UiO : Det medisinske fakultet

Productivity and Quality Analysis of Norwegian Hospital Mergers

A Simar & Wilson Two-Stage Approach with Data Envelopment Analysis and Difference-in-Differences Estimation

Author
Natalie Emanuelsson

Student number
579626

Supervisor
Sverre A. C. Kittelsen

A thesis submitted as a part of the European Master's in Health Economics and Management, with specialization in Health Economics and Policy.

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

July 31st, 2020

2020

Project 4115 “The effect of DRG-based financing on hospital” at the Frisch Centre and NTNU, grant no. 214338 from the Research Council of Norway.

Productivity and Quality Analysis of Norwegian Hospital Mergers

Natalie Emanuelsson

DECLARATION OF OATH

With this declaration, the student confirms having written the thesis him or herself without any outside help. Others' thoughts and ideas are clearly marked as such and the master thesis has not been handed in during the course of another program and has not yet been published. Each master's thesis needs to contain such a declaration and has to be signed by the student in person. An electronic signature cannot be accepted. Exact formulation of this declaration:

"I hereby declare, under oath, that this master thesis has been my independent work and has not been aided with any prohibited means. I declare, to the best of my knowledge and belief, that all passages taken from published and unpublished sources or documents have been reproduced whether as original, slightly changed or in thought, have been mentioned as such at the corresponding places of the thesis, by citation, where the extent of the original quotes is indicated.

The paper has not been submitted for evaluation to another examination authority, nor has it been published in this form or another."

SIGNATURE	
29/07-2020 Natalie Emanuelsson	29.07.2020
Date and signature of student	Date and signature of supervisor

ACKNOWLEDGMENTS

A two-year master program at European Master's in Health Economics and Management is completed, and the process of writing this master thesis have been both very educational and fun. This master thesis is a reminder of all the hard work that I have put down during the past five years as a student at University of Oslo and marks the end of an era where new challenges will be given me in the years to come. I am appreciative for the opportunity to write this thesis for project 4115 "The effect of DRG-based financing on hospital" at the Frisch Centre and NTNU, grant no. 214338 from the Research Council of Norway.

I would like to thank my supervisor Sverre A. C. Kittelsen for exceptional feedback and help with dataset when things were problematic. I am grateful that you have dedicated your own time to help me out when the situation at the Frisch Centre was affected by the Covid-19 virus this spring. I could not have done this without your help.

I would also like to thank my dear friend Veerle Bun for great conversations and critical feedback at the Frisch Centre were we have shared multiple cups of coffee and loud laughter. I wish you good luck in the Netherlands. I would also like to thank my dearest friends and beloved family for proofreading, motivational speeches, and great support during the past seven months. I hope I have made you proud.

Finally, I would like to thank all of my professors and fellow students at the EU-Hem program. It has been two amazing years that I will not take for granted.

It has been a pleasure!

Natalie Emanuelsson

July 2020

ABSTRACT

Background and research question: Hospital mergers have been a hot topic in the Norwegian healthcare sector the past two decades. This study is using a sample of 22 Norwegian hospitals that have merged from 1999 until 2014. A major ownership reform have influenced this period and contributed to both hospital mergers and reorganization. I have examined whether hospital mergers have resulted in increased productivity and quality in Norwegian hospitals.

Data: To perform this study, the Norwegian Patient Register has provided patient-level data. These data are necessary to develop case-mix adjustments for all hospitals included in the analysis and to calculate each patient's travel time from their municipality to the specific hospital. Moreover, hospital data from 1999 until 2014 have been aggregated and merged with the patient-level data from the same time. A sample of approximately 83,000,000 patient records from 1999 until 2014 were used in the analysis.

Methodology: Case-mix adjustment variables such as DRG, patient characteristics, and treatment variables were used to adjust for differences in patient composition between hospitals and to estimate the performance indicators in three models. Productivity scores for both merged and non-merged hospitals were estimated through a Simar & Wilson two-stage approach using Data Envelopment Analysis (DEA-SW), where the Difference-in-Differences estimation (DiD) for productivity were calculated simultaneously. In addition, DiD estimation on quality-scores were calculated to detect the potential differences of quality between merged and non-merged hospitals.

Most important findings: The DEA-SW results showed that merged hospital does not perform with better productivity than non-merged hospitals. However, there is a positive trend over time, which indicate that the productivity off all hospitals in the analysis are increasing. Furthermore, the DiD estimation did not show any evidence of increased quality for merged hospitals compared to non-merged hospitals. The merged hospitals had higher and statistically significant DID-coefficient at a 5% level in all five tables, which means that they have lower quality compared to the non-merged hospitals.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

ATTACHMENTS IN THE APPENDIX

ACRONYMS

ABF……………………………………………………………………..Activity-Based Financing
AE……………………………………………………………………...Allocative Efficiency
ALOS…………………………………………….………...……..Average Length of Stay
CE……………………………………………..……….………………..Cost Efficiency
CETERIS PARIBUS…………………..…………………………Everything Else Being Equal
CRS……………………………………………………………………..Constant Returns to Scale
DEA…………………………………………………………..Data Envelopment Analysis
DGP……………………………………………………………...Data Generating Process
DiD……………………………………………………………..…Difference-in-Differences
DMU………………………………………………………...…. Decision-Making Unit
DRG………………………………………………………...Diagnostic-Related Groups
FTE……………………………………………………………………Full-Time Equivalent
GDP……………………………………………………………………Gross Domestic Product
HF………………………………………………………………………Health Enterprises
IQR……………………………………………………………………Interquartile Range
LOS……………………………………………………………………...Length of Stay
OECD………………………The Organization for Economic Co-operation and Development
OLS……………………………………………………………………Ordinary Least Squares
OUS……………………………………………………………...Oslo University Hospital
RHF……………………………………………………………..Regional Health Enterprises
SE……………………………………………………………………………Scale Efficiency
SFA…………………………………………………………..Stochastic Frontier Analysis
TE………………………………………………………….………...…Technical Efficiency
TP……………………………………………………….…...…Technical Productivity
UNN…………………………………….……...University Hospital of Northern Norway
VRS…………………………………………..……….…..Variable Returns to Scale
QMS………………………………………………….….…....Quality Management System

1 INTRODUCTION

This chapter will include a brief introduction of how the healthcare sector in Norway is organized and the background that has contributed to the Norwegian hospitals and health care sector they have today. The relevance of the study will be discussed. Moreover, the research question will be presented together with the aim of the study. Lastly, a general overview of the upcoming chapters will be provided.

1.1 Norwegian hospitals and the healthcare sector in Norway

In 2014, the Norwegian hospitals offered health services to approximately 5.1 million inhabitants. Out of the 5.1 million Norwegians, about 1.8 million patients were registered in at least one hospital (Statistics Norway, 2020). The provision of health services in Norway have mainly been in the hands of the public sector and funded by the state through general taxation. The specialized health care sector is funded through 60 percent block funding and 40 percent activity-based funding (ABF). The block funding does not consider the activity but base the funding on the number of inhabitants in the region and the age composition. The basic funding is provided to the regional health enterprises (RHF) and then distributed to the local health enterprises (HF) (Helse- og omsorgsdepartementet, 2014). This flexibility of the RHF yields an opportunity to customize the allowances to the real cost of the institutions they are responsible for. On the other hand, ABF depends on the health services provided and the number of treated patients. ABF is based on Diagnostic-Related Groups (DRG), which is a classification system that, in Norway, consists of 758 DRG-groups and works as a reimbursement system. The DRG-groups represents both medical and economic information. Patients that are grouped in the same DRG are approximately equal in the medical sense but also regarding the resources they need for treatment (Sosial- og helsedirektoratet 2007, 14). Additional elaboration on the DRG-system will be given in section 4.3. The goal of ABF is not to have the highest possible activity at the HFs. However, the aim is to stimulate the HFs to fulfill the goals given by the government most efficiently, but also to partially reimburse the RHF for costs regarding volume of treatment.

The composition of the health care financing in Norway is meant to control the activity at the HFs but also to give enough flexibility to the HFs to utilize available capacity within given frames (Sosial- og helsedirektoratet 2007, 12). However, resources are scarce, and an important goal of the government is to ensure efficient use of available resources in all health care systems (Anthun *et al*., 2017, 419; Gutacker *et al*., 2013). Moreover, the share of elderly individuals in the population of Norway is increasing; hence, the health care expenditures per capita are continuously rising. This is the main reason why efficiency and focus on prices have been a more pervasive topic in Norwegian health policy the previous decades. The health care system provides a wide range of services to both people living in the urban areas of the country but also to those living in the more decentralized parts of Norway. However, a small part of the hospital sector in Norway is privately run, and these hospitals have the vast majority of specialists, which includes both private hospitals and private contracted specialists. Statistics presented by Statistics Norway from 2019 state that one out of four consultations are taken by private specialists (Vold. B & Hjemås. G., 2019).

A few major health care reforms have changed how the Norwegian health care system works and how it is organized. Two main health reforms were implemented as a part of the Norwegian health policy during the past two decades and have contributed to the modernization of the health care system they have today. The first health care reform is the Norwegian Hospital Reform of 2002, which mainly involved that the state took over the ownership of the hospitals (Magnussen *et al.*, 2005). The second reform is The Coordination Reform of 2009, which had an overall goal to contribute with more efficient interactions both between services and the level of services instead of having obstacles for cooperation (Helse- og omsorgsdepartementet, 2008, 8). Furthermore, Norway has a health care system that is based on the Scandinavian Welfare model which involves that every inhabitant has an equal right to choose their health care provider and to access health care of a certain quality and availability (NOU 2003:1).

1.2 Efficiency and quality in the Norwegian health care sector

Growing demand for health services in developed countries and increasing expenditures on health care, which amount to a significant proportion of the gross domestic product (GDP), have emphasized the efficiency concerning provision of health services and containing costs. However, the constraint to control costs through increasing efficiency may influence the quality of the health services provided. Hence, quality control could be necessary for the health care sector in Norway (Kittelsen *et al*., 2015, 140). Furthermore, if the health service production provided by a hospital is efficient, one may assume that there exists a trade-off between the quality and quantity of the health services. In other words, one cannot treat the same quantity of patients if the quality of the treatment is increasing. This is due to the opportunity costs that occur. Hence, the alternative is to use more resources (Kittelsen *et al*., 2015, 141). The Norwegian health authorities have a goal to increase hospital efficiency and to improve the quality of the health services they are providing. High cost of health services are not always a sign of inefficiency, but can be associated with improved quality and better health outcomes. Moreover, if health care providers can reduce the costs of the health services, they may release resources that can be used elsewhere (Gutacker *et al.*, 2013, 931). The goal of increasing hospital efficiency was intended to be reached through more efficient structure of the given hospitals (Magnuessen *et al*., 2006). However, according to Riksrevisjonens examination from 2014, there exist differences in efficiency between hospitals in Norway, and these have to reduce in order to increase hospital efficiency even more. Moreover, changing the organization of patient care requires time and patience and seeing the total effect of hospital mergers on productivity and quality of Norwegian hospitals require datasets that cover a wide time span (Riksrevisjonen, 2017, 33).

1.3 The relevance of the study

Hospital merging has been a hot topic in the health care sector in Norway for the previous two decades. The central question regarding this study is whether hospital mergers contribute to better quality and productivity. Though there have been many hospital mergers in Norway, there is not much existing and convincing evidence that larger hospitals yield better quality and productivity, or whether the expected increase in productivity is a trade-off with quality. The reasoning behind this could be that quality is hard to measure and that there is a lack of accurate evaluation measures for quality in the health care sector. If, for instance, mortality is used as a quality measure at the hospitals - hospitals that treat the sickest and costliest patients are unfairly being labeled as being less efficient than other hospitals who treat less severe patients (Nayar *et al*., 2008, 194). Increasing efficiency and quality are aims for most (if not all) policy makers and hospital leaders. Patients are demanding health care of impeccable quality and to receive their health services without long waiting times. The health care sector in Norway has to deal with an increasing share of elderly patients – and the effectiveness of the health services provided have to increase. Hospital merges have occurred regularly over the past decades, which substantiates the assumption that hospital mergers yields positive effects on the health care sector. At the present time, there are debates in the media about closure of hospitals, especially in the Oslo area. The Norwegian inhabitants are concerned about the increasing waiting times and to be discharged to early after receiving treatment, which may result in complications. This study will try to determine whether hospital mergers have increased productivity and quality of Norwegian hospitals.

1.4 Presentation of research question

Norwegian hospitals have struggled to contain cost while providing health services of high quality. Increasing the quality of health services may be related with higher cost. Hence, decreasing the costs could force the hospitals to provide health services with lower quality. The hospitals aim to provide health services with the highest possible level of quality and to contain costs, which may involve increasing the efficiency at the hospitals. Moreover, multiple hospital mergers and reorganizations through new health care reforms have affected the Norwegian health care sector. These reforms may have influenced the efficiency and quality at the specific hospitals. This thesis aims to examine whether hospital mergers, using data from 1999 until 2014, have affected the productivity and quality of Norwegian hospitals. Using datasets that covers the timespan before and after the hospital mergers are essential to assess whether there have been changes in productivity and quality due to the mergers. The true effect of the hospital merges may not occur shortly after the mergers are completed. Hence, it is necessary to examine the effect of hospital mergers after a considerable period. An efficiency frontier estimated though the Simar & Wilson two stage approach using Data Envelopment Analysis and a Difference-in-Differences estimation of productivity and quality makes it possible to determine the performance indicators of the hospitals before and after mergers. The research question examined in this study is as follows;

"Have hospital mergers contributed to increased productivity and quality in Norwegian hospitals?"

1.5 Overview of chapters

This thesis consists of seven chapters where chapter 1 is an introduction to the healthcare sector in Norway. The presentation of the research question will be presented together with the relevance of the study. Chapter 2 concerns the background of hospital mergers, the Norwegian Hospital Reform of 2002 and previous hospital mergers in Norway. Chapter 3 contains information about the theoretical framework. What is known and what is lacking on the subject will be discussed together with previous studies on the field. Finally, further elaboration on the research question will be given. The data will be presented in chapter 4, while the methodology will be provided in chapter 5. Moreover, the results of the analysis will be presented in chapter 6. Finally, chapter 7 includes the discussion section and concluding remarks.

2 BACKGROUND

This chapter will include the Norwegian Hospital Reform of 2002, which have caused reorganizations of the health care sector. In addition, section 2.2 include an elaboration of Norwegian hospital mergers together with an overview of all hospital fusions in Norway from 2000 until 2009. Finally, an international perspective on hospital mergers is provided in section 2.3.

2.1 The Norwegian Hospital Reform of 2002 and hospital levels

The Norwegian health care system has changed drastically during the previous two decades. The Norwegian Hospital Reform of 2002 was mainly a matter of organizational changes rather than changing the Norwegian health policy and its aims. The reform expected to limit hospital costs and to reduce geographical inequalities when offering health services (Ingebrigtsen, 2010). Magnussen *et al*. (2005) state a number of factors before the reform was implemented, which characterized the Norwegian health care system. These include long waiting times, inequity in the supply of hospital services, and lack of financial responsibilities. The reform included two main elements. First, the central government took over the responsibility of the specialist care and all public hospitals from the county governments January 1, 2002. Second, the reform divided the country into five regional health enterprises (RHF). However, in 2007 Helse Sør RHF and Helse Øst RHF merged into one RHF and given the name Helse Sør-Øst RHF (Magnussen *et al*., 2005). Attachment A in the appendix illustrates a geographical overview of the four health regions after the Norwegian Hospital Reform of 2002. Moreover, there are three levels of hospitals, which are local-, central-, and regional hospitals. This structure was not introduced with the reform but existed before the reform was implemented. The local hospitals are supposed to cover the needs for general health care services, surgery, and midwifery. The central hospitals cover more specialist care, and every county should have one central hospital. Lastly, the regional hospitals were developed in each of the health regions. These hospitals cover highly specialized services but also the health care services provided by the local-, and the central hospitals (Legeforeningen, 2014; NOU 1996:5). During the 1990s the patients were given the freedom to choose their health care provider for treatment and also a waiting time guarantee which gave more power for the patients (NOU 2003:1, 46). The main goal of the Norwegian Hospital Reform was to increase the efficiency of the hospitals, but has the aim of increasing efficiency affected the quality of the health services?

