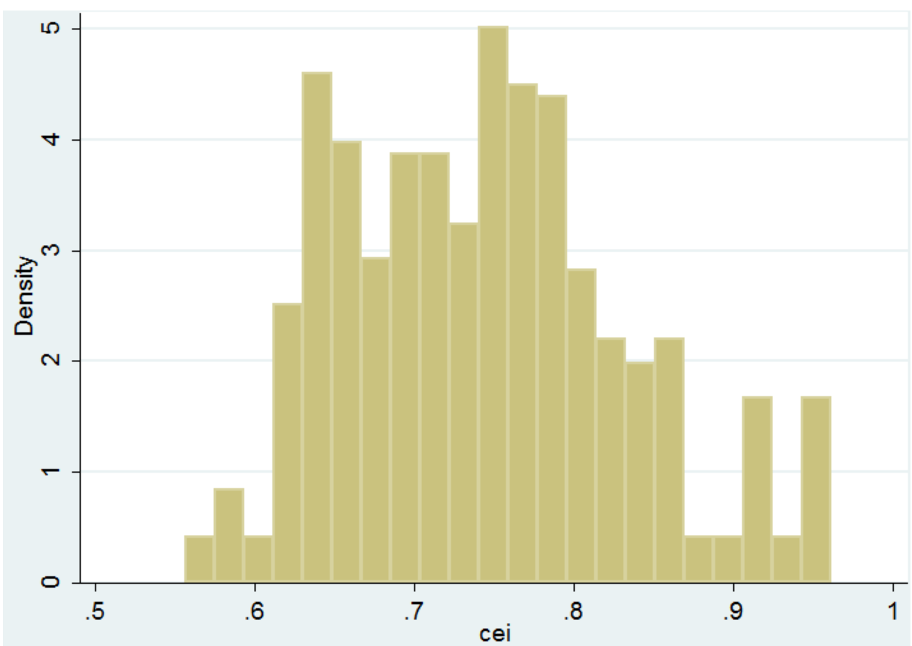

Figure 5-6: Efficiency distribution, linear time trend. Translog 2007-2010.

C. *Assumtions for POLS regression*

The POLS model is:

$$y_{i_t} = \vec{x}_{i_t}\beta + u_{i_t}$$

POLS assumptions 1-4:

1. Linear in the parameters
2. Random sampling from cross sections
3. No exact linear relationship among regressors

$$rank E(\vec{x}'\vec{x}) = k$$

4. Population orthogonal assumption

$$E(\vec{x}' u) = 0$$

This requires only weak exogenity

$$E\left(x' u_{i_t}\right) = 0 \; \forall \; i, t$$

Theorem

Under the assumptions POLS 1-4 the POLS estimator

$$\hat{\beta}_{POLS} = (\vec{x}'\vec{x})^{-1}\vec{x}\vec{y}$$

is a consistent estimator for β i.e.

$$plim\hat{\beta}_{POLS} = \beta$$

(Bergland 2011b)

Assumptions for the inefficiency and noise parameters (Coelli et al. 2005, p.245)

$E(vi) = 0$	zero mean
$E(\sigma_v^2)=0$	homoskedastic
$E(v_i v_j)=0$ for all $i \neq j$	uncorrelated
$E(u_i^2)$ = constant	homoskedastic
$E(u_i u_j) = 0$ for all $i \neq j$	uncorrelated

Under these assumptions we can obtain consistent estimators of the slope coefficients using ordinary least squares (OLS). However, the OLS estimator of the intercept is biased downwards making OLS a bad estimator for the technical efficiency. A better solution is the Maximum Likelihood method (Coelli et al. 2005).

D. Extension of the Revenue cap

A company's costs are calculated according to Equation 23:

$$K_t = (DV_{t-2} + CENS_{t-2})x\left(\frac{KPI_t}{KPI_{t-2}}\right) + NT_{t-2}\,x\,P_t + AVS_{t-2} + AKG_{t-2}\,x\,r_{NVE} \quad \textbf{Equation 23}$$

Where:

K_t = the firms reported costs in eRapp

DV_{t-2} = operating and maintance expenditure in year t-2

$CENS_{t-2}$ = CENS Norwegian Compensation for Energy Not Supplied.

KPI = consumer price index

NT_{t-2} = Power loss per MWh

P_t = area price per MWh

AVS_{t-2} = yearly depreciation

AKG_{t-2} = Avkastningsgrunnlaget

r_{NVE} = NVE's reference rent

NVE regulates each firm's allowed revenue, given in Equation 25 . This way the tariff paid by the customers is kept at a regulated level. The economic regulation is related to the firm's revenue and every year NVE sets each firms allowed revenue level. The firm's economic result depends on how well the firm is able to keep its costs within this allowed revenue. The firm sets the customers tariff themselves, and this combined with the customer's consumption decides the actual revenue. Actual revenue does not cover costs related to costumer initiated improvements or expansions of existing installation that normally cannot be demanded delivered . If the allowed revenue and actual revenue does not suffice as a result of the wrong level of tariff sat by the firm (firm did not manage to set the tariff correctly), this is adjusted with the next year's tariff.

