QUALITY AND COSTS AT THE HOSPITAL LEVEL IN NORWAY

Does there exist a trade-off between quality and costs at Norwegian hospitals?

Author:
Elise Emilsen Bakke

Supervisor:
Sverre A. C. Kittelsen

Thesis submitted as a part of the Master of Philosophy Degree in Health Economics, Policy and Management

Department of Health Management and Health Economics
The Faculty of Medicine
University of Oslo

May 15, 2019

2019

Quality and Costs at the Hospital Level in Norway

Elise Emilsen Bakke

http://www.duo.uio.no/

Printed: Reprosentralen, University of Oslo

ABSTRACT

Background: Norway has throughout history introduced several reforms in order to aim focus on cost containment and improving efficiency in the health sector. However, expenditures are rising and there is a political pressure to contain costs as well as being efficient. As a result, it has led to a larger policy focus on efficiency and costs, but this can affect the level of quality of the health services. It is therefore an interest to find more information on the association between quality and costs across the Norwegian hospitals.

The objective: This paper analyses the relationship between costs and quality across 22 Norwegian hospitals for the years 2008 to 2014. The objective is to study the association between quality and costs at 22 Norwegian hospitals and determine if a trade-off is present between quality and costs. A previous study found no clear cost-quality trade-off. Further on, average cost efficiency is estimated for all hospital across the 7 years in order to see how the level of cost efficiency varies.

Method: Case-mix adjustment was used in order to adjust for DRG, patient characteristics and treatment variables to create hospital performance measures for three models. Case-mixed performance indicators such as emergency readmission within 30 days and mortality within 30 days were used to measure the quality level. A Stochastic Frontier Analysis was used for a 7-year panel data (2008-2014) in order to estimate the inefficiency across the hospitals, where high mortality will be represented as low quality.

Results: SFA results showed that there exists an association between costs and quality across Norwegian hospitals. When performance indicator for 30-day mortality and 30-day emergency readmission increases it means quality is low and costs are low at the same time. Mortality within 30 days was statistically significant at 5% level and stronger than emergency readmission. A trade-off between costs and quality was therefore found. Average cost efficiency was 89.5%, where Oslo University Hospital had the lowest score while the most efficient hospital was Vestfold Hospital Trust. In general, university hospitals in Norway had some of the lowest cost efficiency scores.

Conclusion: The results indicate that a trade-off between cost and quality is present at the hospital level in Norway. When quality is low, costs are low and vice versa. In order to achieve high quality, it is found that costs will increase. Cost efficiency is present across the hospitals. Average cost efficiency was found to be 89.5%. It can be concluded that the hospitals overall are very cost efficient, but university hospitals had some of the lowest scores. Future studies on how hospital specialization affects cost efficiency using same model could be interesting.

ACKNOWLEDGEMENT

I would like to thank my supervisor Sverre A. C. Kittelsen for his guidance and help during the months I stayed at the Frisch Centre. Thank you for providing me with an office space and for being available to help me with my thesis. I really appreciate the advice I was given.

I also want to thank the Department of Health Management and Health Economics and my fellow students. It has been a privilege being a student at the department during the years I studied there.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST ACRONYMS

ABF	Activity Based Financing
AC	Average Cost
CE	Cost Efficiency
DEA	Data Envelopment Analysis
DRG	Diagnosis-Related Groups
GDP	Gross Domestic Product
HF	Health Trust
ISF	Innsatsstyrt Finansiering
MC	Marginal Cost
NOK	Norwegian Kroner
OECD	Organisation for Economic Co-operation and Development
RHF	Regional Health Authority
SFA	Stochastic Frontier Analysis
TE	Technical Efficiency
VAT	Value Added Tax

Chapter 1 INTRODUCTION

The following chapter presents an introduction about the healthcare in Norway, as well as how the health sector is financed in today's society. Furthermore, the research question and aim of the study will be presented, as well as why the subject is relevant to research. At last a general description of previous studies will be presented.

1.1 Healthcare and financing in Norway

Norway is known for having one of the best healthcare systems in the world, based on a classical Scandinavian Welfare model, where every inhabitant has equal right to access health services and to choose their preferred provider. The country uses a large proportion of the country´s GDP on health care expenditure. In 2017 it was estimated to be 65 000 NOK per capita which is 10.4 per cent of total GDP (Statistics Norway, 2018a). In fact, the health care expenditure per capita is constantly growing and this is typical for most OECD countries. As a result of increased expenditure in the health care sector in Norway, there has also occurred a larger focus on efficiency as well as costs which has led to them becoming important policy issues.

The Regional health authorities (RHF), owned by the Norwegian state, have since 2002 had the responsibility for all the public hospitals. Somatic hospitals are financed through activity-based financing (ABF), called ISF-ordningen (ISF) in Norwegian, based on a diagnosis-related group (DRG) system, and block grant. Currently the somatic services are financed 50% through ABF and 50% through block grant. The block grant is given to RHF through the state budget and the sum is determined by the number of inhabitants and age composition in the regions (Helsedirektoratet, 2017). ABF rewards increased activity and gives incentives for the hospitals to treat below the DRG weight. The purpose of the financing system in Norway is to motivate the hospitals to focus on being cost-effective, reduce waiting times and increase activity, with expectations that services are of good quality. All this applies regardless of social status, gender, age, financial situation and ethnicity.

Over the past two decades Norway has gone through a health modernization by introducing several reforms, such as reimbursements reforms, in order to provide the health sector with more incentives to motivate focus on cost containment, efficiency and accessibility. As a result, Norway has today a mixed financing system in order to fulfill these needs as best as possible. However, the hospital expenses are constantly increasing, and there is a political pressure to keep costs down while being cost-effective and delivering services of good quality. In addition, it can be discussed whether the DRG-system may be giving too strong incentives when it comes to increasing efficiency since increased activity gives higher grants, while those who treat below average price are given incentives to produce more. As a result, this could lead to quality being underprioritized at hospitals. The way the hospitals are financed is essential in the way that it needs to give the right incentives at the right place. Hospitals can end up delivering poorer quality as a result of the pressure to keep within their budget. According to economic theory it is claimed that incentives to keep costs down can indeed affect the quality of the services (Kittelsen *et al.*, 2017). Therefore, the financing system is an important tool in order for hospitals to prioritize correctly.

When the hospital reform was introduced in 2002 it gave the state the responsibility of the ownership of all public hospitals and five regional health authorities were created (today four) that were given responsibility for the specialist healthcare in the four regions of Norway. Cost control, allocating resources efficiently and reducing the waiting list were just some of the objectives to create a better health care sector. The reform also resulted in hospitals becoming larger and larger. This is a result of merging hospitals together, and as a consequence, smaller hospitals were closed down. In order to save the Health trust (HF) millions of NOK in terms of efficiency, the regional health authorities have been closing down hospitals to create larger, more specialized hospitals. The centralization has been argued as being profitable, but at the same time it is argued that when hospitals become larger so do the challenges. There is an increased expenditure in the healthcare sector, especially for hospitals, and therefore cost control must be prioritized as well as delivering services of good quality.

1.2 Presentation of research question

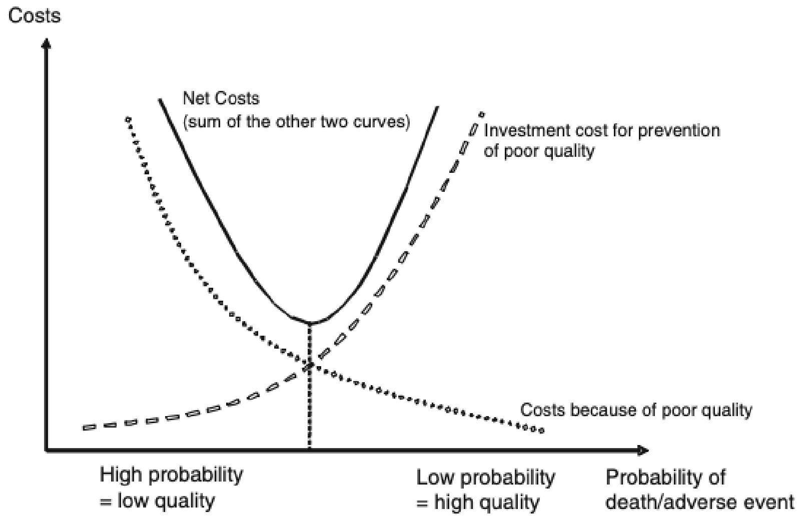
The hospitals in Norway deliver services for 5.3 million inhabitants and in 2017 there were registered 2 million patients across the hospitals (Statistics Norway, 2018b). Patients are expected to be provided with health services of high quality independent of which hospital in the country they are admitted to. Since cost containment is a goal, and reforms throughout the years have tried to solve problems with containing costs, there has been an increased focus on the importance of efficiency and utilizing the capacity of the hospitals. Hospitals struggle with optimizing the relation between cost and quality, where an ideal scenario would be to provide high quality using as few resources as possible, but in order for hospitals to contain costs they must become more efficient. Hospitals can for example be more efficient by reducing the average length of stay or by reducing readmissions. Consequently, this can in worst case scenario lead to poorer quality if reducing costs becomes the most driving goal. Often administrative managers will have to choose where to reduce resources and still deliver an acceptable level of quality. Therefore, research around this has increased in order to attain more understanding about these concerns.

The thesis aims on examining and answering if there exist a trade-off between costs and quality across 22 Norwegian hospitals. By using quality indicators, such as morality within 30 days, and a cost frontier through Stochastic Frontier Analysis (SFA) it is possible to assess hospital performances that captures hospital quality. When quality is high at a hospital it means that there is a trade-off between cost containment and quality improvement, while if quality is low then there could be potential of costs saving of improving quality (Kittelsen *et al.*, 2015a). Trade-off between costs and quality can be found by interpreting the SFA regression results and see how output is associated with hospital costs. If for example mortality within 30 days has a positive coefficient it implies no trade-off because an increase in mortality (low quality) results in increased costs. If the coefficient is negative it will indicate a trade-off between costs and quality because an increase in mortality (low quality) is associated with lower costs, i.e. quality comes at a cost.

Scale elasticity will also be calculated in order to find whether it is optimal to be a large hospital or a small hospital. Scale economies has been a topic in many countries, because analyses can give valuable information about the optimal productive hospital size within a country. If economies of scale exist within hospital production it means that large hospitals are able to increase productivity and reduce costs at the same time, resulting in lower average costs than small hospitals (Kittelsen *et al.*, 2018). Interestingly, because larger hospitals are more specialized, it is argued that they will result in higher quality, but higher quality can also increase the costs, which again results in hospitals being required to contain their costs by using less resources or cutting down services. This study will focus on returns to scale, the relation between output and input. If for example this study found increasing returns to scale, when an increase in output is proportionally larger than increase in input, then it would be consistent with the increased policy focus on merging hospitals together and creating larger and more specialized hospitals in the country. Being able to confirm that it is wise to have larger hospitals would demonstrate that hospital merges are useful and can result in efficiency gains.