2.2 Merging of Norwegian hospitals

The Norwegian Hospital Reform has contributed to hospital mergers around the country. Mergers are driven either by a belief that higher quality of the health services is a result of increased activity at the hospital, or by a desire to decrease duplication of hospital functions and to increase scale efficiency due to possible scale effects (Kjekshus *et al*., 2007, 230). Even though there have been increased centralization and hospital mergers in many countries around the world, the previous analyses on the field show little or no evidence of scale economies. Moreover, the optimal hospital size is proven to be quite small (Kittelsen *et al.*, 2018, 29). Following the Norwegian Hospital Reform of 2002, minor hospitals that were located near larger hospitals were shut down after the mergers and a need for reorganization of health care services occurred (Anthun *et al*., 2017, 419).

In the article *"Scale and quality in Nordic hospitals"* by Kittelsen *et al.* (2018) it is claimed that if there exist economies of scale in hospitals, the average costs would be lower for larger hospitals than for smaller hospitals. Furthermore, having fewer but larger hospitals would be cost saving for the health care sector. Besides, larger hospitals could be socially optimal if the medical outcomes and the traveling times are not negatively affected by a centralized hospital structure (Kittelsen *et al*., 2018, 30). It may be assumed that larger hospitals also treat patients with more severe conditions. Moreover, one may anticipate that hospitals beyond a certain scale also have higher mortality rates and emergency readmissions, everything else being unchanged. Hence, these hospitals may experience higher costs (Kittelsen *et al*., 2018, 41).

Norway has a long coastline and a significant travel distance from one part of the country to the other. In addition, Norway has numerous of fjords and mountains, especially in the north and on the west coast, which makes it even harder to travel within the country. In fact, Norway has the second-longest coastline in the world (Nag, 2018). Hence, the structure of Norway affects how much time patients have to estimate when they are going to the hospital. Furthermore, Norway is exposed to extreme weather conditions from one time to another that could make it even more challenging to travel longer distances by car.

After the Norwegian Hospital Reform, there have been multiple hospital mergers in Norway, which have contributed to the centralization of the Norwegian hospitals. Shutting down smaller hospitals in parts of the country where the population is smaller makes it even harder for the patients in these areas to travel to their respective hospitals. Whether the hospital mergers in

Norway have contributed to structural and physical changes or only organizational changes have to be taken into account regarding traveling times. Attachment B in the appendix gives an overview over which hospitals that have closed due to the reorganization of the health sector and whether there has been an organizational or physical fusion. However, hospital mergers do restrict competition because patients experience a reduction of hospitals to choose from. Hence, the hospitals may lack incentives and motivation to provide health services above a certain quality level to attract patients. Even though patients in Norway are privileged to choose their preferred provider, there are areas in Norway where there are fewer options to choose from and where the closest hospital is more than seven hours away by car (Vågnes, 2019). However, merging hospitals can make the hospitals larger and yield more room for specialist care. Moreover, larger hospitals also have a focus on education, research, and the environment at the hospital. Hence, assuming that these hospitals have greater cost seems appropriate. It is also common to assume that hospitals that have more focus on education and development also provides services with better quality. Whether this is a correlation or causation is hard to elaborate on and is difficult to prove since there are many potential underlying factors that may influence the result.

It may be that it is cost saving to merge departments or units which are not using their capacity to the fullest or to centralize the emergency function so the hospital can cover a larger share of the population. However, the literature is certain that efficiency gains do not come by itself after merging hospitals (Magnussen, 2012, 783). Moreover, there may be a conflict of interest regarding doctors and the management at the hospitals concerning changes in the hospital structure and the way things are done. A common statement is that doctors often are skeptical about reforms that seem to threaten their autonomy. On the other hand, the management of the hospital would like to take the interest of the public and the inhabitants of the country into consideration when making a decision (Magnussen, 2012, 783). Merging hospitals may force the organizational culture at the hospitals to blend. If the cultures are not compatible, it may be difficult to obtain the desired economies of scale. Furthermore, hospital mergers may be driven by political incentives without taking the size of the hospitals, their culture, and prerequisites into account (Legeforeningen, 2014). Kjekshus *et al.* (2014) state that 27% of all Norwegian hospitals were involved in at least one merge from 1992 until 2000. Moreover, 90% of all public hospitals were affected by hospital merging from 2000 until 2010 (Kjekshus *et al.,* 2014). Table 2.1 below illustrates how Norwegian hospitals have merged from 2000 until 2009 with their respective number of employees.

2.1 Overview of hospital mergers with a five-year timeline

Control variable: Non-merged hospital	Control variable: Merged but not affected	Merger year 0	Merger year 1	Merger year 2	Merger year 3	Merger year 4	Merger year 5

	Number of employees		Year of hospital merger

Hospital	In 2000	In 2008	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009
Rikshospitalet	3601	4860										
Radiumhospitalet	1478	1563										OUS
Ullevål	5208	6290										HF
Aker	1889	1226										
Ski	241	122										
AHUS (Lørenskog)	2616	3231										AHUS HF
Stensby	254	162										
Hedmark SSH	1462	1511										
Gjøvik	1071	992										
Lillehammer	1244	1173									Innlandet	
Tynset	220	219									HF	
Kongsvinger	541	514										
Askim	209	47										
Moss	554	513								Sykehuset	Østfold	HF
Halden	240	59										
Østfold SSH	2137	2873										
Asker og Bærum	1235	1361										Vestre
Buskerud (Drammen)	1646	1866										Viken
Ringerike	511	591										HF
Kongsberg	439	430								Helse		
Notodden	267	213								Blefjell		
Rjukan	145	132										
Telemark (Skien)	1295	1358										Tele-mark
Kragerø	77	121										HF
Arendal	1453	1177										
Lister	234	304									Sørlandet	Sy.
Kristiansand	1688	1657									HF	
Larvik	263	246										
Sandefjord	302	88								Sykehuset	Vestfold	HF
Tønsberg	2081	2111										

Stavanger	3185	3796	
Odda	174	180	H. Fonna HF
Haugesund	1212	1413	
Stord	439	460	
Haukland	6427	8044	Helse Bergen HF
Voss	283	355	
Sogn og Fjordane	914	1346	H. Førde HF
Lærdal	194	321	
Nordfjord	208	236	
St. Olavs	4634	4634	
Diakonhjemmet	1565	1650	
Lovisenberg	1660	1700	
Haraldsplass	925	1000	
Volda	361	470	Helse S.Møre HF
Ålesund	1644	1659	
Kristiansund	557	532	Helse N.Møre HF
Molde	1092	1061	
Levanger	1215	1032	Helse Nord-Trøndelag HF
Namsos	698	558	
Hammerfest	502	528	Finnmark Sy. HF
Kirkenes	362	359	
Tromsø (UNN)	3420	3557	UNN HF
Narvik	391	356	
Harstad	706	562	
Vesterålen	296	335	Nordla-nd Sy. HF
Lofoten	265	251	
Bodø (Nordland)	1551	1722	
Mosjøen	280	264	Helgeland Sy. HF
Mo i Rana	354	468	
Sandnessjøen	288	356[1]	

[1] Numbers taken from Kjekshus *et al.* 2014.

2.3 International perspective

The Norwegian health care system share the same ideology as the other Nordic countries where every inhabitant should have equal access to health services, low deductibles, high degree of taxed based financing, and publicly run hospitals. However, there are some differences between the Nordic counties, but the differences are small when seen in a global perspective (Kristiansen & Pedersen, 2000). Since the 1980s, Sweden have about halved the numbers of somatic acute hospitals due to hospital mergers (Ahgren, 2008, 93). Between 1980 and 2007, the number of somatic hospitals in Denmark decreased from 117 to 35 (Kristensen *et al.*, 2012). The number of somatic hospitals in the Nordic countries are decreasing, and the hospitals are becoming bigger and more specialized. Many European countries are considering developing centralization reforms in order to improve the efficiency of their hospitals (Christiansen & Vrangbæk, 2018, 322). According to the Euro Health Consumer Index from 2018, all Scandinavian countries are performing to top positions when ranking 35 of the European counties. The rankings are based on accessibility, where Norway is ranked among the worst in the group, health outcomes, range and reach of services, prevention, patient rights, and pharmaceuticals. Changing the health care systems in the countries that are ranked at the bottom may influence the patient outcomes to the better.

This study concerns the Norwegian health care system, and comparing the results in the analyses performed in this study to countries that have significantly different health care systems, may not be optimal. However, it is important to have in mind that when comparing results across countries that are different from each other, the results may not be as anticipated. The American health care system, for instance, are significantly different from the Norwegian health care system where the primary focus is on competition, profit and increasing market share. Applying the same strategy in these two countries may lead to divergent outcomes.

3 THEORETICAL FRAMEWORK

This chapter will include information on what is known and what is lacking on the subject together with previous studies that have examined efficiency and quality in Norwegian hospitals by using efficiency analysis and scale efficiency measurements. Moreover, studies that have investigated the change in quality when efficiency has been improved will be presented. The terms efficiency and productivity have been applied interchangeably in the literature. However, these terms are distinct from each other (Jacobs *et al.,* 2006). Section 3.2 will not consider this distinction. Moreover, section 3.3 will include an elaboration of the research question in light of the previous studies presented in section 3.2. Finally, in section 3.3.1, main concepts from the research question will be defined.

3.1 What is known and what is lacking on the subject

Increased efficiency and quality of hospitals have been the overall goal in the Norwegian health policy the previous decades and there have been conducted previous studies that have aimed to determine whether these goals have been fulfilled after the Norwegian Hospital Reform. Moreover, cost containment has been an important topic due to an increase of elderly inhabitant in the country. In order to fulfill the goal of cost containment, hospital efficiency has to be increased. This can be done by, for instance, reducing readmission rates or by decreasing their average length of stay at their health institutions. However, decreasing the length of stay for severe patients may yield long-term consequences for both the patient and the hospital due to poor treatment and worsening of health conditions of the patients, which may result in an increase of emergency readmissions for the hospitals.

Merging firms or organizations has been common the previous decades. Hence, it may be assumed that these mergers have positive effects for the ones that are involved. However, a major share of the studies on the field are showing different results. In the article "Lite lønnsom sykehusfusjon" (English: Unprofitable hospital fusion) by B. M. Andersen, it is stated that the mergers of the hospitals in Oslo have affected the number of beds and employees. Moreover, the capacity at the hospitals have been reduced compared to before the merger. The latter is severe especially due to the increase of the population in Oslo (Andersen B. M, 2012).

Based on previous literature and empirical analyses on the field, the optimal sizes of the infrastructure concerning the hospitals is stated to be quite small (Kittelsen *et al.*, 2018, 30). On the other hand, the medical literature supports the hypothesis that better results can be achieved in larger hospitals since there are more specialist care and a greater volume of procedures and easier to provide professional depth (Legeforeningen, 2014). The effect of Norwegian hospital mergers has been examined to determine whether the cost efficiency of hospitals have improved. Kjekshus *et al*. (2007) found a negative and significant effect. However, the data used in the study are based on hospital mergers before the Norwegian Hospital Reform of 2002 were implemented. Furthermore, when comparing hospitals within or between countries, is has been shown important to perform case-mix adjustments of the given hospitals in the study in order to compare the hospitals without influence of differences regarding the composition of patients (Kittelsen *et al.*, 2015).

Most of the studies that I have found concerning hospital mergers look at either the cost-quality aspect or the cost-efficiency aspect. Hence, there is room for studies that examine the productivity and quality aspect of hospital mergers in Norway. The government of Norway have a goal to increase the overall efficiency at the hospitals in the country, and it will be interesting whether the hospital merges have contributed to increased productivity and quality at the hospitals or if these factors have been worsened by the mergers or stayed unchanged.

3.2 Previous studies

Several studies in the past have evaluated the performance of Norwegian hospitals. These studies have used different outcomes and frontier estimation methods. Even though the Data Envelopment Analysis (DEA) estimation technique is more frequently used when estimating the efficiency of hospitals, we do have studies that have used Stochastic Frontier Analysis (SFA) in the Norwegian health care sector as well.

In a study called *"Scale and quality in Nordic hospitals"* by Kittelsen *et al*. from 2018, they used the SFA method to examine whether the elasticity of scale increases in Nordic hospitals when quality variables are included in the analysis. They used a sample of 58 million patient records from 2008 and 2009 to estimate the SFA in 149 hospitals in Norway, Sweden, Finland, and Denmark. Even though this study does not look at the effect of hospital mergers, they do look at the optimal size of Norwegian hospitals. They used the patent data to estimate the quality indicators on mortality within 30 days and emergency readmissions. These indicators adjusted for age, gender, comorbidities, hospital transfers, and DRG. They included dummies for University hospitals and city hospitals. Moreover, they used traveling time in the analysis to include environmental variables. They found that the estimated scale elasticities did not change when they included the quality indicators. However, in one of the models they estimated, the scale elasticity were significantly larger than one (Kittelsen *et al*., 2018). This finding contradicts previous studies.

Another study conducted by L. Kjekshus and T. Hagen (2007) called *"Do hospital mergers increase hospital efficiency?*", examined the effects on cost- and technical efficiency of seven Norwegian hospital mergers from 1992-2000. The mergers included 17 hospitals. First, they used the DEA method to generate efficiency scores on both merged and non-merged hospitals. Second, they used panel data analysis to estimate the effects of the mergers. The results showed that there was a significant and negative effect of the hospital mergers on cost efficiency, but the analysis did not show any significant effect on technical efficiency. However, one hospital showed positive effects on both technical- and cost efficiency where more hospitals were included and where the acute services were centralized (Kjekshus *et al*., 2007). This type of analysis could be renewed by using more up to date data and by examining the quality of the hospitals as well.

The study called *"Data Envelopment Analysis Comparison of Hospital Efficiency and Quality"* by P. Nayar and Y. Ozcan (2007) used a sample of 53 acute care hospitals from Virginia, USA. They used DEA to calculate the efficiency scores for the hospitals and included measures of quality in the analysis. Furthermore, they compared the results from the analysis to the standard technical efficiency DEA model. The results showed that quality outcomes were not being compromised by the efficiency of the hospitals. However, there were self-reported measures of quality used in the study. Hence, using validated quality measures and larger samples could yield more consistency and generalizability of the results in this study.

The article *"Helseøkonomiske effekter av sykehussammenslåinger*" (English: Health economic effects of hospital mergers) by Tor Ingebrigtsen (2010) examined whether hospital mergers have contributed to lower costs and improvements of the treatments at the hospitals. The article audits seven original articles that together include 476 hospital mergers in both the USA, Great Britain, and Norway from 1982 until 2000. The results showed that there is a potential cost reduction of 10% if the merging process leads to fusions with the right dimensions concerning health services and a development of common culture (Ingebrigtsen, 2010).

Kittelsen *et al*. (2015) conducted a study called *"Costs and quality at the hospital level in the Nordic countries"*. They used data from 160 acute hospitals in 2008-2009 to estimate productivity in a bootstrapped DEA analysis. Furthermore, they use case-mix adjusted measures of quality such as mortality, patient safety indices, and readmissions. The findings showed that it is important to use case-mix adjustments. Moreover, productivity seems to differ at both hospital levels and national levels. They also found that there is a weak but statistically significant trade-off between inpatient readmissions within 30 days and productivity. However, there is a link between hospitals with high 30-day mortality and higher costs. Hence, they conclude with no clear cost-quality trade-off pattern in this study (Kittelsen *et al.*, 2015).

A Norwegian article from 2017 by Kittelsen *et al.* called *"Kvalitet og produktivitet i norske sykehus»* (English: Quality and productivity at Norwegian hospitals) look at how incentives related to cost containment can affect the quality of the health services. They found that Norwegian hospitals have high quality levels due to lower mortality rates compared to the other Nordic countries. However, the readmission rate is higher in Norway and the productivity is lower than in Finland and Denmark. In addition, they found a positive and significant covariance between quality indicator for mortality and productivity in Nordic hospitals, and no significant correlation in the Norwegian data. Only travel time has clear relation with the productivity estimates (Kittelsen *et al*., 2017).

Lindlbauer *et al.* (2015) conducted a study by examining German hospitals from 2000 until 2010 by investigating their quality management system (QMS) on performance. First, they calculated efficiency scores for each hospital by using a bootstrapped DEA. Second, they used generic matching to ensure that the findings were due to the certification and not caused by differences in sample characteristics. Finally, they estimated a Difference-in-Differences (DiD) specification to examine whether the certification had an impact on efficiency. The results indicated that hospital efficiency was negatively related to the QMS (Lindlbauer *et al*., 2015). This study conducts analysis using both the DEA-method and the DiD-method, which are one out of a few studies that I have found to use these methods together in the health care sector.