The allowed revenue (TR) is defined as the sum of the revenue cap (IC) plus the costs related to connection to the region grid (KON) and real estate tax minus grid failure costs (KILE). This is illustrated in equation 3

$$TR = IC + KON + E - CENS$$ **Equation 24**

While grid failure costs and real estate taxes are handled as costs outside the manager's control, CENS is highly possible to reduce and therefore enlarge the allowed revenue. When, durability and dimensions are critical factors for estimating the CENS cost and withdrawn in the year the failure takes place.

As mentioned the model has included the difference in depreciation (AVS) and difference in return (AKG) to compensate for the capital investment done in this period, since 2009

The regulatory models way of compensate for investments has been discussed and suffered changes the last couple of years. The discussions have changed form, and from giving too large incentives for investments the arguments are now that these incentives are hardly present. From 2007 to 2009 the cost norm included a compensation parameter (CP) to compensate for losses related to the new investments time lag (Grammeltvedt et al. 2006). This was changed in 2009. "From 2009 the time lags have been removed, so that there is no longer need for the compensation parameter" (Bjørndal et al. 2010, p.322). With the given revenue cap, companies who wish to improve their economic results need to decrease their costs. This can be done either by an efficiency improvement or limit the size of their company. If demand is dependent of prices, an increase in prices would lead to a drop in traded volume inducing lower cost at the same level of revenue (Von Der Fehr 2010). Therefore revenue cap regulation provides incentives to increase efficiency and adjust the size of the firm.

Shown in Equation 25.

$$TI_t = IR_t + KON_t + E_t - CENS_t + (AVS_t - AVS_{t-2}) + (AKG_t - AKG_{t-2})xr_{NVE} \quad \textbf{Equation 25}$$

Put together the total revenue calculated NVE every year is presented in Equation 26.

$$\begin{aligned} TI_t = (1-\rho)\Bigg[(DV_{t-2} + CENS_{t-2})x\left(\frac{KPI_t}{KPI_{t-2}}\right) + NT_{t-2}\, x\, P_t + AVS_{t-2} \\ + AKG_{t-2}\, x\, r_{NVE}\Bigg] + \rho K_t^* + KON_t + E_t - CENS_t \\ + (AVS_t - AVS_{t-2}) + (AKG_t - AKG_{t-2})xr_{NVE} \end{aligned} \quad \textbf{Equation 26}$$

E. STATA 11.1 code

```
**summary of data

Sum

**description of data

Des

**set as paneldata

xtset id year

**remove data

drop if  cust_tot < 300

**scale units to unit means

egen hbar = mean(hv_lines)

egen cbar = mean(cust_tot)

egen dbar = mean(del_energy)

**generate variables

gen w1   = hv_lines/hbar

gen w2   = cust_tot/cbar

gen w3   = del_energy/dbar

gen lw1  = log(w1)

gen lw2  = log(w2)

gen lw3  = log(w3)

gen lw11 = 0.5*lw1*lw1

gen lw12 = lw1*lw2

gen lw13 = lw1*lw3
```

```
gen lw22 = 0.5*lw2*lw2

gen lw23 = lw2*lw3

gen lw33 = 0.5*lw3*lw3

rename   indeksregulerttcmot2011priser tc_deai

gen ltc_deai =log(tc_deai)

**OLS

reg ltc_deai lw? lw1? lw2? lw3?

estimates store reg

**lw11 lw33 not significant

**leverage versus residuals squared

lvr2plot, mlabel(company)

hist r2

predict r

gen r2 = r*r

hist r2

**testing ols for functional form

reg ltc_deai lw? lw1? lw2? lw3?, r

estimates store ols_r

**functional form, Ramsey Reset test

ovtest

gen t = year -2007

gen t2=t*t

xtset id year
```

**Estimating POLS

reg ltc_deai lw? lw1? lw2? lw3?, r cluster(id)

estimates store pols

predict ce, te

gen cei =1/ce

list year company cei

drop ce cei

**Estimating POLS with time trend

reg ltc_deai lw? lw1? lw2? lw3?

reg ltc_deai lw? lw1? lw2? lw3? t t2, r cluster(id)

**Ramsey-Reset functional form

Ovtest

**Test for significance

**frontier estimation

xtfrontier ltc_deai lw? lw1? lw2? lw3?, cost ti

predict ce, te

gen cei=1/ce

list year company cei

xtfrontier ltc_deai lw? lw1? lw2? lw3? t t2, cost tvd

drop ce cei

predict ce, te

gen cei=1/ce

```
list year company cei


**add lagged residuals and test for significance

reg ltc_deai lw? lw1? lw2? lw3? l.u, r cluster(id)

test l.u = 0

**rejected

xtreg ltc_deai lw? lw1? lw2? lw3?, re r

estimates store re

xtreg ltc_deai lw? lw1? lw2? lw3?, fe r

estimates store fe

estimates table pols re fe, se t

estimates restore re

xttest0

help xttest0

estimates table pols re fe, se t

hausman fe re

xtreg ltc_deai lw? lw1? lw2? lw3?, fe

estimates store hfe

xtreg ltc_deai lw? lw1? lw2? lw3?, re

estimates store hre

**hausman test

hausman hre hfe

hausman hfe hre
```

```
estimates restore fe

drop u

predict e, r

predict e,r

predict u,e

**add lagged residual

xtreg ltc_deai lw? lw1? lw2? lw3? l.u, fe r

test l.u.

test l.u=0
```