There are some literatures that mention the theory about a u-shaped relationship between costs and quality. One study from Denmark, by Hvenegaard *et al.* (2011), looked at hospital departments in order to evaluate the relationship between costs and quality. In this article there is presented some theory on the u-shape curve and the relationship between costs and quality. It is described that the u-shape curve represents net costs that are connected with quality at different levels. Figure 1 on page 5 shows a downward sloping curve called net costs and is the sum of the two other curves: investment cost for prevention of poor quality and costs because of poor quality. It is possible to reduce the net cost of treatment by investing in improving quality when the quality level is low. When the u-shaped curve is zero, we can say that the treatment cost is minimized. Quality on the other hand is maximized when investments in prevention is as high as possible. Therefore, a costs-quality trade-off will exist when there is a choice between minimizing costs and maximizing quality at the same time.

Figure 1: The Theoretical Relationship between Costs and Quality (Hvenegaard *et al.*, 2011)



High quality, on the upward slope of the u-shape of the net cost curve, will always have higher (opportunity) costs and an efficient hospital will in reality have a trade-off between costs and quality. In addition, if a hospital is seen as being efficient in their production of services there will most likely exist a trade-off when it comes to costs and quality (Kittelsen *et al.*, 2015a).

1.3 Previous studies

Hospital efficiency and quality have been subjects in many research papers over the years using frontier estimation techniques such as Data Envelopment Analysis (DEA) or Stochastic Frontier Analysis (SFA). A previous study with similar objective has been done between the Nordic countries, a collaboration between Norway, Sweden, Finland and Denmark. The study from 2015, called "Cost and Quality at the Hospital Level in the Nordic Countries" by Kittelsen *et al.* (2015a) focused also on costs and quality at the hospital level where one main objective was to discover if there were any quality-cost trade-offs by looking at differences in quality. Mortality, readmissions and patient safety were the case-mix adjusted quality indicators used. Using DEA, they found a significant difference in productivity at both national and hospital level as well as room for improvement when it came to the used quality indicators (Kittelsen *et al.*, 2015a). When looking at productivity and inpatient readmissions they found that there was a significant trade-off, but only just, and hospitals with high 30-day mortality had higher costs compared to other hospitals. However, they did not find any clear cost-quality trade-off. They did on the other hand find that Norway had the lowest 30-day mortality as well as lower mortality at most of their hospitals and low mortality is associated with higher productivity.

A Norwegian article from 2017 called "Kvalitet og produktivtet i norske sykehus" (English: Quality and productivity at Norwegian hospitals), by Kittelsen *et al.* looked at precisely quality and costs at the hospital level in Norway. They found that Norway had a higher quality level due to a lower mortality rate compared to the other Nordic countries in the study. However, the readmission rate in Norway was high and productivity was lower than in Demark and Finland. Another significant finding was a positive covariance between the quality indicators for mortality and productivity at the Nordic hospitals, but there was not found any significant correlation in the Norwegian data (Kittelsen *et al.*, 2017).

Another study focused on the relationship between size and quality at Norwegian hospitals and found no statistically significant relationship for the used quality indicators (Erichsen, 2016). However, this study used correlation as chosen method. Instead Erichsen claims that other studies have found increased costs and less quality and efficient services when merging hospitals together. Anthun, Kittelsen and Magnussen (2017) analyzed productivity growth, case mix and optimal size of Norwegian hospitals, and found that mean productivity increased by 24.6% between 1999-2014 with a yearly average increase of 1.5%. After the ownership reform hospitals became larger and productivity increased. On the other hand, similar productivity growth did not seem to be found during later hospital mergers.

A doctoral thesis called "Productivity in the Norwegian hospital sector" (Anthun, 2017) found that if you compared Norwegian hospitals with each other, then the technical efficiency was found to be just as high as it was in Finland. Finland was in total most efficient. The productivity differences can be explained by quality differences. In addition, it was discovered that Norway had lowest rates of mortality, but there was not revealed any clear cost-quality trade-off. Hospitals that had high 30-day mortality also had higher costs. Statistics Norway published a paper in 2018 that discussed big difference in cost usage between the hospitals. but this was explained by the geographical differences and the diverse patient groups (Statistics Norway, 2018c). However, they found no significant association when it came to high average usage and mortality and therefore could not conclude that hospitals with high consumption produced better patient outcomes.

Schreyögg and Stargardt (2010) investigated the relationship between hospital costs and health outcomes for acute myocardial infarction in Veterans Health Administrations. The results suggested that there did exist a trade-off between costs and outcomes. There was a negative association between costs and mortality, meaning that the hazard of dying increased when less money was spent. This supported the hypothesis saying that increased resource input used for the patients leads to better outcomes.

Hussey *et al.* (2013) reviewed studies which focused on the direction of the association between costs and quality. Findings showed significant low to moderate association, regardless if the direction was positive or negative. From the article by Hvenegaard *et al.* (2011) the results indicated that they found some evidence that supported a u-shaped costs-quality relationship when it came to mortality, meaning they found a cost and quality trade-off which indicates that taking quality into account can explain efficiency differences. High mortality was therefore linked to lower costs at Danish hospitals, which is the same as low quality is linked to lower costs, which is a trade-off between costs and quality.

Giancotti *et al.* (2017) performed a DEA in order to study scale efficiency and optimal size of the hospital sector for 45 years (1969-2014) based on journals by the Social Sciences Citation Index. The results from the studies indicated that both the size of the hospitals and the offered output mix influenced the efficiency level. Evidence of economics of scale for hospitals was shown in terms of 200-300 beds, meaning diseconomies of scale would occur if beds were below 200 or above 600. In addition, the results supported the policy regarding creating larger hospitals and closing down smaller hospitals.

A study on scale and quality between Nordic hospitals by Kittelsen *et al.* (2018) used quality indicators such as emergency readmission within 30 days and mortality within 30 days and did not find any evidence that medical volume effects on quality indicators increased the scale elasticity. One explanation could be that these volume effects are limited to only a few patient groups or even offset by other patient groups where quality could be reduced by volume. It was also found that increasing hospital size can potentially reduce costs per treatment, without improving or sacrificing quality

Cost efficiency has been a studied subject as well. Linna, Häkkinen and Magnussen (2006), investigated cost efficiency across 47 Finnish hospitals and 51 Norwegian hospitals. Outputs used were outpatient visits, day care and inpatient days, admissions based on DRG system, while input was net hospital operating costs. Chosen method was DEA, and findings showed that Finnish hospitals were 17-25% more efficient than Norwegian hospitals. On the other hand, the cost at Finnish hospitals varied a lot. Kittelsen *et al.* (2015b) analyzed productivity differences between Denmark, Finland, Norway and Sweden. Productivity was decomposed into technical efficiency, scale efficiency and country specific possibility sets (technical frontiers). Data included costs and patient discharges based on DRG. Using both DEA and SFA, they found small differences in scale and technical efficiency between the four countries. However, there existed large differences in production possibilities. For Norway, the technical efficiency score was 89.7%.

At last, Medin *et al.* (2011) studied cost efficiency of university hospitals in the Nordic countries using DEA. Cost efficiency at the hospitals decreased when variable for specialized university hospital was added. It was found that for teaching and research model, the average cost efficiency for Norway was 0.89 for constant returns to scale and 0.93 for variable returns to scale.

1.4 Relevance of the study

Costs and quality at the hospital level is a topic that has been studied in several western countries over the years. This is because the health sector is constantly growing in this part of the world and the national budget is being affected by the massive costs being used in order to keep up with the political pressure as well as the preferences and needs of the population. The focus on reducing costs can in worst case scenario affect the quality at the hospitals when they operate efficiently. Therefore, more focus on incentives connected to the relationship between costs and quality is needed.

The association between costs and quality is often discussed as whether health care spending will impact quality negatively or whether an improvement in quality will decrease health care costs. It is worth to increase costs if quality is superb in relation to costs. Studies have shown that more information about the relation between costs and quality is needed because of conflicting results. This study focuses only on Norway and will be relevant in the way that it can give important information about costs and quality across the Norwegian hospitals. It will give valuable information which can clarify if there is a significant trade-off between the hospitals when it comes to quality and costs and help identify sources to inefficiency. Cost efficiency is an important goal for hospitals, and since other studies have analyzed cost efficiency between several countries, such as Nordic countries, this study will only focus on Norway and 22 hospitals as well as more years.

In addition, merging hospitals together has been a widely discussed topic in Norway, which has resulted in shutting down smaller hospitals over the years and creating larger and more specialized hospitals as an argument to increase productivity and lower costs. Larger firms are often seen as more efficient than smaller firms because of economies of scale (increasing returns to scale). On the other hand, firms can become too large and experience diseconomies of scale. However, there does not exist a lot of empirical analysis that presents evidence to support optimal hospital sizes. With the data used in this thesis, scale elasticity will be measured as well to see if there is decreasing returns to scale (increase in inputs leads to a less than proportional increase in output – good idea to be a small hospital) or increasing returns to scale (when output increases by a larger proportion than the increase in inputs – good idea to be a large hospital).

This thesis will be a part of the project related to paragraph 2.3.1 Methods and data-productivity growth from the project "The effects of DRG-based financing on hospital performance: productivity, quality and patient selection", conducted by researchers from the Department of Community Medicine at NTNU, Frisch Centre and the Group for Health Services Research at SINTEF Technology and Society. One of the aims of this project is to measure the hospital performances in Norway in order to see how efficient they are. Data envelopment analysis and stochastic frontier methods are seen as the most suitable methods to use when measuring productivity and efficiency. Therefore, it will be interesting to attain more knowledge about whether there are any tradeoffs and differences between hospitals in Norway when it comes to costs and quality and if this can explain any differences in productivity at the hospital level.

Chapter 2 DATA

Chapter 2 describes the data used throughout the study. This includes description of cost data, patient-level data and the quality indicators used to measure performance level for 3 models across the Norwegian hospitals.

2.1 Cost data

The cost data is provided by SAMDATA database of Norwegian specialized care, which is produced annually by the Directorate of Health. These costs include all costs related to production within the hospitals, such as operating costs. Cost that are not included are value added tax (VAT), capital costs, ambulance costs, purchased care and costs for teaching and research (Kittelsen *et al*, 2015a). Costs were already deflated.

2.2 Patient-level data

The data used in this study is provided from the Norwegian Patient Register, owned by the Norwegian Directorate of Health. The data is a 7-year period, from period 2008 to 2014, and includes 22 hospitals. Note that there does exist more hospitals in Norway, but several are merged together within a HF resulting in several hospitals being placed under the same name in this dataset. This means that data for some hospitals were aggregated and utilized together.