Lastly, a Turkish study conducted by M. S Gok and B. Sezen in 2012 called *"Analyzing the ambiguous relationship between efficiency, quality and patient satisfaction in health care services: The case of public hospitals in Turkey"* used DEA to analyze the efficiency scores of 348 public hospitals. Moreover, a multiple regression analysis was applied to investigate the relationship between hospital efficiency and structural quality for both small, medium, and large-sized hospitals. Their findings indicated that the trade-off between efficiency and quality depends on the hospital size. In other words, there exists a negative correlation for small hospitals and a positive and significant correlation between quality and efficiency for large hospitals (Gok S. M *et al*., 2012).

3.3 Elaboration on the research question

The study conducted by L. Kjekshus and T. Hagen (2007) examined whether the cost efficiency and technical efficiency had been affected by the hospital mergers in Norway. However, they used data from 1992 until 2000. As presented in table 2.1, most hospital mergers occurred after this period. This study will use more up to date data and also try to determine whether the quality of the hospitals have been affected by the mergers. Moreover, out of the eight studies presented above, five used the DEA estimation method in their analyses. This study will use the DEA Simar & Wilson estimation method (DEA-SW) and a Difference-in-Differences (DiD) estimation. The main difference of the previous studies presented in section 3.2 and this study will be the inclusion of quality in relation to the hospital mergers and the methodology for the analysis.

3.3.1 Main concepts from the research question

The research question in this thesis contains two main concepts while using two main estimation methods. The first main concept is productivity. The terms 'efficiency' and 'productivity' are used interchangeably in the literature. Even though they are not precisely the same thing, the distinction between the terms will not be taken into consideration in this study (Jacobs *et al.,* 2006). Jacobs *et al*. (2006) have defined productivity as *"the ratio of some (or all) valued outputs that an organization produces to some (or all) inputs used in the production process"* (Jacobs *et al*., 2006, 4). Moreover, increasing efficiency requires a reduction of the unnecessary use of resources that are utilized for a given output. The second main concept is quality, which is defined by WHO as; *"the extent to which health care services provided to individuals and patient populations improve desired health outcomes. In order to achieve this, health care must be safe, effective, timely, efficient, equitable and people-centered"* (WHO). In this manner, efficiency of health care refers to the avoidance of waste regarding recourses used to treat individuals and patients. Moreover, the quality of health services should not be affected by individual characteristics such as geographical location, gender or ethnicity. The three concepts presented here are interrelated when it comes to providing the best possible health care to the population of interest. Furthermore, to obtain the best possible practice it is important, not only consider the cost of treatment, but also the health outcomes for the patients. Low cost may be associated with poor health outcomes for the patients, vice versa (Gutacker *et al.*, 2013; Kittelsen *et al.*, 2015; McKay & Deily, 2008). The implementation of the Health Care Reform of 2002 and the activity-based financing (ABF) implemented in 1997 are two mechanisms that have been introduced in the Norwegian health care sector with the purpose of increasing both efficiency and quality at the hospitals. However, increasing efficiency in hospitals may come at the expense of the quality of the health services.

4 DATA

This chapter will include an overview and description of the aggregated datasets. Moreover, quality indicators and patient safety indicators will be presented and discussed. To conduct this study, hospital data from a long-term perspective had to be obtained to investigate the effects of the hospital merges that, for the most part, occurred after the Norwegian Hospital Reform of 2002. In addition, patient-level data had to be aggregated to measure traveling distance for each patient from their municipality to the given hospital and to obtain patient characteristics to adjust for differences in case-mix between hospitals.

4.1 Hospital data

This study are based on quantitative data of Norwegian hospitals obtained from 1999 until 2014. The hospital data is a panel dataset, which means that the hospitals are followed over time (Wooldridge, 2016). This data is not individual data but aggregated on the hospital-level. Hence, ethical considerations regarding confidential information about patients are excluded. The hospital data includes data from 22 hospitals (in 2014) in Norway for over 15 years. This makes it possible to discover long-term effects of the hospital merges. However, there exist more hospitals in Norway but many of them have merged within the same HF. Note that there are a larger number of hospitals in the earlier years compared to the number of hospitals in 2014. This is a result of hospitals merging into health enterprises and implemented under the same name in the hospital data.

4.2 Patient-level data

The Norwegian Patient Register, owned by the Norwegian Directorate of Health, provides the patient-level data used in this study. There are two separate patient-level dataset where the first dataset contains patient information from 1999 until 2007, while the latter dataset contains patient information from 2008 until 2014. The data collected from 1999 until 2007 includes information from 95 different health institutions, while the 2008-2014 data contains information from 83 health institutions. Moreover, in total the dataset includes information from approximately 83,000,000 patients that have received hospital care in Norway from 1999 until 2014. Note that the same patient is registered separately for each hospital visit. Thus, if a patient has visited hospitals or health institutions ten times in a five-year period, the patient is registered ten times in the dataset.

The patient-level data will be used to compute the case-mix adjustments for each hospital. This is essential to compare hospitals without having issues regarding the respective hospital's composition of patients that may influence the result. DRG-weights will be used to distinguish the cost of patient groups and treatment from each other. Since the patient-level data contains information from a more comprehensive number of hospitals and clinics than the hospital data, several hospitals and health institutions have been excluded when merging the two datasets. Attachment C and D in the appendix gives an overview of the full list of omitted hospitals and clinics from both datasets, and the reasoning why they are omitted. In total, 62 and 47 of the hospitals included in the 1999-2007 patient-level data and 2008-2014 patient-level data are used in the analysis, respectively, while 33 and 36 health institutions are excluded.

4.3 Diagnosis-Related Groups (DRG)

DRG is a classification system of patients that classify patient groups. The Norwegian DRG system is called NordDRG and is based on the Nordic system, where the Nordic Centre for Healthcare Classifications are responsible for updating, maintenance, and development of the NordDRG system (Nordic Casemix Centre, 2013). DRG attempts to handle a complex reality where patients are diagnosed with thousands of different diagnoses and as many different treatment options. The system gives an overview of hospital activity. Moreover, the DRG system makes it possible to compare hospital activity even though the hospitals are treating different patients. When patients arrive at the hospital, information regarding sex, age, the individual´s diagnosis, and discharge status are collected to give them a DRG-weight. The DRG-weight gives information about how demanding the patient is compared to the average patient registered in the system (Helsedirektoratet, 2019). The DRG-system is regularly updated due to improvements in treatment methods, but also since the practice of treatments are changing. This might lead to changes in cost of treatment for patients (Sosial- og helsedirektoratet, 2007). The hospitals receive a price that reflects the DRGs of their patients. Furthermore, if the hospitals can treat their patients below the price received, the hospitals keep the surplus. On the other hand, if the hospital spends more than the price, they are facing a deficit. In this study, the data contains 758 DRGs, which are used in the case-mix adjustments of Norwegian hospitals.

4.3.1 Calculation of DRG-weights

DRG-weights represents the mean cost of treatment for a certain patient group. It is the differences between the DRG-weights that are of interest, and not the absolute value of the DRG-weights themselves. The DRG-weights are calculated by an estimation of the mean cost for each hospital stay per DRG. Then, the median value of the means is used in future calculations since the median is more robust when it comes to extreme values than the mean. Finally, the DRG-weights are estimated as the ratio between the median value per DRG and the mean of all median values (Kvæl, L, 2005). To reflect the real cost within each DRG group, DRG-weights are yearly updated to catch up inflation and increased costs of treatment. In 2012, there were calculated new weights of costs based on both costs- and activity data from 2010 from a sample of hospitals. The weights are found by a calculation model that distribute the hospitals' costs down to the given patient stay grouped to one DRG. The model distinguish costs of day-DRGs and outpatient-DRGs. The calculation model for the day-activity have previously been used with small alteration every year. However, the model for costs of outpatient consultations was used for the first time in 2010 (Helsedirektoratet, 2011).

4.4 Quality indicators, length of stay, and travel time

Most patients that consume health services prefer to receive treatment of higher quality rather than lower quality. Thus, when there has been an increase in quality levels of the health services, one may assume that consumers also are willing to pay more for the health services. Patients desire to receive health care services with a certain quality level. Quality is a broadly used term in health care analyses and is a central term in this thesis. Quality refers to better health outcomes for the patients. In general, quality levels of a hospital will be hard to measure because health service is a credence good, which means that the consumer cannot observe the outcomes of the treatments and hence it is difficult for the consumers to assess its utility and quality. This may induce information asymmetry between the patient and the doctors (Emons, W, 1997, 107). Mortality- and readmission rates are commonly used as quality indicators in the literature. Quality indicators are useful tools that provide information about the quality on the field that is being studied for both health care providers, personnel, consumers, and public health policymakers and are often connected to process, structure, and outcomes (Breyer *et al.*, 2019). This study will focus on clinical quality such as in-hospital mortality and readmission. Since hospitals and policymakers always want to keep the mortality rates low, mortality is an appropriate quality indicator to use in the health care sector. Moreover, mortality rates can potentially apply to large patient groups (Kittelsen *et al.*, 2017, 76). Mortality says something about the safety at the hospital and high mortality rates are associated with poor quality of the health services.

Mortality is a clear outcome, but it also has its limitations as a measure of quality. Even though one correct for the patient characteristics one can observe, there may be underlying differences in severity and forecasts of patients between hospitals. Furthermore, mortality rates are considered to have high levels of noise since patients die from other factors that are not associated with their hospital stay (Kittelsen *et al.*, 2017, 78). However, mortality rates are the most commonly used and widely accepted quality indicator. Lowering the mortality rates of a hospital is always an improvement even though some mortalities are unavoidable (Kittelsen *et al.*, 2015, 146). Readmission rates are also a quality indicator that is widely used in health care performance. However, some level of readmission is unavoidable, but having high levels of emergency readmission rates could be a sign that the initial treatment was not adequate (Kittelsen *et al.*, 2015, 46).

This thesis will estimate two indicators of quality and five patient safety indicators in order to measure the levels of performance across the 22 hospitals. Note that the last hospital the patient was registered at are also the hospital that are being held "accountable" for the patient's death in the data. This is especially relevant if the patient has been transferred across different hospitals. The focus will be on emergency readmission- and in-hospital mortality. For the patient-level data collected from 1999 until 2007, patients were registered with an identification number. January 1st each year, these numbers were renewed. Hence, in order to measure the readmission- and mortality rates for patients hospitalized from 1999 until 2007, patients who has been discharged from the hospital in December are excluded from the study since it is not possible to detect their readmission- or mortality rates within 30 days. Moreover, the patient level data from 1999 until 2007 include only information concerning whether the patient was discharged as living or dead. It is not possible to calculate after-discharge mortality for these patients. Hence, the in-hospital mortality of all patients in this study will be based on whether the patient was discharged as living or dead and not as mortality after 30 days of discharge. Even though the patient level data obtained from 2008 until 2014 does not suffer from these inconveniences, the same procedure has to be applied in order to compare the two datasets.

4.4.1 Patient safety indicators

Multiple patient safety indicators can be used to estimate quality for hospitalized patients. The OECD working papers from 2009 present patient safety indicators such as deep vein thrombosis (PSI12vt_pe), sepsis (PSI13sepsis), accidental cut, puncture, or hemorrhage (PSI15ac_punc), obstetric trauma (PS18ob_trau), and bedsores (bed_sores) (OECD, 2009). These five patient safety indicators will be included in this study because they cover a wide range of diseases that may be a cause of insufficient patient care and are based on secondary diagnoses (Kittelsen *et al*., 2015). Moreover, the patient safety indicators mentioned above provides information regarding adverse events and in-hospital complications as a result from childbirth, different procedures, and surgeries (OECD, 2009).

Deep vein thrombosis occurs when a blood clot appears in a vein that is located deep inside the body and is a serious condition. The most common causes are injury or surgery with damage to a blood vessel and the patient is bedridden without movement after surgery (Delgado & Holland, 2018). Sepsis occurs when the body tries to fight infection by releasing chemicals into the bloodstream and when the body’s response to the releasing chemicals is out of balance. This is a potentially life-threatening condition (Mayo Clinic Staff, 2018). Obstetric trauma can be caused by stress after delivery. Furthermore, a third of all women experience their birth as traumatic (Oslo Universitetssykehus, 2017). Post-traumatic stress after delivery can be caused by the birth experience in general, operational interventions, lack of support, and dissociation. However, the potential consequences are related to future pregnancies and women´s ability to create bonds and relationships with their infants. Finally, bedsores are caused by pressure on the skin that results in injuries to the skin and the underlying tissues. Severe cases of bedsores can lead to life-threatening complications and result in blood poisoning.

In addition to the patient safety indicators above, the length of stay (LOS) is commonly used as an indicator of efficiency where longer stays are associated with increased costs per discharge (OECD, 2018). Moreover, the length of stay at the hospital can be associated with the hospitals’ ability to follow up their patients after treatment. This may be valuable, especially for primiparas (women who are giving birth for the first time). According to the Norwegian Institute of Public Health, the average length of stay (ALOS) at the hospital after giving birth in 2017 was 2.8 days and has been stable for the last five years. However, from 1999 until 2008 ALOS decreased from 4.1 to 3.1. It is important to mention that, in general, primiparas have a greater ALOS than multiparous women (Askeland, O. M, 2018).

Travel time is used as an environmental variable in the analysis and represents travel time in hours by car from the home municipality to the given hospital. Merging hospitals may affect the patients travelling time to their closest health institution if their local hospital is being closed due to reorganizations. Travel time is an explanatory variable that varies between hospitals. Hence, it can contribute to distinguish costs and quality levels between the hospitals since the variable has previously shown explanatory power (Kalseth *et al.*, 2011). In this analysis, the travel time estimation follows the same procedure as Kittelsen *et al.* (2015) where Google maps are calculating travel time by using STATA.

Table 4.1 List and definitions of quality indicators and patient safety indicators

Group	Variable name	Definition
Quality indicators		
Readmission	readm30_emgc	Patient admitted acutely to inpatient care in hospital within 30 days of the discharge
Mortality	inhosp_mortality_1	In-hospital mortality from any cause.
Patient safety indicators		(PSI indicators are defined by OECD)
	PSI12vt_pe	Deep vein thrombosis
	PSI13sepsis	Sepsis
	PSI15ac_punc	Accidental cut, puncture, or haemorrhage during medical care
	PSI18ob_trau	Obstetric trauma
	bed_sores	Bedsores

4.5 Descriptive statistics at hospital level

Descriptive statistics of the quality indicators, patient safety indicators and average length of stay (ALOS) introduced in section 4.4 are computed based on each of the Norwegian health enterprises (HF) and presented in table 4.2 below. Note that Diakonhjemmet sykehus and Lovisenberg Diakonale sykehus lacks observations on obstetric trauma, therefore these values are omitted from the table. This may be because these hospitals do not have a maternity ward at their hospitals. The results show that Haraldsplass Diakonale sykehus have the highest numbers of both emergency readmission and in-hospital mortality, while Møre og Romsdal HF and Oslo Universitetssykehus HF have the lowest numbers of emergency readmission and in-hospital mortality, respectively. Furthermore, ALOS is highest at Haraldsplass Diakonale sykehus, while the lowest at Møre og Romsdal HF.

Table 4.2 Descriptive statistics of quality indicators, patient safety indicators and average length of stay. Not case-mix adjusted.