2.2.1 Diagnosis-Related Group

The diagnoses-related group (DRG) system is a patient classification system that classifies patients in groups that are similar when it comes to medical and resource use. The Norwegian DRG system is based on the Nordic system called NordDRG. When patients are admitted to a hospital, they are given a DRG weight based on variables such as the individual´s diagnosis, treatment, gender, age and discharge status. The DRG weight gives information about how resourceful the patient is in comparison to the average patient registered in the system. Since DRG gives information about resource use and activity level it makes it possible to compare hospitals despite that the institutions specialize and treat very different patient groups. A hospital will receive a price based on a patients DRG and if the hospital is able to treat the patient below the price, they keep the profit. If they spend more than the price, they lose money. The higher the activity, the greater reimbursement is given to the hospital. In this study, the data contains 758 DRGs which were used in order to measure the performance of the hospitals for the 3 models.

2.3 Quality indicator

In order to measure the performance and discover any links between costs and quality at the hospitals, one must first find an appropriate quality indicator. Quality is a central word in this thesis, which refers to better health outcome. Improving quality can either reduce the resource use because errors are discovered, or it can result in a larger resource use due to improving health procedures and services. Quality indicators, or performance indicators, are often used as means of evaluation and are useful tools that can provide a lot of valuable information for health care providers, health personnel, public health policy makers as well as health consumers. They say something about the quality of the area that is being studied, such as a hospital department.

Quality will be measured in terms of clinical quality such as mortality. Mortality is an appropriate quality indicator to use because hospitals always wish to keep the mortality rate low and there are available data on mortality from large patient groups. Furthermore, a high mortality rate can be seen as delivering services of poor quality and this says something about the effectiveness and safety. Using mortality can eventually show us variations in healthcare services, such as treatment and processes at the hospitals. If mortality is high there may be an association between quality and costs because patients who die could be more costly (Carey and Burgess 1999, cited in Hvenegaard *et al.*, 2010). Nonetheless, mortality is widely used and accepted as a quality indicator when performing an analysis such as this, which can be seen in several hospital studies.

Mortality is a popular and suitable quality indicator, but it does come with some disadvantages and must be interpreted carefully. Even though we adjust for patient difference (case-mix adjustment) there still can exist systematic differences between the hospitals which can affect the findings in some way (Kittelsen *et al.*, 2017). This can for instance happen when a large hospital, such as a university hospital in Norway, treats a large portion of severly ill patients which can lead to the hospital having a higher mortality and readmission than for example a local hospital. Controlling for differences is essential in order to complete the analysis and get estimates that represent reality.

This thesis will estimate five quality indicators to calculate performance indicators in order to measure the performance level across the hospitals. They are measured out from information from the Norwegian patient register. Note that cause of death will not be known in this study. We only know the health issue resulting to the patient's admission to a hospital. If a patient is transferred across several hospitals, it will be the last hospital the patient was registered at that is "responsible" for the patient´s death. The main focus will however be only on two performance indicators during the result and discussion section. Emergency readmission within 30 days and out of hospital mortality within 30 days. Emergency readmission within 30 days as an inpatient is used because it is seen as representing poor medical quality, because it could for instance be a sign of initial treatment not being sufficient enough which can cause a readmission due to complications (Kittelsen *et al.*, 2018). Out of hospital mortality within 30 days is as mentioned a well-accepted quality indicator and low morality is always a good sign. In fact, Ross *et al.* (2010) found that hospitals with high volume seemed to have lower 30-day mortality for medical conditions. If the ratio of the indicators are high, then it implies low quality.

The first time a patient is registered at a hospital across a year is the only one that counts, which means if they are registered several times, we only view this as 1 patient. The data can include multiple visits by patients, but if the referral date is different it will be viewed as a different diagnosis. In order to avoid multiple visits that may be a result of for example a checkup, patients will as a result be registered with one single visit if they visit hospital several times. Patient mortality date is provided by automatic linkage between patient registry and National Population Registry.

Chapter 3 METHODS

The chapter includes the methodology used in the study. They consist of case-mix adjustments of the quality indicators in order to find the performance measure across the hospitals and stochastic frontier analysis used in order to measure inefficiency and see if a trade-off between costs and quality exists.

3.1 Case-mix adjustment

Comparing performance levels across the hospitals may cause problems because of the diversity of patient groups. Differences in case-mix can be defined as systematic variation that can be explained by patient characteristics, such as gender, age as well as severity of illness and treatment. Case-mix adjustments is essential in order to account for all systematic differences between the hospitals because it captures patient characteristics including illness that in some way affects the outcome (Kittelsen *et al*., 2015a). If differences are not taken into account, it will be difficult to know if the quality indicators are a result of some other factor than differences in quality. One hospital can have a larger group of critically ill patients than other hospitals, and if this is not adjusted for the hospital may end up with weaker results on some of the quality indicators despite that the services are not of poorer quality.

In order to control and capture these differences, we use DRG to help describe the activity level at the hospitals. The data incudes 758 DRGs used in the case-mix adjustment. Adjusting for these differences makes it possible to accept comparison of treatment with a mix of different patients taking into account variables such as gender, diagnoses, severity of disease etc. across all the 54 hospitals. The quality indicators will be case-mix adjusted in 3 models. Model 0 consists of DRG, Model 1 includes DRGs as well as patient characteristics and Model 2 represents the treatment variables along with the other two components. Chapter 4 describes the variables more thoroughly. The inclusion of treatment variables in model 2 will be the main focus throughout this study. This is because in the article by Kittelsen *et al.* (2015a) they found statistical evidence that seemed to favor model 2 and it therefore seems reasonable to use the model in this study as well.

The formula used for case-mix adjusted performance measures, based on Ash *et al.* (2003) cited in the article by Kittelsen *et al.* (2015a), calculates observed-to-expected ratio for each quality indicator for every hospital. For each of the 3 models, $m \in (0,..2)$ the expected values and performance measure is calculated. Each patient i has an observable quality indicator, defined as ω_{ihk}, as well as an expected quality indicator, $\widehat{\omega}_{ihk}^{m}$, when the patient is registered at a hospital $h \in (1, \dots H)$ and given their DRG $k \in (1, \dots K)$.

The case-mix adjusted hospital performance measures P_h^m, are calculated by adding together all observed patient outcomes followed by dividing by the sum of all expected patient outcomes. This is defined in the following equation:

$$P_h^m = \frac{\sum_{k=1}^{K} \sum_{i=1}^{N_{hk}} \omega_{ihk}}{\sum_{k=1}^{K} \sum_{i=1}^{N_{hk}} \widehat{\omega}_{ihk}^{m}} \tag{1}$$

where P_h^m represents the performance indicator for hospital h in model $m \in (0,..2)$ and N_{hk} represents the number of patients within DRG k at hospital h. A low value suggests better quality for the quality indicators, and this also applies for the performance measure, P_h^m. When predicting $\widehat{\omega}_{ihk}^{m}$ for the DRG model, $m = 0$, we take into account that the DRGs are of different composition for each hospital. The predicted quality indicator for a patient, $\widehat{\omega}_{ihk}^{0}$, is estimated as the average value of the quality indicator within each DRG for all the patients across the hospitals (Kittelsen *et al.*, 2015a). The formula used to estimate predicted outcomes for model 0 is:

$$\widehat{\omega}_{ihk}^{0} = \frac{\sum_{h=1}^{H} \sum_{j=1}^{N_{gk}} \omega_{jgk}}{\sum_{g=1}^{H} N_{gk}} \tag{2}$$

and is independent of both patient i and hospital h and equal for all the patients in DRG k.

For the predicted quality measure, $\widehat{\omega}_{ihk}^{m}$, it is possible to condition on patient characteristics. The conditional probability is calculated using a logit model, which seems appropriate because all five quality indicators are binomial variables (Greene 2000; Hosmer *et al.* 2013, cited in article by Kittelsen *et al.* 2015a). Due to many observations, the expected value can be calculated as the predicted value based on the maximum likelihood estimation, as shown in the formula below:

$$\omega_{ihk}^{m} = \frac{e^{\beta_{0k}^{m}+\boldsymbol{\beta}_{k}^{m}\mathbf{z}_{ihk}^{m}+\varepsilon_{ihk}^{m}}}{1-e^{\beta_{0k}^{m}+\boldsymbol{\beta}_{k}^{m}\mathbf{z}_{ihk}^{m}+\varepsilon_{ihk}^{m}}} \tag{3}$$

where ω_{ihk}^{m} represents the quality measure for patient i in DRG k at hospital h; the two coeffiencient vectors β_{0k}^{m} and $\boldsymbol{\beta}_{k}^{m}$ are specific to each DRG k and model $m \in (0,\ldots 2)$; the vector $\boldsymbol{z}_{ihk}^{m}$ represents the variables for individual case-mix adjusting; and at last ε_{ihk}^{m} signifies the error term, assumed here to be distributed normally. Three different models are estimated based on each DRG and the five quality indicators. For model 1 the patient characteristics are the explanatory variables that are captured by z and for model 2 the vector includes patient characteristics as well as treatment variables.

After collecting predicted and observed quality indicators, the performance indicator per hospital and year is obtained by dividing observed quality indicator on predicted quality indicator. It is recommended that predicted is higher than observed because that will result in hospitals having better efficiency and a low performance indicator, which represents better quality.

3.2 Efficiency measurement

Efficiency will be a term that is essential in order to understand the goal of this study. Productivity and efficiency are two concepts that are used interchangeably because they are somewhat closely related. Productivity can simply be defined as how much is created, the ratio of the outputs produced, and the inputs used (Coelli, Rao, O´Donnell & Battese, 2005). Efficiency, often seen as a normative concept, can be described as the maximum level of output based on minimum level of inputs. A productive hospital will for example have high productivity, but for an efficient hospital it will often be related to the amount of work effort (Kjekshus, 2000). In this case, if a hospital is not efficient it simply means that the inputs are not being used efficiently. Reducing inefficiency, such as cost inefficiency, can result in better cost containment but can also be a result of reducing the number of services or quality of the services (Rosko and Mutter, 2008). Productivity can be broken down into differences in technical efficiency, costs efficiency and scale efficiency. This thesis will focus on cost efficiency.

There are two analysis that are capable of estimating efficiency (inefficiency) and in order to do so a frontier is needed. They are called data envelopment analysis (DEA) and stochastic frontier analysis (SFA). Both these methods have their own strengths and weaknesses and measure efficiency differently. DEA has become a popular tool when estimating efficiency in healthcare. The nonparametric method requires no functional form because it assumes no measurement error. Assuming no measurement error is the main difference from SFA. All observations are feasible, and the method can therefore handle multiple outputs and inputs. The efficiency frontier is formed by the best practice units and points that are not on the frontier are considered inefficient. However, the parametric approach SFA is the chosen method in this thesis to estimate the approximate efficiency given the hospital data for the different quality indicators. Since we are looking for a trade-off between quality and costs, a costs frontier is a good way to uncover any trade-off.