	AHUS HF	Diakonhjem.	Finnma. HF	Haraldsplass	Helgela. HF	H. Bergen HF
Quality indicators						
Readn30_emgc	5.38 %	4.44 %	4.14 %	6.04 %	5.61 %	5.06 %
Inhosp_mortality_1	0.27%	0.45 %	0.21 %	0.70 %	0.25 %	0.20 %
PSIvt_pe	0.087 %	0.103 %	0.047 %	0.203 %	0.045 %	0.066 %
PSI13sepsis	0.061 %	0.073 %	0.054 %	0.152 %	0.031 %	0.086 %
PSI15ac_punc	0.032 %	0.013 %	0.010 %	0.015 %	0.009 %	0.021 %
PSI18ob_trau	0.053 %		0.020 %		0.019 %	0.027 %
Bed_sores	0.021 %	0.053 %	0.026 %	0.062 %	0.024 %	0.023 %
Length of stay						
ALOS	1.79	1.81	1.64	2.60	1.61	1.76

	H. Fonna H	H. Førde HF	Lovisenberg	MøreRo. HF	Nordland. HF	Nordmøre HF
Quality indicators						
Readn30_emgc	5.54 %	3.98 %	3.94 %	3.48 %	4.86 %	4.73 %
Inhosp_mortality_1	0.27 %	0.21 %	0.43 %	0.16 %	0.24 %	0.27 %
PSIvt_pe	0.064 %	0.048 %	0.154 %	0.055 %	0.064 %	0.058 %
PSI13sepsis	0.061 %	0.049 %	0.060 %	0.041 %	0.057 %	0.038 %
PSI15ac_punc	0.015 %	0.019 %	0.004 %	0.012 %	0.020 %	0.011 %
PSI18ob_trau	0.026 %	0.017 %		0.010 %	0.020 %	0.007 %
Bed_sores	0.031 %	0.059 %	0.041 %	0.036 %	0.024 %	0.026 %
Length of stay						
ALOS	1.79	1.63	1.91	1.48	1.72	1.68

	Nord-Tr. HF	OUS HF	St. Olavs HF	Stavanger HF	Sunnmøre HF	Innlandet HF
Quality indicators						
Readn30_emgc	5.26 %	5.09 %	4.93 %	5.45 %	4.28 %	5.10 %
Inhosp_mortality_1	0.27 %	0.11 %	0.21 %	0.23 %	0.21 %	0.27 %
PSIvt_pe	0.052 %	0.068 %	0.062 %	0.072 %	0.049 %	0.065 %
PSI13sepsis	0.060 %	0.057 %	0.062 %	0.052 %	0.027 %	0.046 %
PSI15ac_punc	0.020 %	0.022 %	0.017 %	0.024 %	0.014 %	0.022 %
PSI18ob_trau	0.011 %	0.009 %	0.016 %	0.034 %	0.019 %	0.014 %
Bed_sores	0.023 %	0.013 %	0.020 %	0.023 %	0.044 %	0.025 %
Length of stay						
ALOS	1.69	1.52	1.72	1.75	1.74	1.68

	Telemark HF	Sy. Vestfold HF	Sy. Østfold HF	Sørlandet sy. HF	UNN Tromsø HF	Vest.Viken HF
Quality indicators						
Readn30_emgc	5.58 %	5.08 %	5.99 %	5.16 %	4.37 %	5.66 %
Inhosp_mortality_1	0.26 %	0.25 %	0.31 %	0.23 %	0.19 %	0.23 %
PSIvt_pe	0.062 %	0.071 %	0.071 %	0.054 %	0.074 %	0.100 %
PSI13sepsis	0.056 %	0.061 %	0.059 %	0.053 %	0.079 %	0.068 %
PSI15ac_punc	0.012 %	0.012 %	0.012 %	0.016 %	0.020 %	0.024 %
PSI18ob_trau	0.020 %	0.005 %	0.019 %	0.024 %	0.009 %	0.018 %
Bed_sores	0.021 %	0.016 %	0.019 %	0.065 %	0.037 %	0.052 %
Length of stay						
ALOS	1.74	1.63	1.69	1.60	1.64	1.64

The table is based on outputs retrieved from STATA 16.

Descriptive information of patient characteristic, treatment variables, and travel time based on each of the Norwegian HFs are found in attachment E in the appendix. Haraldsplass Diakonale sykehus is the only hospital that has a majority of men. Moreover, Diakonhjemmet sykehus, Lovisenberg Diakonale sykehus, and Haraldsplass sykehus have very low number of patients aged 0 to 19 which may follow from their lack of maternity wards at their hospitals. The majority of the patients in the datasets are aged between 40 and 79. Furthermore, the average travel time are the longest for UNN Tromsø HF and the shortest for Diakonhjemmet sykehus. Descriptive statistics of costs, number of patients and capital for each of the HFs are to found in attachment F in the appendix.

Table 4.3 Descriptive statistics of quality indicators and patient safety indicators sorted by year.

	1999	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**
Patients	3640314	3814268	3969212	4073581	4384821	45694496	4694496	4847595	5047511
Readm30_emgc	4.39 %	4.39 %	4.55 %	4.00 %	4.57 %	4.59 %	4.82 %	4.67 %	5.43 %
Inhosp_mortality_1	0.44 %	0.41 %	0.40 %	0.35 %	0.35 %	0.33 %	0.32 %	0.29 %	0.29 %
PSI12vt_pe	0.064 %	0.065 %	0.067 %	0.060 %	0.072 %	0.072 %	0.071 %	0.072 %	0.076 %
PSI13sepsis	0.048 %	0.045 %	0.055 %	0.050 %	0.058 %	0.057 %	0.061 %	0.063 %	0.066 %
PSI15ac_punc	0.011 %	0.011 %	0.011 %	0.010 %	0.015 %	0.013 %	0.015 %	0.017 %	0.020 %
PSI18ob_trau	0.032 %	0.032 %	0.033 %	0.035 %	0.035 %	0.032 %	0.027 %	0.021 %	0.019 %
Bed_sores	0.016 %	0.018 %	0.018 %	0.018 %	0.022 %	0.023 %	0.023 %	0.024 %	0.024 %
Number of hospitals	55	54	54	36	30	30	29	29	28

	2008	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**
Patients	5178468	5589004	5741045	6072454	6237673	6321031	6507254
Readm30_emgc	5.38 %	5.35 %	5.20 %	4.79 %	4.91 %	4.80 %	4.65 %
Inhosp_mortality_1	0.25 %	0.22 %	0.21 %	0.19 %	0.17 %	0.17 %	0.15 %
PSI12vt_pe	0.077 %	0.071 %	0.074 %	0.071 %	0.068 %	0.066 %	0.064 %
PSI13sepsis	0.062 %	0.060 %	0.063 %	0.060 %	0.061 %	0.057 %	0.059 %
PSI15ac_punc	0.020 %	0.019 %	0.019 %	0.020 %	0.020 %	0.021 %	0.021 %
PSI18ob_trau	0.018 %	0.018 %	0.016 %	0.014 %	0.014 %	0.012 %	0.013 %
Bed_sores	0.026 %	0.027 %	0.031 %	0.034 %	0.031 %	0.032 %	0.030 %
Number of hospitals	28	23	22	22	22	22	22

The tables are based on outputs retrieved from STATA 16.

Table 4.3 above contains descriptive statistics of quality indicators and patient safety indicators (PSI) per year for both merged and non-merged hospitals. In addition, the numbers of patients treated at every hospital included in the analysis per year are presented. Note that the number of patients may be affected by multiple registrations per patient and may not reflect the true number of individuals treated, where every case concerning the same patient are being registered separately. One can see that the number of treated patients has been increasing every year, and the numbers have almost doubled from 1999 until 2014. The share of patients who have been readmitted to the hospitals after 30 days of discharge due to emergency have increased from 2007 until 2013, while in-hospital mortality has decreased gradually over time. In general, the incidences of all PSIs have increased after 2002 except from obstetric trauma, whether these changes are significant or not will be elaborated on in chapter 6.

5 METHODOLOGY

This chapter will include the methodological approaches that consist of three steps. First, case-mix adjustments of quality indicators have to be estimated to control for differences regarding hospitals' composition of patients. Second, an efficiency score and Simar & Wilson regression of each hospital in the data from 1999-2014 will be estimated with the DEA-Simar & Wilson approach. Third, treatment and control groups for the Difference-in-Differences (DiD) analysis have to be determined, where the intervention for the analysis is hospital mergers. A DiD analysis will be performed based on the efficiency estimates between merged and non-merged hospitals. Furthermore, additional analysis of quality will be estimated using DiD and taken into considerations in the result section.

5.1 Case-mix adjustments of the hospitals

The case-mix of patients may vary across hospitals. If the case mix is not adjusted or adequately accounted for in the analysis, the more complex and larger hospitals with a variety of patients may be estimated as being less efficient (Jacobs *et al*., 2006, 112) Hence, adjusting for case-mix when comparing hospitals' efficiency and quality is important. The adjusting variables should capture the characteristics of the patients and their illnesses that could influence the outcome of the analysis. For that reason, using DRG-weights, which are designed to capture differences in patient characteristics that may influence the costs of the hospitals, will also capture factors that influence the values of the quality indicators (Kittelsen *et al.,* 2015, 146).

The obtained data include 758 DRGs that are used in the case-mix adjustments. Kittelsen *et al.* (2015) have previously adjusted for differences in case-mix between Norwegian hospitals in six models. Model 0 captured Nordic DRG, model 1 consisted of Nordic DRGs and patient characteristics, model 2 included treatment variables and the two previous components. The treatment variables are denoted "*transin-"* and "*transout-"* which represents transfers in and out of hospitals and are meant to explain the movement of patients. In model 3, an additional variable concerning length of stay was included, while model 4 also captured municipal variables for patients. Finally, model 5 included all previous components together with travel time between the treating hospital and the resident municipality. Kittelsen *et al.* (2015) found statistical evidence that Model 2 seemed to be favorable. The results from their analysis showed that the R-squared, which represents how much the independent variables explain the variance in the dependent variable, increased the most by including treatment variables together with the previous variables (DRGs and patient characteristics). For that reason, I will use model 2 in the case mix-adjustments of quality indicators. The description of the variables are found in table 5.1 below.

Table 5.1 Case-mix adjustment models and definition of variables used in the calculations

Group	Variable name	Definition
Case-mix adjustment variables		
Model 0 + DRG	DRG	Diagnosis related group based upon common Nordic grouper
Model 1 + Patient characteristics	male agegrp0 agegrp1_9 … agegrp80_89 agegrp90	1=Male, 0= Female Age dummies for the groups: 0, 1-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90 and above.
Model 2: + Treatment variables		
	transinownhospital transinotherhosptal transoutownhospital transoutotherhospital	Dummies for transfer into and out of hospital department stay within one day before or after this stay.
	wcharlsum_sec	Charlson index of comorbidity based upon secondary diagnosis
	n_sec	Number of secondary diagnoses

To calculate the observed-to-expected ratio of each quality indicator for all hospitals, I will follow Ash *et al*. (2003) as cited in Kittelsen *et al*. (2015). For model 2 *(m* = 2) the performance measures will be estimated. Each patient is denoted by *i* and has an observable quality indicator, which is represented by ω_{ihk}. Moreover, *i* has an expected quality indicator given by $\widehat{\omega}^2_{ihk},$ which is subscripted by hospital $h \in (1,\ldots,H)$ and DRG $k \in (\mathrm{I},\ldots K)$ and superscripted by model 2.

The case-mix adjusted hospital performance measures are represented by P_h^2, where the hospital is denoted by h and model 2 is denoted by 2. N_{hk} represents the number of patients in DRG k at hospital h. The first step is to summarize all observed patient outcomes and all expected patient outcomes, respectively. The second step is to divide the first component on the latter component, and is expressed by the following equation:

$$P_h^2 = \frac{\sum_{k=1}^{K} \sum_{i=1}^{N_{hk}} \omega_{ihk}}{\sum_{k=1}^{K} \sum_{i=1}^{N_{hk}} \widehat{\omega}_{ihk}^2} \tag{5.1}$$

Lower values indicate better qualities, which applies both for the quality indicators and the performance measure, P_h^2. The way we predict $\widehat{\omega}_{ihk}^2$ affect the performance measures in model 2. In Model 0, we examine and take advantage of the different DRG compositions at each hospital. Thus, the predicted quality indicator for patient *i*, $\widehat{\omega}_{ihk}^0$, is given by the mean value of the quality indicator within each DRG for all patients across all hospitals (Kittelsen *et al.,* 2015). The formula below is used to obtain predicted outcomes for Model 0, which does not depend on *i* nor *h*, and is equal for all patients in DRG *k*.

$$\widehat{\omega}_{ihk}^0 = \frac{\sum_{g=1}^{H} \sum_{j=1}^{N_{gk}} \omega_{jgk}}{\sum_{g=1}^{H} N_{gk}} \tag{5.2}$$

Furthermore, the predicted quality measure can be additionally improved by conditioning on patient characteristics. It is appropriate to use the logit model to estimate the conditional probability since all quality indicators in this study are binomial variables (Greene, 200; Hosmer *et al*., 2013, cited in Kittelsen *et al.,* 2015). Since there are many observations included in the data, it is possible to estimate the expected value as the predicted value, which is based on the maximum likelihood estimation, as shown in the following formula;

$$\omega_{ihk}^2 = \frac{e^{\beta_{0k}^2 + \boldsymbol{\beta}_k^2 \mathbf{z}_{ihk}^2 + \varepsilon_{ihk}^2}}{1 - e^{\beta_{0k}^2 + \boldsymbol{\beta}_k^2 \mathbf{z}_{ihk}^2 + \varepsilon_{ihk}^2}} \tag{5.3}$$

Where ω_{ihk}^2 represents the quality measure for patient *i* in DRG *k* at hospital *h*. The coefficient vectors β_{0k}^2 and $\boldsymbol{\beta}_k^2$ are specific to the DRG *k* and model 2. Furthermore, the vector $\boldsymbol{z}_{ihk}^2$ corresponds to individual case-mix adjusting variables. Finally, ε_{ihk}^2 is the error term in the formula, which is here assumed to be normally distributed (Kittelsen *et al.,* 2015). The predicted value can be found by setting $\varepsilon_{ihk}^2 = 0$. This study will be based on model 2, which includes DRG, two quality indicators, five patient safety indicators, four treatment variables, the Charlson index, which are based upon secondary diagnosis, and number of second diagnoses.

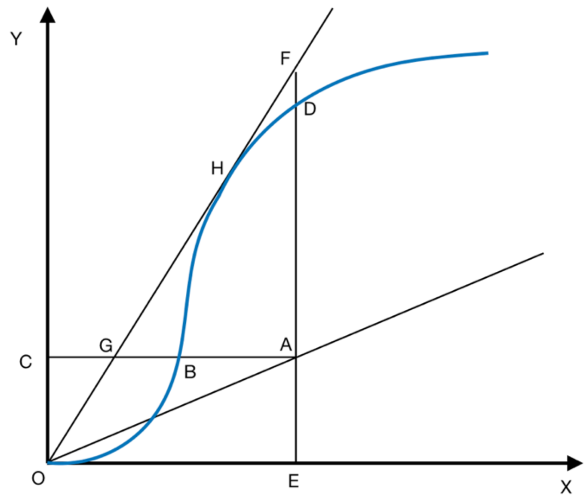
5.2 Measurement of efficiency

The terms efficiency and productivity have been used interchangeably throughout the literature and they are used differently between professional environments. Based on microeconomic production theory, productivity is defined as the ratio between production and resource use, while efficiency is being defined as the ratio between actual productivity and the best possible productivity that is feasible given the available resources (Anthun *et al.*, 2016, 6). An efficient hospital provides most health services within their given budget. The main problem by measuring efficiency arise when the institution consists of multiple services or multiple inputs, which are the amount of inputs that are used to produce a certain output (Anthun *et al.*, 2016, 6).

Farrell (1957) proposed that efficiency could be divided into two terms: allocative efficiency (AE) and technical efficiency (TE)[2]. Allocative efficiency refers to the optimal mix of input recourses. However, the input prices are not known is this study and allocative efficiency cannot be measured. High technical efficiency indicates that there is no excess input of recourses used to obtain a certain production level (Johannesen *et al.*, 2017, 120). TE and cost efficiency (CE) are representing the same in this study. The third term related to efficiency is scale efficiency (SE). Being scale efficient indicates that the size of the unit is optimal and that any modification on the units' size will result in inefficiency.

Efficiency measures can be either input-orientated or output-orientated. Whereas input-orientated measures address the question; *"By how much can input quantities be proportionally reduced without changing the output quantities produced?"* (Coelli *et al*., 2005, 54). The output-orientated measures address the question: *"By how much can output quantities be proportionally expanded without altering the input quantities used?"* (Coelli *et al*., 2005, 54).

[2] The terminology used in the paper "The Measurement of Productive Efficiency" from 1957 by M. J. Farrell differs from that used here. Farrell used the term *overall efficiency* instead of *economic efficiency* and *price efficiency* instead of *allocative efficiency*. The terminology used in this thesis corresponds with the terminology mostly used in recent literature.



5.1 TE, TP and SE *(Based on Anthun et al., 2016, 9)*

Figure 5.1 above illustrates TE for both input- and output orientation. Moreover, it shows technical productivity (TP) and scale efficiency (SE) for both input- and output orientation. As mentioned above, efficiency is the ratio of actual productivity to the best possible productivity given the available inputs and outputs. The efficiency depends on the technology and the frontier (OBHD). TE can be subdivided into two orientations, where the first is input technical efficiency (5.4) and the latter is output technical efficiency (5.5). In figure 5.1, TE_{IN} and TE_{OUT} of a firm or organization is calculated by the ratios:

$$TE_{IN} = \frac{CB}{CA} \tag{5.4}$$

$$TE_{OUT} = \frac{EA}{ED} \tag{5.5}$$

Equation 5.6 below yields a number between zero and one where a value of one indicates that the firm is fully technical efficient. Point H in figure 5.1 is an example of a fully efficient firm because it lies on the isoquant curve. However, the production frontier of fully efficient firms is not known in practice and has to be estimated by using either DEA or SFA (Coelli *et al.*, 2005). Furthermore, Coelli *et al.* (2005) state that the input-orientated measure of TE of a firm can be articulated in terms of input-distance function $d_i(\boldsymbol{x}, \boldsymbol{q})$ as:

$$TE = 1/d_i(\boldsymbol{x}, \boldsymbol{q}) \tag{5.6}$$

meaning that a firm or organization is considered to be technical efficient if the observation lies on the frontier and $d_i(\boldsymbol{x}, \boldsymbol{q})$ is equal to one. If the information on input prices is present, measurement of cost efficiency (CE) of the firm would be achievable. However, in this study, CE and TE are the same due to only having costs as input variable in the analysis.