Average cost efficiency for the hospitals during 2008 to 2014 will also be estimated using SFA. SFA seems appropriate to use rather than DEA, which is a popular method to use to estimate efficiency, because SFA identifies effectiveness of the description of the model and can decompose deviations from efficiency into random noise and inefficiency. A firm is efficient if resources are used right. A hospital is not cost efficient is they produce output at a higher cost level than those predicted by the cost function (Jacobs *et al.* 2006). Any point on the cost function is cost efficient. Cost efficiency can in general be explained as the ratio of costs to output. A more economic expression shows that it is technical efficiency combined with allocative efficiency $CE = ACxTE$. Cost efficiency needs to be a value between 0 and 1 (0$< CE < 1$).

In Stata, the cost elasticity of the output are the DRG-point coefficients. Hirschey (2009) defines cost elasticity as the "percentage change in total cost associated with a 1 percent change in output" and it says something about the effect of cost on efficiency. When estimating the elasticity of scale using the Cobb-Douglas cost function, the four DRG-point coefficients (the betas) need to be added together. Elasticity of scale is found by taking 1 and dividing by the sum of the four betas. An elasticity of scale value equal to 1 suggests constant return to scale because costs and output increase proportionally. If the value is larger than 1, cost increase less than output and increasing returns to scale is present. A value less than 1 demonstrates decreasing returns to scale because costs increase more than output and firms get less product of each input (or more cost of each output).

3.2.1 Stochastic Frontier Analysis (SFA)

Stochastic frontier analysis (SFA) was first developed by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). In 1989, Wagstaff was the first to perform an SFA study of a health care organization, examining 49 Spanish hospitals (Rosko and Mutter, 2008). Unlike DEA, this method assumes a functional form and uses maximum likelihood estimation to estimate a cost function. The method is capable of separating both random stochastic error and efficiency from the residual/error term. (Jacobs, 2001). A random error can be defined as an uncontrollable problem at a hospital that affects the output variable.

When estimating SFA for panel data there is a choice to either estimate a production function or a cost function. Kumbhakar and Lovell (2005) present a good introduction of the production function in SFA, which was first proposed by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). Theory usually starts by presenting firms as being able to maximize production, minimize costs and maximize profits. A production function of a firm can be written as $f(z_i;\ \beta)$, where z_i, represents the vector of inputs and β is the vector of technology parameters. The function represents the maximum output that the firm is able to produce, also defined as technical maximum.

This is not always true for every firm, and deviations can occur, such as inefficiency. Inefficiency will result in firms not being able to optimize as they wish even though input is the same. Therefore, $q_i \leq f(z_i;\beta)$ and ratio $\frac{q}{f(z_i;\beta)} \leq 1$ is defined as technical efficiency $(0 \leq TE \leq 1)$. Often technical efficiency $(TI = 1 - TE)$ is described as *"...percentage shortfall of output from its maximum, given the inputs."*, cited by Parmeter and Kumbhakar (2014). This is seen as important because inequality $q \leq f(z_i;\beta)$ expressed as logarithm, $lnf(z_i;\beta) - u_i$, where $u_i > 0$ represents the technical inefficiency (Parmeter and Kumbhakar, 2014). The equation for technical efficiency for a production function is, as presented by Coelli *et al* (2015), given as:

$$TE_i = \frac{q_i}{\exp(z_i\beta + v_i)} = \frac{\exp(z_i\beta + v_i - u_i)}{\exp(z_i\beta + v_i)} = \exp(-u_i) \qquad (4)$$

where q_i is the output of firm i; z_i represents the inputs; β is a vector of unknown parameters; the variable u_i is non-negative and associated with technical inefficiency and exp represents the exponential. TE takes a value between 0 and 1. Since we are using a cost function, technical efficiency will be the same as cost efficiency, and therefore costs efficiency will be the used term.

If a firm produces less due to inefficiency the term for frontier production will look like:

$$q_i = f(z_i;\ \beta)\varepsilon_i \tag{5}$$

$$\varepsilon_i = v_i - u_i \tag{6}$$

where ε_i represents the level of technical efficiency for the firm i and takes a value between 0 and 1. v_i represents random noise, which can be positive or negative, while u_i represents the inefficiency. If ε_i = 1, then the firm has achieved maximum output, while if ε_i < 1 then they are not making the most out of their inputs z_i given their technology in the function. Since output is assumed to be positive ($q_i > 0$), then the degree of technical efficiency will also be positive ($\varepsilon_i > 0$) (Statacorp, 2017). Random shocks can affect the firms output, which can be written as:

$$q_i = f\left(z_{i,}\beta\right)\varepsilon_i \exp(v_i) \tag{7}$$

where v_i represents the random noise that can affect output.

Given the data, a cost function will be estimated. A cost function seems more suitable to work with because cost can be estimated and can handle multiple outputs. Estimating a production function can cause some difficulties if there are several outputs. The cost function can help identify the relationship between output level and costs for a firm and is an essential tool in order to analyze efficiency. A cost function is identical to a production function when assuming cost minimizing behavior and the frontier will measure how far the firm is from full-cost minimization (Jacobs *et al.*, 2006). In Coelli *et al.* (2005), the expression for cost-minimization is written as:

$$c(w, y) = \min_x w'x \tag{8}$$

where $w = (w_1, w_2, \ldots, w_n)$ represents the input prices; y is the output and $\min_x w'x$ represents a combination of input and output in order to get the minimum cost of input that produces the output. Cost function can be written as $c = f(y)$, where c is costs and y signifies the output. If a firm has multiple outputs and inputs the cost form can be written as $c = f(w, y)$, where c represents costs; *w* is the input price and y is the output. It represents the minimum costs in order to produce output given the input price. An advantage of using SFA is that not only is technical inefficiency in focus, but also cost-inefficiency. Cost efficiency, the ratio of minimum cost to observed costs, can be expressed as:

$$CE_i = \exp(-u_i) \tag{9}$$

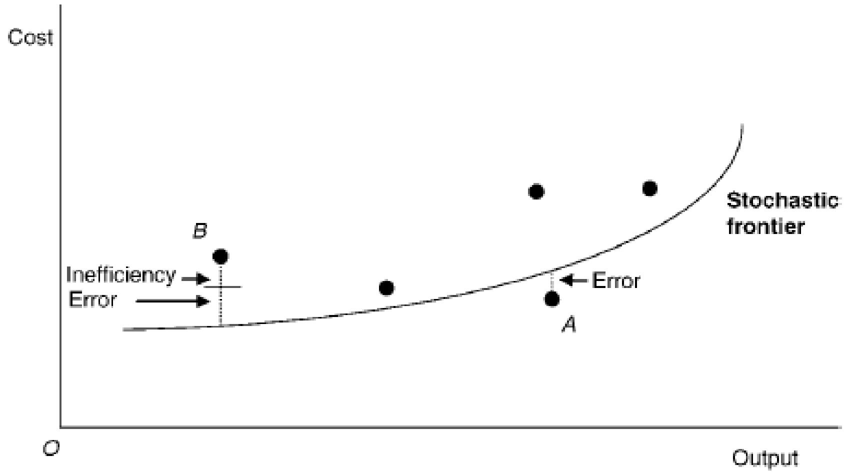
One main reason why SFA was chosen was because it is capable of separating inefficiency from random stochastic error/noise in the residual, as seen in figure 2 on page 21, something that is not possible with DEA. Coelli et al. (2005) states:

"*That statistical noise arises from the inadvertent omission of relevant variables from the vector x, as well as from measurement errors and approximation errors associated with the choice of functional form."*

A cost frontier takes the performances of the firms and weighs them relatively to the best that can be economically achieved. Observation A in the figure represents the random stochastic error term/noise v_i which is found below the stochastic frontier. v_i is able to capture any factors that are not in control of the hospital. Observation B is recognized as inefficiency u_i and random error v_i and is placed above the cost frontier. u_i is a non-negative variable that represents the technical inefficiency at the hospitals.

When estimating the cost frontier, the error term ε_i is decomposed into inefficiency u_i and random stochastic noise $v_{i,}$ with zero covariance (Jacobs *et al.* 2006). This simply means that random stochastic noise and inefficiency will affect the performance at the hospitals. As a result, SFA can end up calculating a higher average cost efficiency.

Figure 2: Stochastic Frontier (Jacobs et al. 2006, p.54)



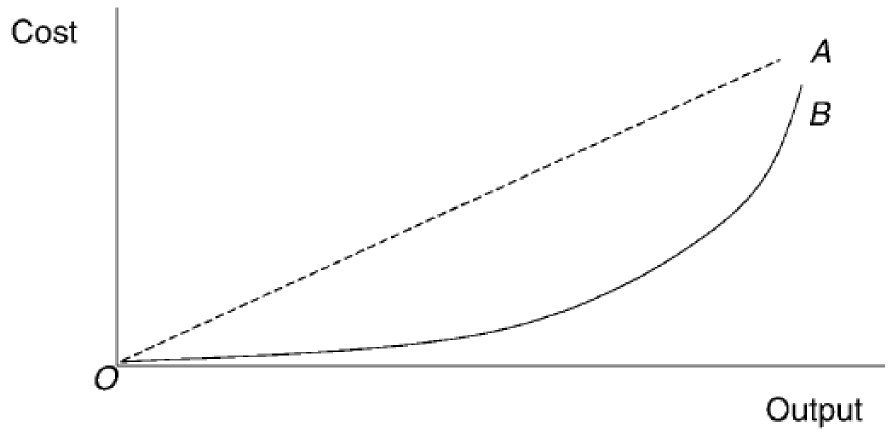
As mentioned by Coelli (1996), in order to compose the stochastic frontier cost function, the error term specification for production frontier $(v_i - u_i)$ must be transformed into $(v_i + u_i)$. This creates cost function expressed as:

$$c_i = c(w_i, q_i) + (u_i + v_i) \tag{10}$$

where c_i is the cost for a firm; w is a vector of input prices; q is the output vector, while the last parenthesis includes the error term. Furthermore, there is a choice on how to transform the variables. They can either be used in their natural units or transforming them into logarithmic form (Jacobs *et al*, 2006). Assigning a functional form to the cost function will affect the flexibility of the curve and is essential in order for the data to fit the model in the best way. Choice of form will also affect the performance of the hospitals on the efficiency frontier. Using natural units implies that there will be a linear relationship between dependent variable and explanatory variable (Jacobs *et al*, 2006).