Two commonly used methodologies can be applied to measure efficiency. The first is the Data Envelopment Analysis (DEA) that was developed in 1978 and is a method used to measure efficiencies of decision-making units (DMUs). DEA is a data-driven and non-parametric approach where the shape of the efficiency frontier is determined by the data (Anthun *et al.*, 2017, 420). In other words, DEA does not require assumptions regarding the functional form of the production function. On the other hand, DEA assumes no measurement errors or random noise, and the results are sensitive to outlier observations (Jacobs *et al.*, 2006, 112). If the chosen method is based on these outlier observations, the result may be biased. Hence, the measurement of the DMUs has to be accurate and concise to reduce the risk of measurement errors and biased results. The second approach is the Stochastic Frontier Analysis (SFA) that has a specific functional form. Moreover, SFA can only estimate one input or output and has the ability to incorporate measurement errors and outliers (Anthun *et al.*, 2017, 420). Figure 5.2 below illustrates the differences of the SFA- and DEA methodology. Using the DEA method in the health care sector in Norway seems to be applicable since substantial measurement errors are unlikely and the method yields an opportunity to estimate multiple inputs or outputs (Anthun *et al.*, 2017, 420).

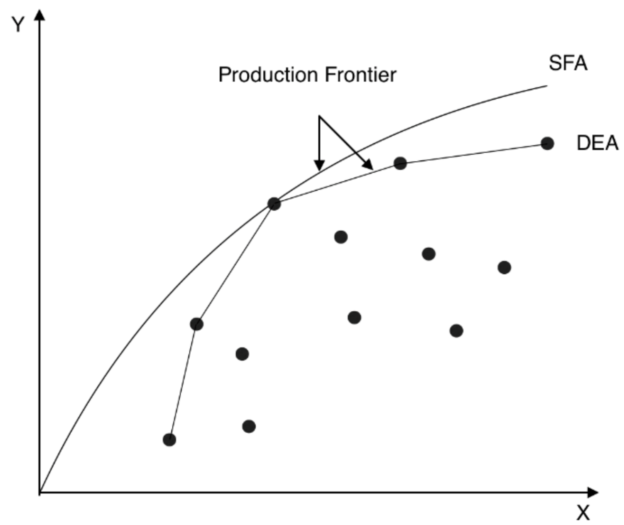


Figure 5.2 Comparison of the SFA- and DEA method.

5.3 Simar & Wilson two-stage approach with Data Envelopment Analysis

DEA is based on linear programming and was introduced in 1978 by Charnes, Copper and Rhodes. The two-stage approach consists of estimating efficiency scores using DEA and a Simar & Wilson regression to explain the variation in efficiency scores in the second stage (Badunenko & Tauchmann, 2018, 33). DEA-SW was implemented in STATA in 2019 by Badunenko and Tauchmann, and contains the new command simarwilson, which will be used in this analysis.

5.3.1 Data Envelopment Analysis (DEA) and DEA Simar & Wilson (DEA-SW)

To be more efficient means that you produce the same amount of output with less input (Jacobs *et al.*, 2006, 91). If it is possible to produce the same level of all outputs by using less of at least one input and not more of any other inputs, the unit is inefficient. Alternatively, if it is possible to produce more of at least one output and not less than any other output by using no more of any input, the unit is inefficient. If the situations above are not possible, the unit is technically efficient (Petersen, 1990). The DEA approach bases the frontier on "best-observed practice". Hence, DEA is only an approximation to the true efficiency frontier. Thus, the efficiency scores obtained using the basic DEA methodology are facing an issue of sampling variation of the estimated frontier (Badunenko & Mozharovskyi, 2016). As mentioned in section 5.2, the DEA methodology uses the wording of decision-making units (DMUs) for the units that are included in the analysis. In this study, DMUs will represent all hospitals included in the analysis (Jacobs *et al*, 2006, 91).

However, the basic DEA approach has limited capability and the calculations are slow with larger data sets. This study will include personal level data from approximately 83 million observations and hospital-level data with a comprehensive set of variables. Hence, using the DEA Simar & Wilson methodology (DEA-SW), which performs radial technical efficiency analysis like the user-written command dea offers, will complete the analysis more efficiently (Badunenko & Mozharovskyi, 2016). The DEA-SW approach compute radial technical efficiency measures in DEA-models. The radial measures of efficiency are developed by Debreu (1951) and Farrell (1957), which define technical efficiency corresponding to the isoquant. In other words, radial measures require efficiency observations to lie on the isoquant (Ferrier *et al.*, 1994, 450).

For each data point k $(k = 1,\ldots,K)$, vector $x_k = (x_{k1},\ldots,x_{kN}) \in \Re^N$ denotes N inputs, vector $y_k = y_{k1},\ldots,y_{kM}) \in \Re^M$ denotes M outputs, and T represents the technology. We assume that the data (y, x) is distributed in a way that outputs are producible by inputs under T,

$$T \equiv \{(x, y) : y \text{ are producible by } x\} \quad (5.7)$$

$$P(x) \equiv \{y : (x, y) \in T\} \quad (5.8)$$

$$L(y) \equiv \{x : (x, y) \in T\} \quad (5.9)$$

where, condition (5.8) represents the output possibility set that the technology T is fully characterized by and produce outputs of a given set of inputs. Condition (5.9) corresponds to the input requirement set, which represents the inputs needed to produce a set of outputs. The two latter conditions symbolize that the outputs and inputs available are feasible. The frontier is defined by an upper boundary of the output possibility set and a lower boundary of the input requirement (Badunenko & Mozharovskyi, 2016, 3).

As mentioned in section 5.2, the DEA-method does not require assumptions regarding the functional form of the production function. However, it assumes the properties of the production possibility set (T). Halsteinli *et al.* (2001) present three basic assumptions concerning the properties of the production possibility set and these are illustrated in figure 5.3 below.

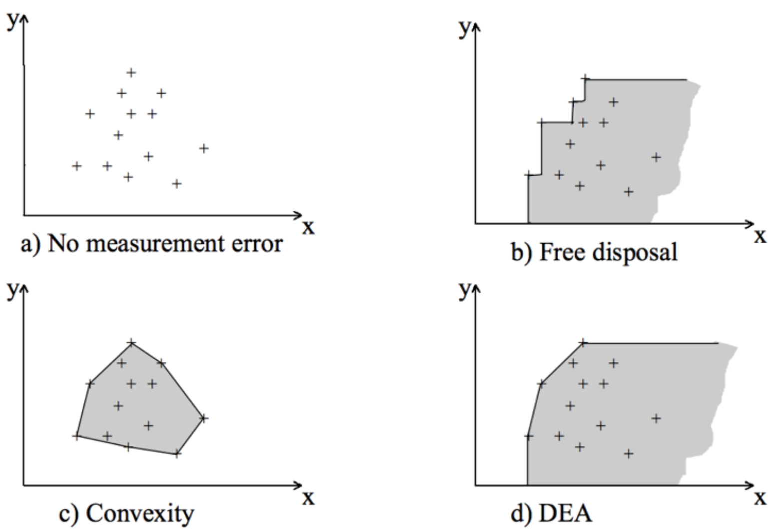


Figure 5.3 The DEA assumptions on the possibility set *(Halsteinli et al. 2001).*

First, all observed inputs and outputs for each DMU is feasible (a) and is presented by;

$$\left(\mathbf{x_i^0}, \mathbf{y_i^0}\right) \in T \tag{5.10}$$

Second, there is free disposal of inputs and outputs (b) which means that all $\mathbf{x}$ that are larger and all $\mathbf{y}$ that are smaller are feasible, and is presented by;

$$\left\{(\mathbf{x}, \mathbf{y}) | \mathbf{x} \geq \mathbf{x_i^0}, \mathbf{y} \leq \mathbf{y_i^0}\right\} \in T \tag{5.11}$$

Lastly, the production possibility set is piecewise linear convex through the observations that contribute to the production function (c), which implies linear functions in all places. In other words, observations cannot be observed to the left or below of the given observation. This can be expressed as equation 5.12 below. The DEA method combines the three elements (d).

$$\left\{(\boldsymbol{x}, \boldsymbol{y}) \begin{array}{l} |\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \lambda_2 \boldsymbol{x}_2 + \cdots + \lambda_I \boldsymbol{x}_I, \\ |\boldsymbol{y} = \lambda_1 \boldsymbol{y}_1 + \lambda_2 \boldsymbol{y}_2 + \cdots + \lambda_I \boldsymbol{y}_I \\ |\lambda_1 + \lambda_2 + \cdots + \lambda_I = 1 \\ |\lambda_1 \geq 0 \end{array}\right\} \in T \tag{5.12}$$

The DEA-SW approach is characterized by determining the efficiency scores in two stages. First, a frontier is identified by either using the input or output orientation. Second, each DMU is given an efficiency score that has been driven by a comparison of its ratio to that of the efficient DMUs (Jacobs *et al*., 2006, 98). An efficiency score is a non-negative number between zero and one where the most efficient DMUs will be given an efficiency score of one. If the DMU has been given a DEA-SW score of one, by definition, the hospital is on the frontier. A hospital that is behind the frontier is considered inefficient since it has a score of less than one. As an example, a DEA-SW score of 0.75 is interpreted as being 25% inefficient (Anthun *et al*., 2017, 420). The efficiency scores that are collected by using the DEA-SW approach compare each DMU to peers that generate a similar mix of outputs. This means that if any outputs are unique to the DMU, there will be no other peer to compare the results with and the DMU will be considered as fully efficient (Jacobs *et al*., 2006, 112). In this analysis, I will determine TE of hospitals, which means that I will investigate the hospitals´ ability to produce the maximum quantity of outputs given a set of inputs (Jacobs *et al*., 2006, 94). Since output prices are rarely accessible, most studies use TE. Which measure of efficiency you are using in your analysis does not affect how the results are given, the efficiency score will still be bounded by 0 and 1 (Jacobs *et al*., 2006, 95)

5.3.1.1 Variable-, non-increasing- and constant returns to scale

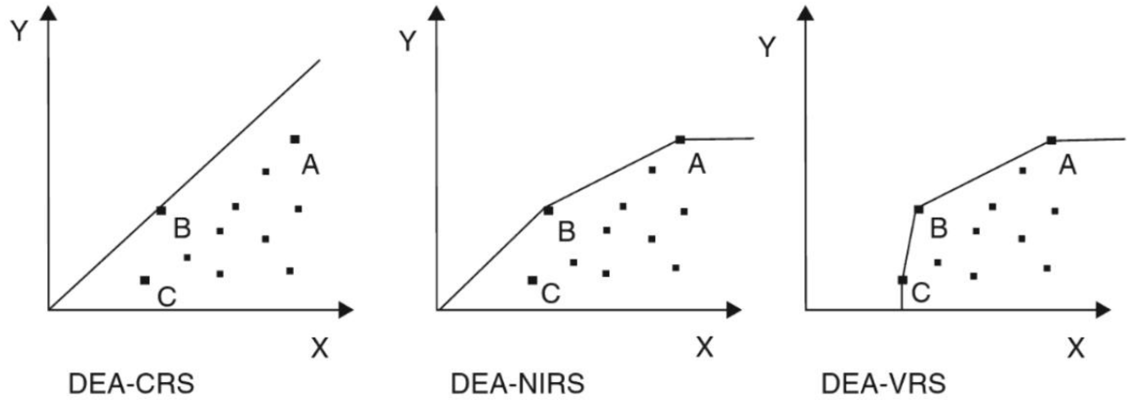


Figure 5.4 CRS, NIRS, and VRS under the DEA methodology (McCombie *et al.,* 2002)

The DEA method can be subdivided in terms of returns to scale. In economics, it is common to differentiate between constant returns to scale (CRS), non-increasing returns to scale (NIRS), and variable returns to scale (VRS). The DEA-CRS model was introduced in 1978 by Charnes, Cooper, and Rhodes and is based on the assumption that all DMUs are operating at their optimal scale. Banker, Charnes, and Cooper presented the DEA-VRS model in 1984. They introduced the assumption that not all DMUs are operating at their optimal scale. Moreover, in the DEA-VRS model, technical productivity could be broken down into scale efficiency and technical efficiency (Ji. Y & Lee. C, 2010). Where VRS allows you to make a distinction in the conversion ratio, which depends on the size of the hospital, CRS implies a linear conversion from outputs to inputs (Anthun *et al*., 2017, 420).

However, unpredictability, imperfect competition, financial constraints, and barriers to entry/exit and, hospital mergers characterize the health care sector. In addition, Norway has factors such as diversity in population pattern, administrative division, and planning which also affect the health care sector. These characteristics often yield a result that does not correspond to hospitals operating at an efficient scale. Hence, it is not reasonable to assume that all Norwegian hospitals are homogenous, in other words, they are not operating at the same level of efficiency (Badunenko & Mozharovskyi, 2016). However, when examining the effect of hospital mergers, two potential factors have to be taken into consideration. First, when multiple hospitals are merging into one HF, the merged hospital will have a mix of all the different case-mixes of the merging hospitals. Second, the size of the merged hospital will increase due to the fusion of other hospitals. If there exist diseconomies of scale, the frontier will bend downwards. Hence, if two hospitals merge, *ceteris paribus,* the hospital may move closer to the frontier due to the increase in efficiency compared to the hospitals that are considered as the best of its size. After taking these assumptions and potential factors of hospital mergers into consideration, CRS seems appropriate in this analysis of Norwegian hospital mergers.

5.3.1.2 Input- or output orientation

The DEA-model can be further subdivided into input orientation and output orientation. As defined by the book *"Measuring Efficiency in Health Care*", written by Jacobs *et al*. in 2006, input-oriented technical efficiency measures are stated to *"keep output fixed and explore the proportional reduction in input usage which is possible"* (Jacobs *et al*., 2006, 92). On the contrary, output-oriented technical efficiency measures are stated to *"keep input constant and explore the proportional expansion in output quantities that are possible"* (Jacobs *et al*., 2006, 92). If you are using a CRS approach in the analysis, the DEA results will be the same regardless if you are using an input or output orientation. However, if you are using a VRS approach the two orientations are not equivalent to each other. If you are investigating the TE in the input-orientation, the result depends on the horizontal distance to the frontier. If you are assessing the output-orientation, the result depends on the vertical distance to the frontier (Jacobs *et al*., 2006, 105). In general, studies that have investigated hospital efficiency have assumed a minimization of inputs for a given level of outputs (1) rather than maximization of outputs for a given level of inputs (2). The first represents the input orientation and is justified by the assumption that most hospitals do take the demand for health care as given. Moreover, the hospitals have to manage their inputs in order to meet the demands form the population (Lindlbauer *et al*., 2015, 1030). Therefore, the DEA-SW model will be estimated by using an input orientation in this analysis.

5.3.2 Simar & Wilson regression

Using a two-stage approach that combines both the efficiency measurement by DEA and a regression analysis that uses the DEA estimated efficiency as the dependent variable has become very popular. The reasoning behind this commonly used two-stage approach is argued to be the identification of determinants of inefficiency rather than determining the magnitude of inefficiency for each DMU (Badunenko & Tauchmann, 2018, 2). Traditionally, Ordinary Least Squares (OLS) or a censored regression have been used as the second stage regression. Simar and Wilson (2007) criticize these models for two reasons. First, none of the studies justifies the two-stage approach by describing the fundamental data generating process (DGP). Second, they stress the problem that arises when DEA efficiency estimates are correlated. This is a result of having a common sample of data that are treated as being independent observations. Consequently, the standard approaches to inference are invalid (Simar & Wilson, 2007, 32). Simar and Wilson have developed a new two-stage estimation that implies a truncated- rather than a censored regression model that considers the above issues. The new two-stage approach consider a setting where a researcher observes three types of variables, $\mathbf{x}_i$, $\mathbf{y}_i$ and $\mathbf{z}_i$, for a sample of i=1,…, N DMUs, where $\mathbf{x}_i$ is a vector of P inputs to production, $\mathbf{y}_i$ is a vector of Q outputs from production, and $\mathbf{z}_i$ represents a vector of K environmental variables. The environmental variables may affect each DMU's ability to efficiently integrate the consumed inputs to the produced outputs (Badunenko & Tauchmann, 2018, 4). The production-possibility frontier is assumed to be homogenous across DMUs, where the boundary of the production-possibility set symbolizes all combinations of inputs and outputs that are fully efficient. In this context, this means that no outputs can be increased without increasing at least one input or decreasing at least one other output (Badunenko & Tauchmann, 2018, 4). In the Simar Wilson regression, one may measure the Farrell distance input direction as V_i, which represents the input consumption $\mathbf{x}_i$ of DMU i has to be proportionally reduced so $\mathbf{y}_i$ and $\mathbf{x}_i$ are projected onto the frontier (Badunenko & Tauchmann, 2018, 4).