If for example it is assumed a linear relationship between the number of treated patients and costs, this would mean that the marginal cost (MC) for each treated patient are the same. However, costs changing in a constant rate may not be a reality, so assuming a logarithm functional form can solve this. Functional form also helps to deal with heteroscedasticity, which is when random error does not have a constant variance. Coefficients are then transformed from natural units to either elasticities of percentage change (Jacobs *et al*, 2006). Figure 3 below shows two lines, where A represents the natural functional form and line B is the logarithmic functional form.

Figure 3: Natural vs. logarithmic Functional Form (Jacobs *et. al*, 2006, p.44)



It is, however, important to see the cost function in relation to scale properties (Jacobs et al., 2006). One common functional form in SFA is Cobb-Douglas, which is log linear and uses a logarithmic form. It is then assumed that the hospitals have the same scale elasticities. Due to this we shall assume a Cobb-Douglas form. According to Coelli *et al*. (2005), the Cobb-Douglas cost frontier can be written as:

$$\ln c_i = \beta_0 + \sum_{n=1}^{N} \beta_n \ln w_{ni} + \sum_{m=1}^{M} \phi_m \ln q_{mi} + v_i + u_i \qquad (11)$$

where ln denotes the logarithm, β_0 and β_n are parameters to be estimated, w_{ni} is the n-th input price, q_{mi} is the m-th output, ϕ_m refers to the probability density function.

The function will be non-decreasing, homogenously linear and concave if β_n is non-negative and satisfy following:

$$\sum_{n=1}^{N} \beta_n = 1 \tag{12}$$

When taking logarithm of both sides of the function, we get linearity. In this case the parameters are transformed into elasticities, not natural units. The Translog function is another functional form which includes cross products. It is praised for being flexible but can be seen as more complex to use than the Cobb-Douglas form. Cobb-Douglas on the other hand is able to save degrees of freedom. Since Cobb-Douglas will be used it should be recognized that this functional form will affect the inefficiency effect and estimated costs when working with the data.

When performing SFA, a regression for average cost function will be calculated, where three different types of inefficiency terms can be analyzed. They are called exponential, truncated normal and half normal. All three are examined in Stata with the panel data.
For exponential, u_i is distributed independently exponentially with variance σ_u^2.
For half-normal, u_i follows a half normal distribution and distributed as $N + (0, \sigma_u^2)$.
Finally, for the truncated normal model, u_i is independently $N + (\mu, \sigma_u^2)$ distributed with truncation point at 0 (Kumbhakar and Lovell, 2000).

A firm, or in this case a hospital, is efficient when $u = 0$ and inefficient if $u > 0$. This thesis assumes a half-normal distribution. Assumptions regarding this distribution are the following: random stochastic noise v_i is assumed to be independent with normal distribution with zero mean and zero variance, $v_i \sim iidN(0, \sigma_v^2)$. Inefficiency u_i is represented having a half-normal distribution, $u_i \sim iidN^+(0, \sigma_u^2)$ (Coelli *et al.*, 2005).

Jacobs *et al.* (2006) defines the expected mean value of inefficiency for the half-normal distribution, where the residual is conditional upon the composite residual, which can be expressed as:

$$E(u_i|\varepsilon_i) = \frac{\sigma\lambda}{(1+\lambda^2)}\left[\frac{\phi(\varepsilon_i\lambda/\sigma)}{\Phi(-\varepsilon_i\lambda/\sigma)} - \frac{\varepsilon_i\lambda}{\sigma}\right] \tag{13}$$

where $\phi(.)$ represent the probability density function and $\Phi(.)$ represents the cumulative distribution function of normal distribution; the total error variance is defined as $\sigma_s^2 = \sigma_v^2 + \sigma_u^2$ and the contribution of the inefficient component is $\lambda = \frac{\sigma_u}{\sigma_v}$. If $\lambda = 0$ there exists no technical inefficiency and all observations are placed on the frontier (Coelli *et al.*, 2005). Any deviation will be a result of noise.

Chapter 4 VARIABLES

This chapter presents all the chosen variables that will be included in the analysis when performing stochastic frontier analysis in STATA 15.

4.1 Case-mix adjustment variables

Table 1: Patient characteristics and treatment variables

Group	Variable name	Description of variable
Model 0: DRG	DRG	Diagnosis-related group
Model 1: Patient characteristics	+ Male agegrp0 agegrp1_9 agegrp10_19 agegrp20_29 .... agegrp80-89 agegrp90	0= Female, 1=Male Age group of 0 Age group 1 to 9 Age group 10 to 19 Age group 20 to 29 (Age groups 30-39, 40-49, 50-59, 60-69, 70-79) Age group 80 to 89 Age group 90 and above
Model 2: Treatment variables	+ TransinOwnHospital TransinOtherHospital TransoutOwnHospital TransoutOtherHospital Charlson NumSec Diagnses	Dummies for transfer in and out of hospital department. Stay within one day before or after this stay. Not original coding but calculated from dates of patient registry directly. Charlson index based on secondary diagnosis Number of secondary diagnoses

Table 1 above includes all 3 models that were developed in order to find the case-mix adjusted hospital performance indicator. Most of the included variables are retrieved from the individual patient data. The number of DRGs used were 758, with the purpose of capturing patient differences that may affect the cost. The patient characteristics include gender as well as age groups. The age groups are in 10-year groups, from age below 1 up to age above 90. Although the treatment variables are partly endogenous, they are still permitted to alter for risk because they can reflect severity. The *transout* and *transin* variables represent transfer in and out of hospital or department and are coded in order to explain the movement of the patients (where the patient came from and where they went next). Note that movements such as to/from home, a health clinic center or a nursing home (or non-hospital) are not needed in order to utilize the data available. Comorbidity is incorporated in the number of secondary diagnosis and the Charlson index based on secondary diagnosis (Charlson *et al.,* cited in Kittelsen *et al.,* 2015a). Model 2 will be the main model used in the SFA result section.

4.2 Dependent and independent variables

Table 2: Description of dependent and independent variables

Group	Explanation of variable	Variable name	Description of variable
Dependent variable **Hospital costs**	Cost function in ln	lnCosts	Log transformed total hospitals costs in NOK, already deflated.
Independent variables	Mortality 30 days	perf_u0_mort30last perf_u1_mort30last perf_u2_mort30last	Out of hospital mortality, 30 days after admission.
Performance indicators for model 0, 1 and 2	Mortality 90 days	perf_u0_mort90last perf_u1_mort90last perf_u2_mort90last	Out of hospital mortality, 90 days after admission.
	Mortality 180 days	perf_u0_mort180last perf_u1_mort180last perf_u2_mort180last	Out of hospital mortality, 180 days after admission.
	Morality 365 days	perf_u0_mort365last perf_u1_mort365last perf_u2_mort365last	Out of hospital mortality, 365 days after admission. Calculated for model 0, 1 and 2.
	Readmission 30 days Emergency	perf_u0_readm30_emgc perf_u1_readm30_emgc perf_u2_readm30_emgc	Patients admitted acutely to inpatient hospital care within 30 days of discharge
DRG point	Emergency	lnv_emrg_p	Log transformed emergency patients in DRG point
	Elective	lnv_elective_p	Log transformed elective patients in DRG point
	Day patients	lnv_day_p	Log transformed day treatment patient in DRG pint for patients that only received day treatment without overnight stay

	Outpatients	lnv_out_p	Log transformed outpatient in DRG point where patients only received care without surgery and overnight admittance

The case-mix adjusted performance indicators for model 0, 1 and 2 were calculated by dividing observed quality indicator by the predicted quality indicator. The variable names are written without u# in the result chapter.

4.3 Other variables included in SFA

The table below contains the statistics that are included in the SFA result report, following a half-normal model, in Chapter 5.

Table 3: Other outputs used in SFA

Variable name	Description of variable
sigma_v	Standard deviation of component v, σ_v
sigma_u	Standard deviation of component u, σ_u
lnsig2v	Log likelihood variance v, $\ln \sigma_v^2$
lnsig2u	Log likelihood variance u, $\ln \sigma_u^2$
sigma2	Total error variance $\sigma_s^2 = \sigma_v^2 + \sigma_u^2$
lamda	Ratio of the standard deviation of the inefficiency component to the standard deviation to the random stochastic noise component: $\lambda = \frac{\sigma_u}{\sigma_v}$

4.4 Summary of data

Table 4: Summary statistics for variables in SFA (n=154)

Variable	Mean	Std. Dev.	Min	Max
Patients	272,513.3	208,458.8	30,738	1,049,609
readm30_emgc	15,473.16	11,828.78	2,408	62,015
mort30last	995.195	543.524	220	2,231
mort90last	1,292.227	697.75	329	2,886
mort180last	1,436.636	778.975	336	3,504
mort365last	1,602.63	903.75	385	4,598
perf_u0_readm30_emgc	1.023	0.109	0.711	1.306
perf_u0_mort30last	1.049	0.146	0.616	1.462
perf_u0_mort90last	1.055	0.148	0.598	1.495
perf_u0_mort180last	1.058	0.159	0.579	1.501
perf_u0_mort365last	1.057	0.206	0.532	1.732
perf_u1_readm30_emgc	1.014	0.104	0.705	1.297
perf_u1_mort30last	1.001	0.122	0.666	1.369
perf_u1_mort90last	1.009	0.122	0.653	1.369
perf_u1_mort180last	1.011	0.133	0.629	1.455
perf_u1_mort365last	1.011	0.179	0.579	1.676
perf_u2_readm30_emgc	1.008	0.997	0.720	1.312
perf_u2_mort30last	0.967	0.129	0.626	1.345
perf_u2_mort90last	0.978	0.133	0.614	1.345
perf_u2_mort180last	0.982	0.142	0.598	1.429
perf_u2_mort365last	0.983	0.183	0.553	1.648

Table 4 above summarized the mean values of different variables for the period 2008-2014 based on 22 Norwegian hospitals. On average there were 272,513 patients over the duration of 7 years. The average readmission emergency within 30 days was 15,473 patients over the period for 154 hospital observations. For 30 days mortality the average was 995 patients. The 5 first quality indicators (readm30_emgc, mort30last, mort90last, mort180last and mort365last) are the observed patient outcomes. The performance indicators are also calculated for all 3 models, and they are found by dividing observed by predicted given the patient mix of each hospital. The performance indicator values are in ratios and a high value implies that observed outcome is higher than predicted outcome, which means quality is low. Model 2 has lower performance indicator values compared to the other two models, so treatment variables affect the quality. It is always better that predicted is higher than observed outcome in order to get a lower performance indicator. In the result section, there will only be used two performance indicators in the SFA. They are emergency readmission within 30 days and mortality within 30 days.

Table 5: Summary of costs and DRG points in NOK

Variable	Mean	Std. Dev.	Min	Max
costs	3,138,052	2,728,894	589,178	14,251,069
v_emrg_p	27,164.13	17,416.35	6,241.669	83,062.45
v_elective_p	17,450.42	19,423.05	2,592.768	10,5061.7
v_day_p	3,781.943	2,704.384	454.1408	12,617.76
v_out_p	9,353.477	7,898.253	921.163	41,604.05

Table 5 shows the mean for hospital costs and the four DRG points: emergency patients (v_emrg_p), elective patients (v_elective_p), day treatment patients (v_day_p) and outpatients (v_out_p).

Chapter 5 RESULTS

This chapter presents the main estimation results which were calculated using Stochastic Frontier Analysis (SFA) in STATA 15.

5.1 SFA results

When performing SFA, the case-mix adjusted performance indicators for emergency readmission within 30 days and out of hospital mortality within 30 days were included in the analysis along with other independent variables such as DRG point for emergency (v_emrg_p), DRG point for elective (v_elective_p), DRG point for day patients (v_day_p) and DRG point for outpatients (v_ouy_p). The DRG points are included because they are likely to drive the costs at the hospitals. Other quality indicators were also tested, but they gave insignificant results when added with the DRG points, readmission 30 days and out of hospital mortality within 30 days. That is why only two quality indicators are used.

A cost function was performed, and hospital cost was the dependent variable in the analysis. The *frontier* command was executed following a half-normal distribution in STATA 15. The results are obtained with Cobb-Douglas cost function which was used by transforming costs (lnCosts) and the DRG points into logarithmic form, as you can see in table 6. The performance indicators are not transformed into logarithmic form because binomial logic regression was used when they were calculated. As mentioned earlier, when the performance indicators are high, this will indicate low quality. If we find any deviation from the frontier, this will either be explained by being stochastic noise or a result of inefficiency.

Table 6: SFA regression. Dependent variable is lnCosts (Model 0, 1 and 2) (n=154)

	Model 0	**Model 1**	**Model 2**
	DRG	+Patient characteristics	+Treatment variables
Independent variable:			
lnv_elective_p	0.268*** (0.027)	0.284*** (0.031)	0.277*** (0.033)
lnv_emrg_p	0.349*** (0.046)	0.304*** (0.051)	0.293*** (0.048)
lnv_day_p	0.275*** (0.047)	0.240*** (0.049)	0.299*** (0.053)
lnv_out_p	0.037 (0.055)	0.12** (0.056)	0.076 (0.064)
perf_readm30_emgc	-0.246*** (0.087)	-0.183* (0.097)	-0.187* (0.099)
perf_mort30last	-0.401*** (0.082)	-0.236** (0.102)	-0.211** (0.098)
_cons	6.837*** (0.975)	6.406** (3.061)	6.388*** (0.216)
lnsig2v	-4.655*** (0.115)	-4.519*** (0.357)	-5.485*** (0.592)
lnsig2u	-17.014 (11828)	-13.254 (5783.6)	-3.922*** (0.439)
sigma_v	0.975 (0.006)	0.104 (0.019)	0.064 (0.019)
sigma_u	0.000 (1.195)	0.001 (3.828)	0.139 (0.031)
simga2	0.009 (0.001)	0.011 (0.006)	0.024 (0.006)
lambda	0.002 (1.196)	0.013 (3.846)	2.173 (0.048)
Log likelihood	139.919	129.496	129.878
chibar2	0.00	0.00	1.00
Prob>=chibar2	1.000	1.000	1.159

Significant coefficients are marked at 0.10 (*), 0.05 (**) and 0.01 (***) level.

Standard error in parenthesis

An SFA regression was completed with all 3 models in order to get some general information about the models, even though the main focus will be on model 2. Results are presented in table 6 above. The results are based on 154 hospital observations between 2008 to 2014 for 22 hospitals. Model 0 includes DRG, model 1 includes DRG as well as patient case mix such as gender, age, while model 2 includes DRG, patient characteristics and treatment variables. The table gives information about how costs varies for different levels of output.

The independent variables of DRG points in table 6 are significant at 1% level for all three models and they influence the cost of a hospital. The positive coefficients indicate an increase in costs. DRG points are as mentioned transformed into logarithmic values and this suggests that, for model 2, a 1% increase in DRG point for elective treatment will increase total costs by 0.277% when other dependent variables are held constant, for emergency the costs will increase by 0.293%, for day patient treatment costs will increase with 0.299% and for outpatient treatment costs will increase 0.076%. There are no major differences between the DRG points across the three models. That coefficients for DRG points are positive because when the average price of a patient increases, so will hospital costs and therefore a positive association exists between costs and DRG points. Furthermore, the performance indicators were added together along with the DRG points. When looking at the quality indicator for emergency readmission within 30 days (perf_readm30_emgc) and mortality within 30 days (perf_mort30last), there are some differences to notice.

For model 0 the coefficients for emergency readmission within 30 days (perf_readm30_emgc) and mortality within 30 days (perf_mort30last) are significant at 1% level and the coefficients are negative. This indicates that an increase in emergency readmission within 30 days and out of hospital mortality within 30 days will result in lower costs. For model 1 the quality indicator coefficient was also negative, indicating lower costs. Emergency readmission is only significant at 10% level while mortality within 30 days is significant at 5% level. Interestingly, the same goes for model 2. An increase in mortality within 30 days across Norwegian hospitals for model 2 results in lower costs and is significant at 5% level. Emergency readmission is statistically significant at 10% level, so there exists stronger evidence in favor of the performance indicator for mortality. When the performance indicators are high it means that quality is low. The negative relation between performance indicators and costs indicates that an increase in mortality within 30 days (low quality) is associated with lower costs, indicating that there exists a costs-quality trade-off.

The log likelihood variance of random stochastic noise “v” (lnsig2v) is significant at 1% level in all models, however, the log likelihood variance of the inefficiency term “u” (lnsig2u) is only significant for model 2 (-3.922), which suggests that cost inefficiency exists in the model including treatment variables. As mentioned earlier, the SFA model is able to sperate inefficiency from random stochastic noise. In order for the hospitals to be considered efficient, sigma_u, the standard deviation of inefficiency component, must be equal to zero (u=0). sigma_u is close to zero for model 0 (0.000) and model 1 (0.001), but not close as close to zero for model 2 (0.139), indicating that there exists inefficiency. The chi-bar-square (Chibar2) is 0 and the p-value (prob>=chibar2) is 1 for model 0 and 1. Sigma2, the total error variance $\sigma_s^2 = \sigma_v^2 + \sigma_u^2$ needs to be positive, which it is in all models.

Lambda, the ratio of standard deviation of inefficiency to stochastic noise is 0.002 for model 0 and 0.013 for model 1, which is below 1 and close to 0, indicating that stochastic random noise is more important in the production than inefficiency (v>u) when it comes to the decomposition of the total error. The effects caused by noise (lnsig2v) are significant for all models. When lamda is zero it means that all observations are placed on the frontier and no inefficiency. On the other hand, for model 2 lamda is 2.173, which larger than 1 and this reveals that there is presence of inefficiency at the hospitals when treatment variables are included in the model. Inefficiency is therefore more influential than random stochastic noise. Overall, adding treatment variables in model 2 changes several coefficients and makes lnsig2u significant meaning that inefficiency exists across the hospitals.

Scale elasticity was found by adding the coefficients of the four DRG points for model 2 and dividing 1 by the sum of the coefficients. The scale elasticity was found to be 1.06 ($= \frac{1}{0.945}$). The value is above 1 which means that that production is experiencing increasing returns to scale and it is beneficial to be a large hospital. As a result, large hospitals are able to produce output at a higher productivity level.

5.2 Cost Efficiency

The cost efficiency estimation is based on model 2 and follows a half-normal distribution. Equation 9 on page 20 is used to estimate the efficiency.

First, cost efficiency was estimated based on model 2 from table 6, in order to see how the average cost efficiency across the 22 hospitals over the years. Results are presented in table 7 below. Overall, Norwegian hospitals are operating cost efficient with a mean cost efficiency of approximately 89.5% for model 2. The minimum value is 72.7%, meaning that for one year a hospital with the lowest cost efficiency level operated at 72.7% while the most efficient hospital is as much as 95.5% cost efficient. As mentioned earlier, the assumptions made so far can calculate a higher cost efficiency level in SFA and misspecifications could result in an inconsistent estimate of average cost efficiency. DEA could possibly calculate a lower cost efficiency because the error term is not decomposed into random stochastic noise and inefficiency. Remember that the true frontier is unobservable, but we are given an approximate estimation of the cost efficiency.

Table 7: Average cost efficiency for model 2 for 22 hospitals (n=154)

VARIABLE	Mean	Std. Dev.	Min	Max
Cost efficiency	0.895	0.058	0.727	0.975

The table shows average cost efficiency of 7 years for all 22 hospitals. The cost efficiency at each hospital over the years showed quite a variation. For some hospitals the cost efficiency has increased over the years while for others it decreased. For Oslo University Hospital the cost efficiency went from 82.1% in 2008 to 80.6% in 2014. In 2012 the cost efficiency was as low as 73%, which was the lowest cost efficiency registered across the observational years. A smaller hospital, Finnmark Hospital Trust, went from 81.4% cost efficiency in 2008 to 76.6% in 2014. A hospital that experienced an increase in cost efficiency was St. Olavs University Hospital. In 2008 the hospital was 84.8% cost efficient, while in 2014 the cost efficiency was 92.9%, resulting in an increase of 8.1%. Overall, 10 of the hospitals experienced an increase in cost efficiency in 2014 than in 2008, and 12 hospitals experienced a slightly lower cost efficiency in 2014 compared to 2008. In 2014 the most cost-efficient hospital was Vestfold Hospital Trust, while the least cost efficient was Oslo University Hospital.

Table 8: Average costs efficiency for each hospital 2008-2014

HID	Hospital	Cost Efficiency	Ranking
1	Oslo University Hospital	0.787	22
2	Vestre Viken	0.929	8
5	Akershus University Hospital	0.828	19
11	Innlandet Hospital Trust	0.948	3
12	Østfold Hospital Trust	0.866	17
23	Vestfold Hospital Trust	0.950	1
24	Telemark Hospital Trust	0.925	9
25	Hospital of Southern Norway	0.949	2
26	Stavanger Hospital Trust	0.921	10
27	Fonna Hospital Trust	0.947	4
28	Bergen Hospital Trust	0.891	15
29	Førde Hospital Trust	0.912	12
30	Nord-Trøndelag Hospital Trust	0.938	6
34	St. Olavs University Hospital	0.883	16
35	Møre og Romsdal Hospital Trust	0.944	5
36	Finnmark Hospital Trust	0.794	20
37	University Hospital of North Norway	0.793	21
39	Nordland Hospital Trust	0.848	18
40	Helgeland Hospital Trust	0.892	14
52	Lovisenberg Diaconal Hospital	0.931	7
51	Diakonhjemmet Hospital	0.918	11
54	Haraldsplass Deaconess Hospital	0.896	13
	Average CE	0.895	

Table 8 shows the average cost efficiency for each hospital across 7 years. The column to the left represents the ID number for the hospitals and the column to the right is the ranking list where value 1 is the most cost-efficient hospital and value 22 is the least cost-efficient hospital. The most cost-efficient hospital in average was Vestfold Hospital Trust with a score of 95%. The hospital that was least cost efficient on average was Oslo University Hospital, with a score of 78.7%. The average cost efficiency score for all hospitals from 2008-2014 was

89.5%. This means that in order to operate fully efficient, the hospitals can only reduce input cost by 10.5%

Chapter 6 DISCUSSION

This chapter discusses the main findings of the study, which were presented in chapter 5.