Figure 5.5 below illustrates the true and estimated inefficiency. The artificial sample used in the illustration consists of 20 DMUs. Hence, the sample size is quite small. However, the DEA estimated frontier (dark blue) systematically underestimates the true frontier (marked in black). The bias-corrected estimated frontier (light blue) lies consistently above the DEA estimated frontier - indicating that bias correction will adjust the estimated inefficiency upwards. As an example, DMU B lies on the DEA estimated frontier and is apparently fully efficient. However, according to the bias-corrected estimated frontier, DMU B is considered inefficient (Badunenko & Tauchmann, 2018, 6). Note that the DEA estimated frontier is convex while the production-possibility set under bias correction is not convex.

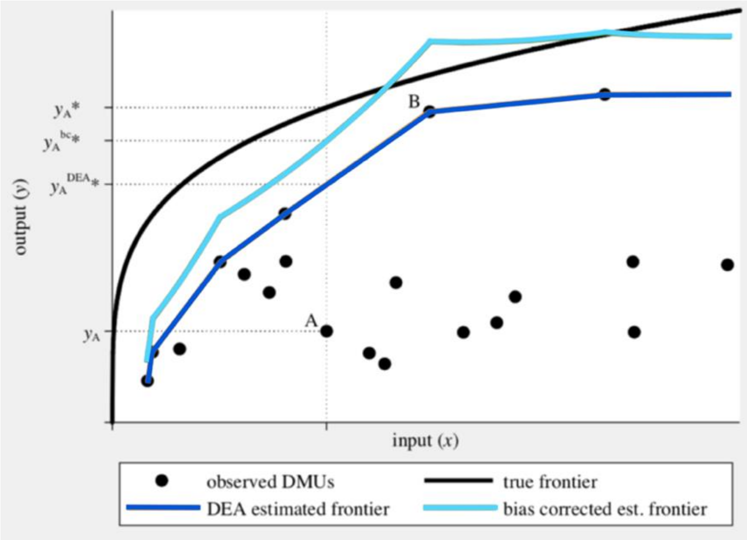


Figure 5.5 Illustration of true and estimated inefficiency *(Badunenko & Tauchmann, 2018, 6).*

Simar and Wilson present the key idea about the DGP for which the efficiency linearly depends on the environmental variables and can be expressed as; $\theta_i = \boldsymbol{z}_i\beta + \varepsilon_i$, where θ_i refers to the efficiency, β is a column vector of coefficients and ε_i is unobserved factors that is assumed to be normally distributed. The Simar and Wilson (2007) approach consist of two different procedures: algorithm 1 and algorithm 2, where the latter is more complex and involved. Algorithm 1 excludes the DMUs from the regression analysis where the estimated DEA scores are equal to one. On the other hand, algorithm 2 includes bias-corrected DEA scores and rest on a bootstrap procedure (Badunenko & Tauchmann, 2018, 7). This study will use both algorithm 1 and 2 for the Simar-Wilson regression, which include 10,000 bootstrapped replications. However, only the results from algorithm 2 will be presented in chapter 6.

There is only one input variable in the dataset. Hence, costs are the only input variable in the analysis. As for the output variables, DRG-points for regrouped patients treated are used under four different groups: emergency patients, elective patients, day patients and outpatients. Contextual variables are added to the analysis to explain the variation in the input variable since these variables may be associated with costs. The patient safety indicators included in the analysis represents the performance indicator estimated in model 2 for each variable.

Table 5.2 List of output- and input variables used in DEA-SW analysis.

Group	Variable name	Definition
INPUT	cost	Deflated costs in NOK
	v_emrg_p	DRG-points regrouped emergency inpatients
OUTPUT	v_elective_p	DRG-points regrouped elective inpatients
	v_day_p	DRG-points regrouped day patients
	v_out_p	DRG-points regrouped outpatients
Contextual variables	Average_traveltime	Average travel time (by car) in hours between hospital and center of home municipality
	ALOS	Average length of stay defined as discharge date – admission date + 1.
	Treated	0 = non-merged, 1 = merged.
	DID	Treated hospitals after merging.
	Perf_u2_readm_30_emgc	Patient admitted acutely to inpatient care in hospital within 30 days of the discharge
	Perf_u2_inhosp_mortality	In-hospital mortality from any cause.
	Perf_u2_PSI12vt_pe	Deep vein thrombosis
	Perf_u2_PSI13sepsis	Sepsis
	Perf_u2_PSI15ac_punc	Accidental cut, puncture, or haemorrhage during medical care
	Perf_u2_ PSI18ob_trau	Obstetric trauma
	Perf_u2_Bed_sores	Bedsores
	i.year /YearDiff	Year dummies with 1999 as reference/ time trend each year

5.4 Difference-in-Differences estimation of hospital mergers

The Difference-in-Differences (DiD) estimation is applied to measure sudden environmental- or policy changes on a group of individuals or institutions and has been a common tool to estimate causal relationships in applied research (Clower, 2019). DiD is a version of fixed effects estimates which are variables that are constant across institutions or individuals (Angrist & Pischke, 2009). The data used in the analysis of this study occurs from quasi-experiments. A quasi-experiment refers to when an exogenous event affects the environment of individuals, geographical regions or firms. Exogenous event is used as a term when the intervention is out of the hands of the firms. A quasi-experiment includes dividing firms into two groups. The control group are not affected by the policy intervention and works as the baseline group, while the treatment group are affected by the policy change (Wooldridge, 2009, 453). The intervention in this study is hospital mergers. In order to control for the systematic differences that may occur between the two groups, data before and after the intervention need to be obtained. Hence, the DiD estimation require observations of the control group before and after the policy intervention and observations of the treatment group before and after the intervention (Wooldridge, 2009). A comparison of the two groups before and after the intervention may explain whether the hospital mergers have caused differences in efficiency and quality.

The intuition behind the DiD estimation is that we are not able to observe the outcome of the treatment group if they were not treated. Hence, the control group needs to be similar to the treatment group regarding characteristics because the control group is used as a proxy for the treatment group to visualize what would have happened to the treatment group if there were no treatment (Clower, 2019). When estimating DiD one need to assume a common trend over time. In other words, we assume that the treatment and the control group will evolve in the same way through time and the intervention is the only thing that affects the results. For that reason, the difference of the differences between the control group and the treatment group would be zero if there were no intervention (Abadie, 2005).

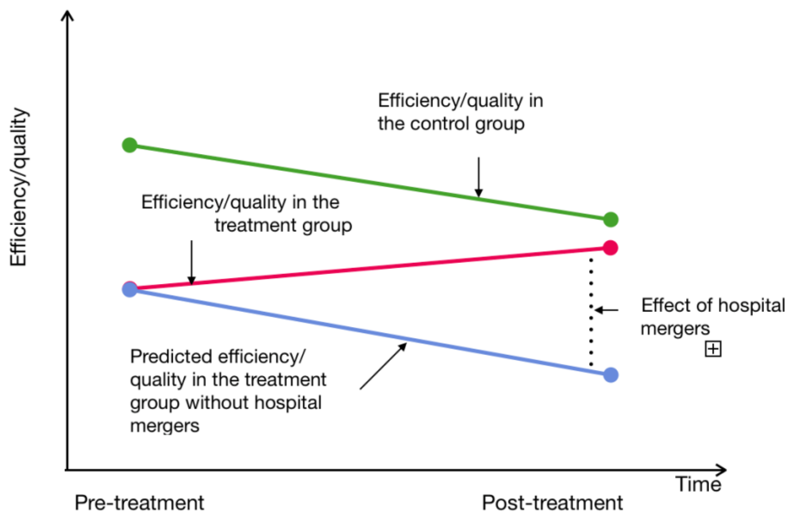


Figure 5.6 Difference-in-Difference estimation and intervention effect.

The error term (ε_i) are affecting the treatment and control group in the same way. ε_i changes over time and affects the dependent variable. DiD estimation allow us to control for factors that otherwise would have caused endogeneity problems in the model. This means that there is a correlation between the explanatory variable and the error term in the model (Lindlbauer *et al.*, 2015, 1034). In this analysis DiD estimates are used to identify the efficiency/quality changes of non-merged hospitals relative to the efficiency/quality changes of merged hospitals.

This study will determine whether merged hospitals have better efficiency and quality than non-merged hospitals. The treatment group represents hospitals that have merged the previous decade while the control group represents hospitals that are non-merged or not yet merged. The coefficient of interest is the interaction term, which represents the changes in productivity/quality in the treatment group relative to the changes in productivity/quality in the control group. The DiD estimator is achieved by creating an interaction between the time the intervention occurred and the merged hospital. In this model, the interaction term will be denoted *"DID"*. The coefficient of "*DID"* will determine whether the merged hospital preform with greater productivity/quality than non-merged hospitals, and whether the effect is statistically significant or not. The DiD-estimation for productivity will be estimated simultaneously within the Simar & Wilson regression. Hence, "DID" will be included as an explanatory variable. Let C denote the control group and T denote the treatment group, letting dT represent equal unity for group T, which are zero if group C. Then, let d2 represent the time dummy variable that are equal one for the post-treatment period and zero otherwise. The equation of interest is then.

$$y_{it} = \beta_0 + \delta_o d2 + \beta_1 dT + \delta_1 d2 * dT + a_i + u_{it} \quad (5.13)$$

where, y_{it} is the outcome variable of interest for institution i at time t, β_0 represents the intersection for t=1, and δ_1 captures the effect of the intervention. The component a_i captures the unobserved factors that are specific to institution i over time and can affect y_{it}. According to Wooldridge (2009, 453) is the component, a_i, often referred to as the fixed effect or unobserved heterogeneity in the model. Moreover, u_{it} represents the idiosyncratic error which are unobserved factors that vary over time and across institutions and affects the outcome variable of interest (Wooldridge, 2009, 456). Note that only the parameters y_{it} and dT are observable (Abadie, 2005). Equation 5.13 represents the DiD-theory when all units in the treatment group are treated at the same time. However, this is not the case for Norwegian hospital mergers. There are a few adjustments that have to be made in order to get a more accurate and unbiased result. The units that are not yet treated are used as a control group for the units that are already treated. Hence, the variable dT are not of particular interest. Moreover, the time of treatment cannot be defined for the control group, and the part of equation 5.13 that includes d2 is removed and dT*d2 are re-interpreted as the actual treatment time for each unit. The DiD-equation used in this study is;

$$y_{it} = \beta_0 + \beta_1 dT + \delta_1 d2 * dT + a_i + u_{it} \quad (5.14)$$

In order to get an approximate percentage effect, log linear regression of equation 5.14 can be computed where 100*$\hat{\delta}_{DID}$ is the percent increase (or reduction) of efficiency due to hospital mergers.

There are two possible ways to estimate the interaction term "DID" (Wooldridge, 2009, 454). First method is to calculate the differences in averages between the treatment- and the control group in both period 1 and 0, and then difference the result. This can be expressed as;

$$\hat{\delta}_{DID} = (\bar{y}_1^T - \bar{y}_1^C) - (\bar{y}_0^T - \bar{y}_0^C) \quad (5.15)$$

where the superscripts T and C represent treatment group and control group, and the subscripts 1 and 0 represent post-treatment and pre-treatment, respectively. The bar denotes the average. The second method is to calculate the change of the averages between period 0 and 1, and then the changes. This can be expressed as;

$$\hat{\delta}_{DID} = (\bar{y}_1^T - \bar{y}_0^T) - (\bar{y}_1^C - \bar{y}_0^C) \quad (5.16)$$

Where the first part of the equation refers to the before-after difference for the treatment groups, while the latter refers to the before-after for the control group. $\hat{\delta}_{DID}$ represents the DiD-coefficient in the analysis. The second method of estimating DiD will be used in this study. The estimate $\hat{\delta}_{DID}$ will not be depend on which method to differentiate, the result will be the same regardless of using either 5.15 or 5.16 (Wooldridge, 2009, 454).

Table. 5.3 Illustration of the DiD estimator.

	Before (0)	After (1)	After – Before
Treatment group (A)	$\beta_0 + \beta_1$	$\beta_0 + \delta_0 + \beta_1 + \delta_1$	$\delta_0 + \delta_1$
Control group (B)	β_0	$\beta_0 + \delta_0$	δ_0
Treatment – Control	β_1	$\beta_1 + \delta_1$	δ_1(DID)

(*Woodlridge, 2009, 454*).

Table 5.3 shows how the coefficient $\delta_1 = (\delta_0 + \delta_1) - \delta_0$ is determined, which represents the interaction term in the analysis. Bertrand *et al.* (2004) present in the article *"How much should we trust difference-in-differences estimates?*" three possible reasons for why estimations of equation 5.13 could lead to correlation problems. First, DiD estimation often relies on long time series. Second, the typically used dependent variable is commonly correlated. Finally, the treatment variable rarely changes within an institution over time (Betrand *et al.*, 2004).

Table 5.4 Overview of control- and treatment groups used in the DiD-analysis.

Control (non-merged hospitals)	Haraldsplass Diakonale Sykehus
	Lovisenberg Diakonale Sykehus
	Diakonhjemmet Sykehus
	Stavanger Universitetssykehus HF
Treatment (merged hospitals)	Helgeland Hospital HF
	Nordland HF
	UNN HF
	Helse Finnmark HF
	Helse Nord-Trøndelag HF
	Helse Møre og Romsdal HF
	Helse Førde HF
	Helse Bergen HF
	Helse Fonna HF
	Vestfold Hospital HF
	St. Olavs Hospital HF
	Sørlandet Hospital HF
	Telemark Hospital HF
	Vestre Viken HF
	Østfold Hospital HF
	Innlandet Hospital HF
	AHUS HF
	OUS HF

Table 5.4 above gives an overview over which hospitals that are considered as being in the control group and the treatment group in the DiD-analysis. From 1999 until 2014, only 4 out of 22 hospitals included in this study were not merged, either physically or juridical. This corresponds to the findings by Kjekshus *et al.*, (2014) that 90% of all public hospitals were affected by hospital merging from 2000 until 2010. Stavanger Universitetssykehus HF is the only health enterprise that did not merge together with other Norwegian hospitals.

6 RESULTS

This chapter will include the results of the case-mix adjustments of the quality indicators. Moreover, the results from the DEA-SW analysis with DiD estimation of productivity will be presented. In addition, the results from the DiD estimation on quality will be presented and discussed. Finally, critical reflection of the results will be provided. It is important to note that the data used in this thesis cannot be used to trace individuals.

6.1 Case-mix adjustments of the hospitals

The performance indicators for hospitals in all three models are calculated by dividing the observed by the predicted values through a logit estimation where lower values are associated with better quality. Even though all three models are calculated, only the results from model 2 will be presented in the following section and used in the DEA-SW analysis. The mean over hospital of the performance indicators for the three models are stated in table 6.1 below. Higher values of the performance indicator are associated with lower quality since the observed outcome is higher than the predicted outcome.

Table 6.1 Mean values over hospitals, and standard deviation of performance indicators for model 0,1 and 2, using 1999-2014 patient-level data.

	Model 0		**Model 1**		**Model 2**	
	Mean	Std.Dev	Mean	Std.Dev	Mean	Std.Dev
Readm_30_emgc	0.9895	0.2244	0.9696	0.2146	0.9798	0.2182
Inhosp_mortality_1	1.1374	0.3544	1.0828	0.3232	1.1052	0.3546
Psi12vt_pe	0.9872	0.4681	0.9329	0.4283	0.9993	0.4713
Psi13sepsis	0.9244	0.3550	0.9024	0.3373	1.0293	0.3835
Psi15ac_punc	0.8563	0.7801	0.7813	0.6744	0.7953	0.6033
Psi18ob_trau	0.9786	0.7922	0.8666	0.7274	0.9027	0.7310
Bed_sores	0.9832	0.6925	0.8668	0.5890	0.9331	0.5822

The table is based on outputs retrieved from STATA 16.