6.1 Cost-quality trade-off

The aim of this study was to examine if there exists a trade-off between quality and costs. This was done based on a panel data of 22 Norwegian hospital from 2008 to 2014 (154 observations in total) using stochastic frontier analysis as chosen method. Cobb Douglas was used, and certain variables were transformed into logarithmic form. Table 6 summarized the estimation for the stochastic frontier of the 3 models. Model 2, which includes DRG, patient characteristics and treatment variables, was chosen in the beginning of the study to be analyzed more narrowly, due to the study by Kittelsen *et al.* (2015a) where statistical evidence favored model 2. First of all, lambda, the ratio of u and v, was larger than 0 and which means that the total error consist of more inefficiency, and inefficiency seems to be more important than stochastic random noise across the hospitals for model 2.

The DRG points were all statistically significant at 1% for model 2, meaning that when DRG points increase so will the total hospital costs. DRG point for day patients had the highest coefficient. A 1% increase in DRG point for day patients increases costs with 0.299%. A higher DRG point increases total costs because more patients are treated at the hospital. In order to see if there exists a cost-quality trade-off, the quality indicator coefficient for mortality within 30 days and 30-day emergency readmission was interpreted. The coefficients were negative, indicating that there exists an association between costs and the two performance indicators. Emergency readmission was statistically significant, but only at 10% level. As a result, because mortality within 30 days is significant at 5%, it is evidently stronger than emergency readmission and will be the main focus of the discussion.

As mentioned earlier, the performance indicators were calculated by dividing observed patient outcomes by predicted patient outcomes. A lower value is always better because that means that observed mortality is lower than the predicted mortality. High performance indicator will therefore be equal to low quality. Therefore, an increase in performance indicator for mortality within 30 days resulted in a reduction in costs and was statistically significant at 5% level. Since results showed that high performance indicator coefficient is associated with low quality and low quality results in lower costs, a trade-off between quality and costs exists.

This means, according to theory, that the hospitals are efficient in their production because a trade-off is present.

Regarding trade-off, Hvenegaard *et al.* (2011) found that higher mortality rate (low quality) was linked to lower costs at Danish hospitals indicating a trade-off between quality and costs, which is the same as the results found in this study. Figure 1 shows the u-shape relationship between quality and costs, and in order to provide higher quality there will occur a trade-off between cost containment and quality improvement. Positive trade-off will be present if higher quality requires higher costs, or lower costs results in lower quality. A negative relationship between costs and quality can in some degree be a result of complementarity and working to reducing costs as well as improving quality means they are two compatible goals. The results show that improving quality, placed on the upward slope of the u-shape of the net cost curve, will always have a higher cost, which means there will be a trade-off between cost containment and quality improvement and a hospital with trade-off will produce services efficiently.

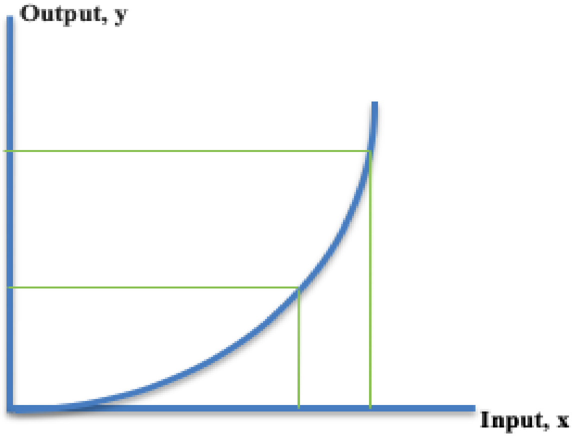
The relationship between cost and quality varies however from study to study, whether it is negative or positive, but this is probably due to how quality is measured. In this study, findings showed a trade-off where low quality was associated with lower costs, or vice versa, high quality results in higher costs. Improving quality, on the upward slope of the u-shape of the net cost curve will always have a higher (opportunity) cost, which means there will be a trade-off, and therefore hospital services are efficient. High quality is in this case costly and containing costs may be a challenge because it may negatively affect the quality level. If a hospital is to reduce their costs it can cause poor patient outcomes. One potential situation would be if hospitals were placed at the downward slope because at least here costs can be reduced by investing in low quality and at the same time quality can be improved. On the other hand, if this was the case, then the hospitals would not be efficient. A trade-off between costs and quality will need to be present in order for production to be efficient.

Correlation was also tested to see how the variables are related to each other. Appendix 1 also shows a slightly negative correlation between performance indicators and costs, where low quality is associated with lower costs. The trade-off finding between quality and costs is similar to the findings of Schreyögg and Stargardt (2010), where they also found negative association between costs and mortality, indicating a trade-off. They concluded that less money spent results in an increase in risk for mortality. The hospitals in this study can therefore use more resource in order to prevent negative health outcomes for hospital patients. Providing treatment is resourceful but will in best case scenario lead to better outcomes. Finding a balance between high-quality services at reasonable costs is essential, and reimbursement methods linked to patient outcomes (quality) could be one way of gaining higher quality. The reimbursement method would in this case not offer any incentives linked to activity, i.e. treating many patients, but instead shift the focus on the quality of treatment (better patient outcomes).

The most interesting part of the trade-off result is that it is not the same finding found in the similar study by Kittelsen *et al.* (2015a) using the same model. They found a significant, but weak, trade-off between productivity and inpatients readmission within 30 days due to a positive association between high productivity (low costs) and high readmission (low quality), which is similar to our findings for readmission as well. In addition, high mortality (low quality) was linked to higher costs, therefore no trade-off between quality and costs because it is possible to increase quality without increase costs. Of course, the difference in findings can be due to the method choice such as DEA or the data. They found that high productivity (low costs) was positively associated with lower rates of mortality, while this thesis found negative association between costs and quality.

At last, in addition to finding a significant trade-off between quality and costs, it was also found that there exists an increasing return to scale for model 2, when including mortality within 30 days and emergency readmission within 30 days with the four DRG points. Producing many services is cheaper per unit than if the hospital produced on a smaller scale. Increased returns to scale will for example mean that a hospital is able to increase production with 20% by only increasing input by 10%. This can be shown in figure 4, where an increase in output, production, increases more than input, costs. This could mean, for example, that merging hospitals is a good thing, as Giancotti *et al.* (2017) stated. Kittelsen *et al.* (2018) found that larger hospital size could reduce costs associated with treatment without affecting quality.

Figure 4: Increasing Returns to Scale



6.2 Cost efficiency

On the whole, average cost efficiency across the 22 hospitals from 2008 to 2014 was found to be 89.5%, which means that in order to operate fully efficient the hospitals can only reduce their input costs by 10.5% without decreasing output. We know from the SFA output that cost inefficiency was present, due to a lambda value larger than zero and a significant log likelihood of variance of inefficiency term (lnsig2u). Both these indicate that inefficiency is more important than stochastic noise and this was only present for model 2. However, the high average cost efficiency shows that more Norwegian hospitals are cost efficient than inefficient. The average cost efficiency estimate was very close to the findings of Kittelsen *et al.* (2015b). Cost efficiency can be seen as equivalent to technical efficiency, hence, Kittelsen *et al.* calculated technical efficiency to be 89.9%, which is very close to the score found in this study. This strengthens the result. Cost efficiency in the hospitals are significantly dependent on the variables. This means that mortality within 30 days and emergency readmission within 30 days influence the cost efficiency across the hospitals over the years.

The most cost-efficient hospital was Vestfold Hospital Trust, with 95%, which is above the average. The hospital with the lowest value was Oslo University Hospital with 78.7%. It is however not so surprising that Oslo is at the bottom of the list. The hospital consists of three university hospitals in Oslo: Rikshospitalet, Ullevål University Hospital and Aker University Hospital. In total, Oslo University Hospital consists of 1500 beds. One theory could be that the low score is a result of the institutions specialization and treating some of Norway´s most resourceful patients. Perhaps the hospital focuses a lot on patient outcomes and delivers services of high quality. Since we found a cost-quality trade-off it makes sense that cost efficiency is low at Oslo University Hospital, because quality is costly and focusing on treatment and patient outcomes can affect the cost efficiency. Oslo University Hospital is internationally known for having one of the best Heart and Lung Clinic and many patients from the whole country are sent to this hospital in order to get the best treatment.

Being a hospital that treats a large proportion of patients each year and at the same time accepting the most advanced and expensive patients may be a cause that can explain the cost efficiency score. Giancotti *et al.* (2017) stated that diseconomies of scale would occur if the number of beds were above 600, hence, when production grows (more patients) there comes a point where costs per unit also will increase. Also, teaching hospitals and specialization were found to be more costly and possibly they deliver high quality which is also cost driving. Perhaps it is a difficult task managing hospitals when they become too large. This finding is somewhat close to what Erichsen (2016) found, where higher costs and less efficient services was found when hospitals became larger.

The other 5 university hospitals also did not score as high compared to the average efficiency level. It can be discussed whether specialization perhaps affects cost efficiency, meaning that specialization results in lower cost efficiency. It would be interesting if a similar study as this one was done by including a variable for specialization and testing whether it has a positive or negative affect on cost efficiency. It may not be an easy task increasing cost-efficiency, however, working to find information about where costs can be reduced to achieve cost-effectiveness and delivering services of good quality is needed in order to provide more health for the money. The reason for Vestfold Hospital Trust being at the top could be due to its smaller size (400 patient beds) and their use of resources. According to the findings of Giancotti *et al*. (2017), Oslo University Hospital has too many beds than optimal. Earlier it was found that it is a good idea to be a large hospital, but because optimal hospital size is difficult to estimate, it could be that some hospitals simply are too large than optimal in terms of beds.

6.3 Limitations

One limitation in this study is based on the assumptions made throughout the research. First, a Cobb-Douglas cost function was assumed. Using a Translog cost frontier could perhaps have been better due to its flexible form and avoiding modeling errors but due to lack of time it seemed appropriate to use the Cobb-Douglas cost function. DEA could also have been performed, but it seemed more suitable to use SFA because it separates inefficiency term and random noise from the residuals. It could be recommended to use DEA if further research of this topic is to be done as well as study other factors that might influence efficiency in the analysis, such as specialization at hospitals. The study also assumed a half-normal distribution but using truncated-normal or exponential distribution could influence the cost efficiency estimate and alter the estimates of the residual components. A larger yearly dataset could also strengthen the results giving more robust findings. Some data errors can occur due to lack of data variability and a wider research on this topic could be recommended in the future, especially taking a more thoroughly look on the u-shape relation between costs and quality.

Using hospital mortality as a quality measure may result in some challenges in the sense that there are differences when it comes to medical conditions, whether it is a local hospital or not, that costs depend on how quick the patient dies after being admitted to the hospital and the fact that most resources are used for patients during their last days before death. As a result, this means that costs are an endogenous variable that is affected by the health outcomes. However, in this study it is less of an issue if we choose to measure mortality independent of death occurring in hospital or if it happens after discharge. Another limitation is whether cost minimizing behavior is assumed, and this does not hold because the analysis identifies units that deviate from cost efficiency. In addition, the cost data only represents the overall costs of each hospital and is not separated into a specific treatment such as emergency, which can cause analysis to be less accurate.

At last, since this is a panel data, it would be normal to use a time-varying decay model or a time-invariant model. The reason why the decision to use half-normal distribution was chosen was because it gave more precise estimates and it resulted in more significant results than the time models. Cost efficiency using the half-normal model was for example closer to the results from the article by Kittelsen *et al.* (2015b) than if time-varying decay model was used.

These assumptions are the reason for why half-normal seemed more acceptable to use. Overall, the results from this study may be due to method choice.

Chapter 7 CONCLUSION

This thesis examined if there exists a cost-quality trade-off across 22 hospitals in Norway for the period 2008-2014 and calculating the average cost efficiency across the seven years. The frontier method SFA was employed and was based on the Cobb-Douglas method in order to estimate efficiency scores for the hospitals and used a half-normal distribution. Mortality within 30 days and readmission within 30 days were the two main quality indictors used in the result section. Case-mix adjusted performance indicators were calculated for 3 models. Model 2, which included DRG, patient characteristics and treatment variables, was the main model used to answer the hypothesis of the thesis because of significant results in previous study.

The results showed a trend of negative coefficients for mortality within 30 days and emergency readmission within 30 days, indicating an association between costs and quality. Mortality was statistically stronger than emergency readmission. An increase in mortality within 30 days (low quality) will therefore result in lower costs. The findings therefore confirm that there exists a trade-off between quality and costs given the assumptions made throughout the thesis. In addition, increasing returns to scale was found, suggesting that it is a good idea to be a large hospital. Inefficiency was present for model 2 and played a bigger part than stochastic random noise. On the whole, average cost efficiency across the 22 hospitals from 2008 to 2014 was 89.5%, which means that the hospitals can reduce input costs by 10.5%. It can also be explained as cost inefficiency being 10.5% across the 7 years. This estimate is very close to the findings of Kittelsen *et al.* (2015b), where they found technical efficiency to be 89.7%. The high average cost efficiency shows that more hospitals are being cost efficient than not being efficient. However, the university hospitals have some of the lowest average cost efficiency scores, so perhaps specialization affects the costs efficiency level across Norwegian hospitals. Further study on this would be interesting to obtain more information about the association between specialization and cost efficiency.

REFERENCES

Anthun, K. S. (2017). *Productivity in the Norwegian hospital sector. Financing, quality, coding and comparability.* Doctoral theses at NTNU, 2017:302. ISBN 978-82-326-2673-1 (electronic ver.).

Anthun, K. S., Kittelsen, S. A. C., Magnussen, J. (2017). Productivity growth, case mix and optimal size of hospitals. A 16-year study of the Norwegian hospital sector. *Health Policy* 121(4). 418-425.

Ash, A. S. (2003). Comparing outcomes across providers. In *Risk Adjustment for Measuring Health Care Outcomes,* LI Iezzoni. Health Administration Press. Chicago; 297-333.

Coelli. T. J. (1996). Centre for efficiency and productivity analysis (CEPA) Working papers. A guide to Frontier version 4.1: A computer program for stochastic frontier production and cost function estimation. No. 7/96. Department of Econometrics. University of New England.

Coelli, T. J., Rao, D. S. P., O'Donnell, C. J., and Battese, G. E. (2005*). An Introduction to Efficiency and ProductivityAanalysis.* USA: Springer Science & Business Media.

Erichsen, A. B. (2016). *Forholdet mellom størrelse og kvalitet av norske sykehus.* Fagforbundet.
https://brage.bibsys.no/xmlui/bitstream/handle/11250/2432707/Erichsen.pdf?sequence=1

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society.* Series A (General), 120(3), 253-290.

Giancotti, M., Guglielmo, A., Mauro, M., (2017). Efficiency and optimal size of hospitals: Results of a systematic search. *PLoS ONE* 12(3): e0174533.
http://doi.org/10.1371/journal.pone.0174533

Helsedirektoratet. (2017). *Innsatsstyrt finansiering.* Regelverk. IS-2689.

Helsedirektoratet. (2012). *Norway and Health. An Introduction.* Norwegian Directorate of Health. IS-1730E.

Hirschey, M. (2009). *Fundamentals of managerial economics, 9th edition*. Ohio USA: Cengage Learning.

Hussey, P. S., Wertheimer, S., Mehrotra, A., (2013). The association between health care quality and cost: a systematic review. *Annals of Internal Medicine.* 158(1): 27-34.

Hvenegaard A., Arendt, J. N. Street, A. Gyrd-Hansen, D. (2011). Exploring the relationship between costs and quality: Does the joint evaluation of costs and quality alter the ranking of Danish hospital departments? *The European Journal of Health Economics,* 12(6), p.541-551.

Jacobs R. (2001). Alternative Methods to Examine Hospital Efficiency: Data Envelopment Analysis and Stochastic Frontier Analysis. *Health Care Management Science*, 4, pp.103–115.

Jacobs, R., Smith, P., and Street, A. (2006). *Measuring Efficiency in Health Care: Analytical Techniques and Health Policy.* Cambridge University Press, Cambridge.

Kittelsen, S. A. C., Anthun, K. S., Goude, F., Huitfeldt, I. M. S., Häkkinen, U., Kruse, M., Medin, E., Rehnberg, C. and Rätto. (2015a). Costs and Quality at the Hospital Level in the Nordic Countries. *Health Economics.* 24(Suppl. 2): 140–163.

Kittelsen, S. A. C., Anthun, K. S. and Huitfeld, I. M. S. (2017). Kvalitet og produktivitet i norske sykehus. Universitetet i Oslo. Avdeling for helseledelse og helseøkonomi. Institutt for helse og samfunn. *Michael.* 2017; 14: Supplement 19, 75-85.

Kittelsen, S. A. C., Anthun, K. S., Häkkinen, U., Kruse, M., Rehnberg, C. (2018). Scale and quality in Nordic hospitals. *Nordic Journal of Health Economics,* Vol.6, No. 2, pp. 29-44.

Kittelsen, S. A. C., Winsnes B. A,. Anthun, K.S., Goude, F., Hope, Ø., Häkkinen, U., Kalseth, B., Kilsmark, J., Medin, E., Rehnberg, C., and Rättö, H. (2015b). Decomposing the productivity differences between hospitals in the nordic countries. *Journal of Productivity Analysis* 43 (3): 281-293.

Kjekshus, L. E., (2000). Om å sammenlikne sykehusprestasjoner. *Tidsskrift for Den norske legeforeningen.*

Kumbhakar, S. and Lovell, C. (2000). *Stochastic frontier analysis.* Cambridge [England]: Cambridge University Press.

Linna, M., Häkkinen, U., and Magnussen, J. 2006. Comparing hospital cost efficiency between Norway and Finland. *Health Policy,* 77, 268-278.

Medin, E., Anthun, K. S., Häkkinen, U., Kittelsen, S. A. C., Linna, M., Magnussen, J., Olsen, K., Rehnberg, C., (2011). Cost efficiency of university hospitals in the Nordic countries: a cross-country analysis. *European Journal of Health Economics*. 12(6): 509-519.

Parmeter, C. F. and Kumbhakar, S. C. (2014). Efficiency Analysis: A Primer on Recent Advances. *Foundations and Trends® in Econometrics*, 7(3-4), pp.191-385.

Rosko, M. D., Mutter, R. L., (2008). Stochastic Frontier Analysis of Hospital Inefficiency: A review of Empirical Issues and an Assessment of Robustness. *Medical Care Research and Review*. DOI: 10.1177.

Ross, J. S., Normand, S-L.T., Wang, Y., Ko, D. T, Chen, J., Drye, E. E., Keenan, P. S., Lichtman, J. H., Bueno, H., Schreiner, G. C., and Krumholz, H. M,. (2010). Hospital volume and 30-day mortality for three common medical conditions. *New England Journal of Medicine* 362 (12): 1110-1118.

Schreyögg, J., & Stargardt, T. (2010). The trade-off between costs and outcomes: the case of acute myocardial infarction. *Health services research*, *45*(6 Pt 1), 1585–1601. doi:10.1111/j.1475-6773.2010.01161.x.

Statistics Norway. (2018a). *NOK 65 000 spent on health per capita.* Published 14 March 2018. https://www.ssb.no/en/nasjonalregnskap-og-konjunkturer/artikler-og-publikasjoner/nok-65-000-spent-on-health-per-capita

Statistics Norway. (2018b). *Patient statistics.* Statistics. Health. Updated 26 March 2019. https://www.ssb.no/en/pasient

Statistics Norway. (2018c). *Store forskjeller i pengebruk mellom norske sykehus.* Published 12 December 2018. https://www.ssb.no/helse/artikler-og-publikasjoner/store-forskjeller-i-pengebruk-mellom-norske-sykehus

StataCorp. (2017). *Stata Base Reference Manual Release 15.* Statistical Software. College Station, TX: StataCorp LLC. P.822-836.

APPENDIX

Appendix 1: Correlation for model 2 (n=154)

	lnCosts	lnv_emrg_p	lnv_elective_p	Lnv_day_p	lnv_out_p	perf_u_mort30last	perf_u_readm30_emgc
lnCosts	1.000						
lnv_emrg_p	0.9544	1.000					
lnv_elective_p	0.9598	0.8950	1.000				
lnv_day_p	0.9648	0.9448	0.9036	1.000			
lnv_out_p	0.9733	0.9457	0.9272	0.9710	1.000		
perf_u_mort30last	-0.0561	0.0922	-0.1010	-0.0264	-0.0501	1.000	
perf_u_readm30_emgc	-0.3614	-0.2261	-0.4625	-0.2350	-0.3619	0.1389	1.000

Values +1 indicate perfect positive correlation, -1 indicates perfect negative correlation, 0 means no correlation.