The performance indicator outcomes of the quality indicators and patient safety indicators for the 22 Norwegian hospitals are listed in table 6.2 below. Sunnmøre HF and Nordmøre HF merged in 2011 and changed name to Møre og Romsdal HF. However, numbers from all three health enterprises are presented in the results below.

Table 6.2 Performance indicators of the two quality indicators and seven patient safety indicators for model 2.

	AHUS HF	**Diakonhjem.**	**Finnmark HF**	**Haraldsplass**	**Helgeland HF**	**H. Bergen HF**
Perf_u2_readm30_emgc	1.020	0.995	0.915	1.004	1.148	1.029
Perf_u2_inhosp_mortality_1	0.962	1.090	0.940	0.992	1.024	0.809
Perf_u2_psi12vt_pe	1.070	1.165	0.795	1.148	0.734	0.846
Perf_u2_psi13sepsis	0.959	1.122	1.347	1.333	0.692	1.083
Perf_u2_psi15ac_punc	1.498	0.804	0.578	0.575	0.697	0.814
Perf_u2_psi18ob_trau	2.130		0.949		1.985	1.249
Perf_u2_bed_sores	0.733	1.317	0.997	0.913	0.843	0.698
	H. Fonna HF	**H. Førde HF**	**Lovisenberg**	**Møre&Roms. HF**	**Nordland. HF**	**Nordmøre HF**
Perf_u2_readm30_emgc	1.077	0.935	0.851	0.757	0.999	1.087
Perf_u2_inhosp_mortality_1	0.893	0.887	1.204	0.756	0.858	0.851
Perf_u2_psi12vt_pe	0.931	0.782	1.336	0.892	0.773	0.695
Perf_u2_psi13sepsis	1.112	1.046	0.922	0.853	0.821	0.543
Perf_u2_psi15ac_punc	0.816	1.163	0.379	0.790	1.033	0.620
Perf_u2_psi18ob_trau	1.073	1.733		1.557	1.386	0.509
Perf_u2_bed_sores	0.956	1.312	1.181	1.186	0.654	0.776
	Nord-Tr. HF	**OUS HF**	**St. Olavs HF**	**Stavanger HF**	**Sunnmøre HF**	**Innlandet HF**
Perf_u2_readm30_emgc	1.051	1.000	1.004	1.128	1.004	1.019
Perf_u2_inhosp_mortality_1	0.900	0.668	0.925	0.945	0.823	0.884
Perf_u2_psi12vt_pe	0.607	1.050	1.027	0.886	0.802	0.893
Perf_u2_psi13sepsis	0.893	0.869	1.185	0.753	0.549	0.833
Perf_u2_psi15ac_punc	0.929	1.109	0.927	0.963	0.880	1.051
Perf_u2_psi18ob_trau	0.899	0.809	1.140	1.103	1.139	1.171
Perf_u2_bed_sores	0.605	0.522	0.881	0.638	1.760	0.793
	Telemark HF	**Sy. Vestfold HF**	**Sy. Østfold HF**	**Sørlandet sy. HF**	**UNN Tromsø HF**	**Vest.Viken HF**
Perf_u2_readm30_emgc	1.189	1.087	1.164	1.094	0.936	1.015
Perf_u2_inhosp_mortality_1	0.877	0.794	1.068	0.887	0.860	0.819
Perf_u2_psi12vt_pe	0.857	0.886	0.930	0.796	1.006	1.230
Perf_u2_psi13sepsis	0.956	0.918	1.047	0.940	1.178	1.113
Perf_u2_psi15ac_punc	0.760	0.666	0.661	0.890	0.882	1.185
Perf_u2_psi18ob_trau	1.193	0.395	1.609	1.362	1.183	1.206
Perf_u2_bed_sores	0.640	0.511	0.620	1.931	1.186	1.527

The table is made based on outputs retrieved from STATA 16.

The numbers in table 6.2 are mean values for each of the HFs where lower numbers imply higher quality. Note that Diakonhjemmet sykehus, Haraldsplass Diakonale sykehus and Lovisenberg Diakonale sykehus have missing values for obstetric trauma. This may be due to lack of maternity wards at their hospitals. According to the output above, Telemark HF has the poorest quality when it comes to emergency readmission after 30 days of discharge, while Helse Møre og Romsdal HF has the best. Oslo University hospital HF has the best quality when it comes to in-hospital mortality, while Lovisenberg Diakonale sykehus has the lowest. This may be due to their majority of elderly patients as presented in attachment E in the appendix.

6.2 DEA-SW

Due to missing information in some hospitals in the aggregated dataset, 502 out of 506 hospitals are included in the DEA-SW analysis. There were estimated four different analyses on productivity of the hospitals where the two first were based on algorithm 1 and the latter two on algorithm 2. Moreover, year dummies and year trend were included separately under both algorithm 1 and 2. The results presented in table 6.3 below rely on the output from algorithm 2 with year dummies as explanatory variable.

Table 6.3 Simar & Wilson regression on estimated Farrell input-oriented productivity scores using CRS

Variable	Coefficient	Bootstrap Standard Error	P-value	95 % Confidence interval	
DEA_tebc_dummy					
Average_traveltime	-.067*	.008	0.000	-.083	-.050
ALOS	.051*	.015	0.000	.021	.081
treated	.010	.016	0.553	-.022	.041
DID	.002	.012	0.841	-.021	.027
perf_u2_readm30_emgc	.081*	.018	0.000	.045	.117
perf_u2_inhosp_mort~1	-.012	.017	0.477	-.044	0.022
perf_u2_psi12vt_pe	-.008	.008	0.330	-.025	.008
perf_u2_psi13sepsis	-.035*	.011	0.001	-.056	-0.013
perf_u2_psi15ac_punc	-.0006	.006	0.928	-.013	.012
perf_u2_psi18ob_trau	-.006	.007	0.358	-.019	0.007
perf_u2_bed_sores	.010	.007	0.138	-.003	.023
2000.year	.040*	.016	0.012	.008	0.071
2001.year	.009	.016	0.561	-.023	.041
2002.year	.037	.021	0.076	-.004	.079
2003.year	.053*	.022	0.015	.011	.096
2004.year	.072*	.022	0.001	.029	.116
2005.year	.070*	.023	0.003	.024	.115
2006.year	.073*	.244	0.003	.025	.121
2007.year	.073*	.025	0.004	.024	.123
2008.year	.125*	.025	0.000	.076	.174
2009.year	.121*	.027	0.000	.067	.175
2010.year	.118*	.028	0.000	.062	.173
2011.year	.162*	.029	0.000	.104	.219
2012.year	.127*	.030	0.000	.068	.185
2013.year	.148*	.030	0.000	.088	.206
2014.year	.149*	.031	0.000	.090	.210
Constant	.615*	.055	0.000	.508	.722
/Sigma	.081*	.003	0.000	.073	.084

Significant coefficients at a 5% significance level are marked with *.

Using year dummies will yield a more robust result compared to the year trend. However, including year trend, which is only one variable, may result in more degrees of freedom. The results are estimated with a 5% significance level. The bootstrapped output presented above gives 502 inefficient and zero efficient DMUs. This is a result of also having bootstrapped the DEA-model itself with 10,000 bootstrap replications. The output indicates that travel time from the patients' municipality to the given hospital by car has a negative impact on productivity and is significant with a p-value less than 0.050. In other words, longer travel time for the patient gives a reduction of productivity at the hospitals. Average length of stay for the patients has a positive and significant effect indicating that increasing ALOS yields better productivity for the hospitals in this study. Moreover, the DID coefficient is positive but not significant at a 5% significance level, signifying that hospital mergers do not affect the productivity at the hospitals. Furthermore, emergency readmission rate after 30 days of discharge are positive and significant which implies that readmission have a positive effect on productivity. On the other hand, sepsis is negative and significant at a 5% significance level, which indicates that sepsis has a negative effect on productivity. The year dummies included in the analysis are mainly positive and significant. This signifies that productivity scores are gradually increasing over time. The output from algorithm 2 with year trend show that the productivity increases with 1.01 % each year, and is significant at a 5% significance level. Hence, there is a positive progress on productivity even when not distinguishing between years.

Figure 6.1 Boxplot of productivity score for merged hospitals per year

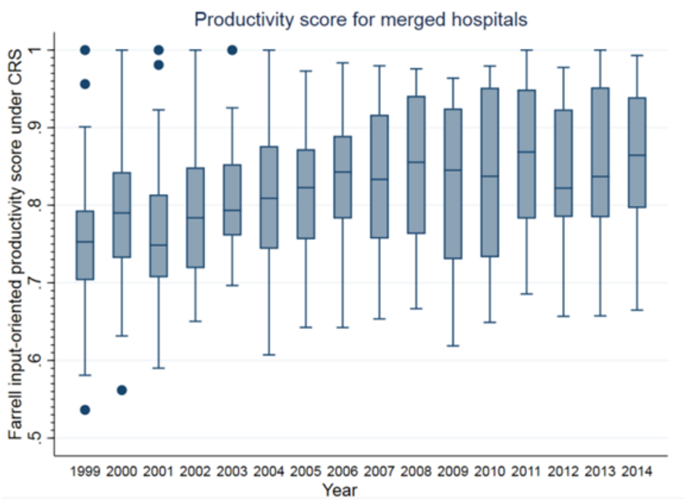


Figure 6.2 Boxplot of productivity score for non-merged hospitals per year

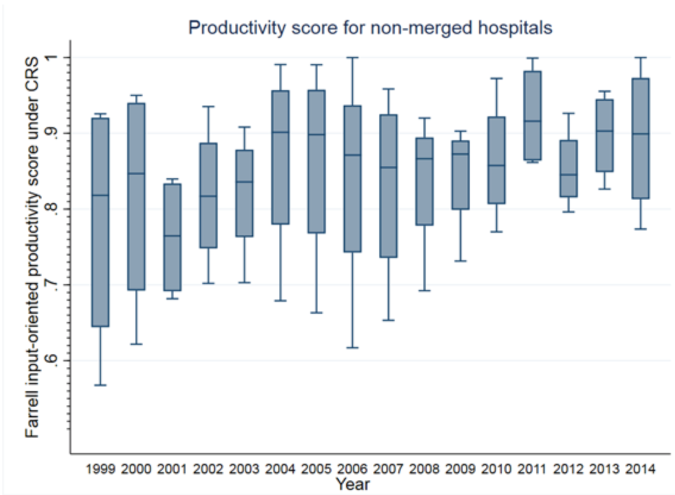


Figure 6.1 and 6.2 above illustrates the productivity scores obtained under CRS in the DEA-SW analysis for merged and non-merged hospitals, respectively. There are a few outliers from 1999 until 2003 for the merged hospitals, which represents extreme values that are deviating substantially from the rest of the observations in these years. The interquartile range (IQR) is narrower for the merged hospitals from 1999 until 2008 implying that the estimates are more precise compared to the ones of the non-merged hospitals. The horizontal line within each box represents the median productivity score. The merged hospitals in figure 6.1 have a more gradually increase in productivity score compared to the non-merged hospitals in figure 6.2.

6.3 Difference-in-Differences estimation of quality

Hospital mergers are treated as an exogenous shock in the DiD analysis in order to examine the effect of hospital mergers on quality. This section will examine whether Norwegian hospital mergers have affected the quality of the hospitals with different dependent variables. The explanatory variable "DID" represent the Difference-in-Difference estimation (DiD) of quality. "ALOS" is the average length of stay for the patients, while "patients" works as an indicator of hospital size. Furthermore, "average_traveltime" represents the travel time in hours from the patient's municipality to the given hospital. Finally, the year dummies are catching up the development of quality over time. In all five tables presented in this section, quality is improving yearly and is not included in the tables in this chapter. The year dummies will be commented for each dependent variable and are fully presented in attachment G in the appendix. Note that lower values in the following tables are associated with better quality.

Table 6.4 Quality output for the sum of all seven performance indexes

Explanatory Variable	**Coefficient**	**Standard Error**	**P-value**	**95% Conf. Interval**	
DID	.110*	.026	0.000	.059	.162
ALOS	-.054	.032	0.088	-.117	.008
patients	-1.40	7.92	0.078	-2.95	1.57
average_traveltime	-.071*	.021	0.001	-.112	- .031
Constant	*1.167**	*.077*	*0.000*	*-.223*	*.005*

Significant coefficients at a 5% significance level are marked with *.

Table 6.4 above includes the performance indicators obtained in model 2 for the two quality indicators and the five patient safety indicators. The coefficient "DID" is both positive and significant at a 5% significance level, which indicates that the merged hospitals have lower quality compared to the non-merged hospitals in this study. Increased ALOS are associated with better quality. However, this study finds a positive effect of ALOS on quality, but the effect is not significant. Hence, increasing ALOS does not have an effect on quality. Moreover, bigger hospitals are related to better quality with a negative coefficient. Nevertheless, the effect is not significant. Furthermore, the results imply that the quality is being positively affected by increased travel time, and is significant at a 5% significance level. The year dummies for the seven performance indexes show a positive trend on quality compared to the quality in 1999. However, the year dummies are not significant after 2008.

Table 6.5 Quality output for the sum of the two main performance indexes

Explanatory Variable	Coefficient	Standard Error	P-value	95% Conf. Interval	
DID	.085*	.021	0.000	.045	.126
ALOS	-.129*	.025	0.000	-.179	-.080
patients	-2.75*	6.24	0.000	-3.98	-1.15
average_traveltime	-0.130*	.016	0.000	-.162	-.098
Constant	*1.615**	*.061*	*0.000*	*1.495*	*1.735*

Significant coefficients at a 5% significance level are marked with *.

Table 6.5 above includes the performance indicators obtained in model 2 for the sum of the two quality indicators readm30_emgc and inhosp_mortality. The results concerning "DID" and "average_traveltime" are more or less the same as for all seven performance indexes, and merged hospitals are performing with lower quality than non-merged hospitals in this study. However, "ALOS" and "patients" are still negative but are now significant at a 5% significance level. This implies that longer average length of stay gives better quality for the hospitals. Furthermore, bigger hospitals have better quality and are now significant. The year dummies in the output indicate that the quality is gradually increasing every year compared to 1999.

Table 6.6 Quality output for emergency readmission within 30 days of discharge

Explanatory Variable	Coefficient	Standard Error	P-value	95% Conf. Interval	
DID	.139*	.026	0.000	.088	.190
ALOS	-.092*	.032	0.004	-.154	-.030
patients	-9.25	7.85	0.239	-2.47	6.18
average_traveltime	-.084*	.020	0.000	-.124	-.044
Constant	*1.204**	*.077*	*0.000*	*1.053*	1.355

Significant coefficients at a 5% significance level are marked with *.

Table 6.6 above only includes the quality indicator for emergency readmission within 30 days of discharge. The "DID" coefficient is still positive and significant at a 5% significance level, which implies that merged hospitals give higher readmission rates. The average length of stay is negative and significant as in table 6.3.2. Increased average travel time is significant and does not result in higher readmission rates. The year dummies indicate that the readmission rates are decreasing over time expect from in 2007. However, they are not significant other then in 2002 and 2007, which implies a stable development over time.

Table 6.7 Quality output for in-hospital mortality

Explanatory Variable	Coefficient	Standard Error	P-value	95% Conf. Interval	
DID	.032	.029	0.272	-.023	.090
ALOS	-.167*	.036	0.000	-.237	-.097
patients	-4.58*	8.87	0.000	-6.33	-2.84
average_traveltime	-.177*	.023	0.000	-.222	-.132
Constant	*2.026**	*.087*	*0.000*	*1.856*	2.197

Significant coefficients at a 5% significance level are marked with *.

Table 6.7 above only includes the quality indicator for in-hospital mortality. The coefficient "DID" are still positive but are no longer significant at a 5% significance level. This implies that hospital mergers do not have an effect on in-hospital mortality in this study. However, average length of stay, the size of the hospital, and average travel time are negative and significant and reduce in-hospital mortality. The year dummies indicate a strong reduction over time compared to 1999, where all dummies are significant.

Table 6.8 Quality output for average travel time

Explanatory Variable	Coefficient	Standard Error	P-value	95% Conf. Interval	
DID	.284*	.057	0.000	.172	.396
ALOS	.018	.071	0.796	-.121	.158
patients	-.165	1.76	0.108	-6.30	6.21
Constant	*.503**	*.171*	*0.003*	*.167*	.839

Significant coefficients at a 5% significance level are marked with *.

Table 6.8 above only includes average travel time. Hence, "average_traveltime" is removed as explanatory variable in the equation. "DID" are still positive with a significant p-value. This implies that the travel time from the patient's municipality to the given hospital is increasing for the merged hospitals. Moreover, "ALOS" is also positive but not significant. Hence, average length of stay does not affect travel time. On the other hand, average travel time is shorter for larger hospitals with a negative coefficient, but the p-value is not significant. The year dummies indicate an increase in travel time after 2006, but these are not significant at a 5% significance level.

Figure 6.3 Mean main performance score for merged hospitals

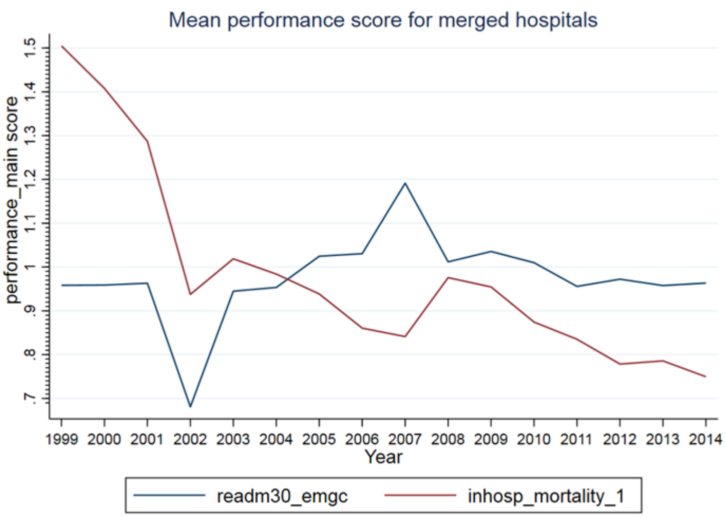
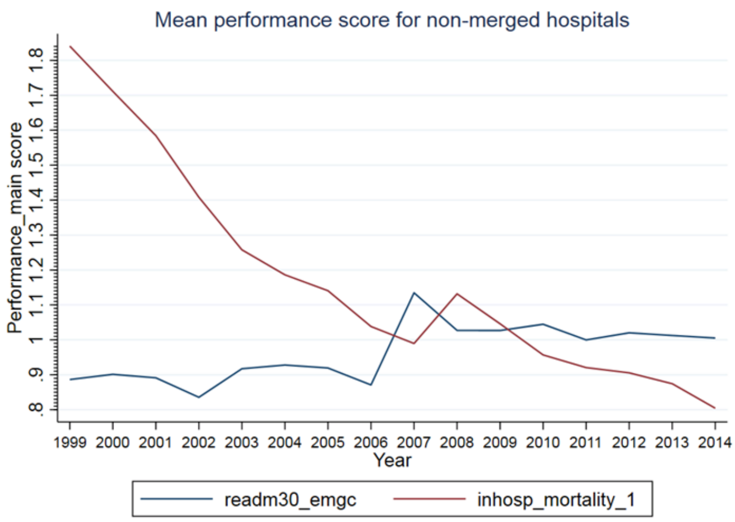
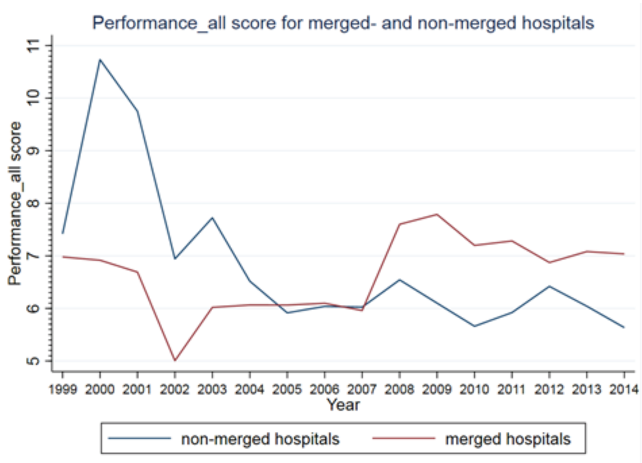


Figure 6.4 Mean main performance score for non-merged hospitals



In-hospital mortality has gradually decreased over time for both merged and non-merged hospitals, which are beneficial. Mortality is unavoidable, but an improvement is always desirable. Emergency readmissions are stable over time. However, there is a peak between 2007 and 2008 for both in-hospital mortality and emergency readmission rates in both figures. This may be a result of a change in outpatient registrations in 2008. Overall, non-merged hospitals are performing with better quality after 2005 when considering all seven performance indicators in figure 6.5 below.

Figure 6.5 Mean performance_all score for merged- and non-merged hospitals



6.4 Critical reflection

Sections 6.2 and 6.3 have presented the results from the DEA-SW analysis on productivity and the DiD estimation on quality, respectively. The DEA-SW analysis indicate that hospital mergers do not have an effect on productivity. Increasing efficiency of Norwegian hospitals was one of the main goals of the Norwegian Hospital Reform of 2002. This study has not found significant evidence in favor of increased efficiency or productivity of the merged hospitals. However, the treatment- and control groups in this study have a significantly different numbers of units within each group. Moreover, previous empirical literatures are not consistent when it comes to the effect of hospital mergers on efficiency and productivity.

There are different results of the impact of hospital merges on quality in the five tables presented in section 6.3. One of the main reasons for merging hospitals is that it may result in increased quality of the health services. It is not certain that the increase in quality will affect every patient. However, the patients that, in the starting point, should have died of their illnesses may not die due to the increase in quality. Moreover, if there are long travel distances for the patients to the hospital, the patients may refuse to be discharged to early if they have high risks. If there are a great strain on the patient by sending them home after treatment, it may be difficult to transport the patient back to the hospital on short notice. Hence, the hospital might want to keep the patient at the hospital longer than necessary. This may result in an increase in in-hospital mortality. Nevertheless, in table 6.6 that have emergency readmission rates as the dependent variable, average travel time is negative and significant, which implies that the hospitals do not keep patients at the institution because it may be difficult to transport the patient back to the hospital on short notice. Furthermore, the travel time is shorter for larger hospitals. This may be due to the fact that bigger are located in larger cities hospitals in general. Additionally, average travel time increases for the merged hospitals. However, this is not unexpected, as smaller hospitals have been closed due to the mergers. Average travel time has increased over time. During the 1990s, Norwegian policy makers introduced the freedom to choose their preferred health provider. As a result, more patients may travel longer to receive treatment at hospitals that are specialized on certain treatments or that are superior on complications under treatment.

7 DISCUSSION AND CONCLUSION

This chapter will focus on the study's general relevance and an interpretation of the results obtained in chapter 6 in relation to the research question presented in section 1.3. Limitations of the study will also be discussed together with the validation of the study in the light of the limitations. Moreover, policy recommendations and "what to be learnt" from this study will be presented. Finally, ethical considerations of the study and concluding remarks will be discussed and presented.

7.1 Interpretation of results concerning the research question

This analysis has tried to determine if the hospital mergers that have occurred between 1999 and 2014 have resulted in increased productivity and quality at Norwegian hospitals. The DEA-SW analysis indicates that hospital mergers do not have an effect on productivity since the DID coefficient is not significant. Moreover, average travel time and ALOS are associated and have an impact on the productivity of the hospitals, but the result itself cannot provide information regarding the causation of these two variables. Furthermore, there is a yearly trend in the results where there is a progress in productivity, and where the productivity increases with approximately 1.01% each year. Emergency readmission after 30 days of discharge was positive and significant. Having a larger number of readmissions is associated with lower quality, which in this study has resulted in higher productivity. This is a trade-off since poorer quality yields lower costs, and better quality results in higher costs. On the contrary, sepsis has a negative and significant effect on productivity. Hence, poorer quality drives costs. In all of the five quality outputs presented in section 6.3, the DID coefficient is positive, which implies that non-merged hospitals are performing with better quality than merged hospitals. Hence, hospital merges have not contributed to increase quality of the hospitals in this study. The empirical literatures are not consistent concerning the effect of hospital mergers on efficiency, productivity, and quality.

This study corresponds to the findings obtained in the study performed by L. Kjekshus and T. Hagen (2007). They found that there was no significant effect of hospital mergers on technical efficiency. Moreover, the study performed by Kittelsen *et al.* (2017) found that Norwegian hospitals are considered to have higher quality than the other Nordic countries since the mortality rates are lower. This study however, does not compare Norwegian hospitals to other

Nordic countries, but the findings show a strong reduction in in-hospital mortality over time for both merged and non-merged hospitals, while emergency readmissions have stayed fairly stable over time. Even though there are mixed and limited empirical evidence concerning the effect of hospital mergers on quality, hospital mergers seem to reduce the quality at the hospitals (Varkevisser & Schut, 2009).

7.2 Limitations of the study

This study does have limitations when it comes to the measurement of mortality. Since the data obtained from 1999 to 2007 only contains information regarding whether the patient was discharged as living or dead, in-hospital mortality was the only possible measurement of mortality. Individuals who have passed away due to complications that have occurred after discharge are not captured in the analysis. Nevertheless, if the patient is being readmitted to the hospital and then died at the hospital the individuals are captured in the data. However, the data does not contain information concerning how many individuals who have died at home. Even though the data from 2008 to 2014 are able to capture mortality after 30-, 90-, 180- and 365 days of discharge, the same procedure had to be applied to this data in order to make the two patient-level datasets comparable.

Another limitation of the study concerns the registration of patients in the dataset. In the data obtained from 1999 until 2007, the patient is registered with a new patient number every year, meaning that every patient is given a new patient number on January 1. This leads to complications of measuring emergency readmissions after 30 days of discharge since patents who are discharged in December are given a new patient number in January. For that reason, patient who are discharged in December have been excluded from the analysis. However, the data obtained from 2008 until 2014 does not suffer from this inconvenience, but the same procedure has to be applied to this dataset as well.

Furthermore, the patient-level datasets include information of patient who have been hospitalized at any health institution or hospital in Norway from 1999 until 2014. However, the hospital data from the same period does not include all these hospitals and health institution. Hence, a substantial number of health institutions had to be omitted in order to merge the three datasets together resulting in a reduction of approximately 6,000,000 patients in the patient-level data.

Moreover, it was not feasible to estimate travel time for all patients in the datasets. From 1999 until 2014, there have been multiple changes in municipality names and regions were municipalities have merged. Not all these changes are included in the estimation of travel time resulting in a lack of travel time estimation for some of the patients in the analysis. Furthermore, if may be an underlying reason for the selection into mergers. I may be that the hospitals that was expected to perform badly are the ones who have merged into bigger hospitals to take advantage of possible scale effects. Finally, there is a dissimilar distribution of treatment- and control groups in the DiD analysis. Having two groups that are very different in size could make it difficult to detect the treatment effect and to obtain significant values in the analysis.

7.2.1 Validation of the study in light of the limitations

In-hospital mortality capture sudden changes in the patient state of health. However, the long-term effect of the treatment is not captured. This may influence the result of the analysis, especially if certain hospitals have a substantially higher number of, for instance, mortality rates after 30 days of discharge. Whether there is a correlation between in-hospital mortality and after-discharge mortality is not within this thesis frames and are hard to elaborate on due to other potential underlying factors that may influence the result. When that said, in-hospital mortality is a quality indicator that can be used as a measure of the hospital's quality of treatment even though this kind of measure is much weaker compared to the additional after-discharge mortality rates.

Furthermore, not using all twelve months concerning emergency readmission rates after 30 days of discharge should not have a substantial effect on the result. An assumption that December is somewhat similar to the previous eleven months is applied. There is no reason to assume that December is significantly different from the eleven former months when it comes to readmission rates. Moreover, the omitted hospitals from the patient-level data mainly includes health institutions that do not provide somatic health services but are highly specialized health institutions. Institutions that provide rehabilitation services or private hospitals are not included in the analysis. However, most Norwegian hospitals that have merged into health enterprises are included in the data.

Not having obtained travel time for all patients in the patient-level dataset should not significantly affect the mean travel time for the hospitals. Most of the patients in the analysis are listed with an estimated travel time from their municipality to the given hospital. However, including travel time estimation for all patients could affect the mean travel time for each of the hospitals in the analysis, but here I assume that the travel time would not significantly affect the outcome by excluding some of the patients travel time.

Finally, the inequitable distribution of treatment- and control group in the DiD-estimation could have influenced the results of both the productivity- and quality estimation. However, there is no rule that states what the ratio between the treatment- and control group should be. Nevertheless, having a more equitable distribution of the groups may be beneficial, especially when the distribution of the two groups are very different from each other.

7.3 Policy recommendations

Even though there are only Norwegian hospitals included in this analysis, the results of the study may be interesting for other countries if they have similar organization structures of their health care sector as the Norwegian system. This study has tried to determine whether hospital mergers have resulted in increased productivity and quality at Norwegian hospitals. When hospitals are too big, the expected economies of scale may result in scale disadvantages. However, larger hospital may also attract patients that are more demanding and the efficiency of these hospitals may be underestimated. As mentioned in section 2.2, hospital mergers are driven be a desire to reduce duplication of hospital functions since this could result in a reduction of costs. Nevertheless, hospital leaders cannot base their decision only on cost savings but has to take the quality of the health services, availability, and patient safety into consideration. If hospital mergers should be an effective and useful instrument, there has to be some sort of volume effect for the merged hospital. If this happens through closure of hospitals or departments - or centralizing certain treatment to one or more hospitals is out of this thesis score. Increasing efficiency and quality is an overall goal for the Ministry of Health in Norway. However, the patients', dependents' and, inhabitants' needs and wishes has to be taken into account when policy makers consider changing the health care sector – especially when the changes are affecting the quality of care and, for instance, travel time. There are a few policy recommendations that follow form the results presented in chapter 6.

The first recommendation concerns the implementation of the reorganizations. Based on the results provided in this study, hospital mergers have no significant effect on productivity and negative effect on quality. By examining the effect of hospital mergers on a small sample of hospitals before reorganizing the entire health care sector could give an idea of what the total effect of hospital mergers could be. Moreover, since it takes time to discover the true effect of hospital mergers, patience is required for the policy makers and the society as a whole. When hospitals are merging there may be that some hospitals are performing better while some are performing worse and the net effect are closer to zero. However, the hospitals may also experience more specialization as an alteration to mergers. Furthermore, when the organizational structure of the hospital are chaining, the incentives to provide treatment may change. Larger hospital that covers a greater share of the population may not have the same incentives to provide the same quality of treatment when there are fewer hospitals to choose from for the patients than smaller hospitals. Smaller hospitals may focus more on the quality of treatment to attract new patients. Hence, it may be beneficial to change the funding system by giving better incentives to provide health services above an acceptable quality level. However, there will always be downsides related to a funding system, but policy makers have to develop a funding system that have more benefits and disadvantages for most people in general.

7.3.1 What to learn?

Most hospital mergers in Norway are assumed to take place gradually over time and is a time-consuming action. When comparing hospitals, it is important to perform case-mix adjustments in order to sort out differences that are mainly a result of differences in patient composition. Patient characteristics in model 1 presented in the case-mix results in section 6.1 are shown to be important. Moreover, using datasets that together contains approximately 80 million patients is time-consuming and demanding, but indeed feasible. This study has yield somewhat unexpected results. By considering this study, one can learn that Norwegian hospital mergers have not worked as a convincing tool to achieve the goal of increased hospital efficiency. Even though the effect is positive, it is not significant. Furthermore, when taking the results of this study into consideration, the hospital mergers have not positively affected the quality of the hospitals. These findings may come in conflict with the assumption that larger hospitals are performing with better quality.

7.4 Further studies to be needed

This study has examined how hospital mergers affects productivity and quality at Norwegian hospitals. Many factors have to be taken into account when it comes to running a hospital. Productivity and quality are not the only two aspects that are important to improve. Lowering waiting times and increase the availability are two factors that are valuable for the patent. It could be interesting to see how waiting times and availability are affected by hospital mergers, especially since hospital mergers often results in covering a larger share of the population for the hospitals. Furthermore, a prerequisite of economies of scale by hospital mergers is that there actually is potential for it in the health care sector. Hence, it could be interesting to examine which hospitals who has potential for economies of scale and to see how well these hospitals perform in relation to fulfil these cost-saving actions. Moreover, if it is possible to access patient level data that makes it feasible to measure the quality indicators more accurate, it could be interesting to see if the results of this study are significantly different from a more comprehensive set of quality indicators. Finally, it could be interesting to examine how traveling distance for certain patient groups, for instance pregnant women, have been affected by the Norwegian hospital mergers and whether the patient satisfaction have decreased or increased after these mergers.

7.4.1 Ethical considerations

The hospital data used in this study is not individual data but aggregated at hospital level. The patient-level data used in the analysis includes personal information about, for instance, sickness and home-municipality of the patient. It is essential to have in mind that these data should not be saved and stored over time. Moreover, it is important to state that tracing people based on the data used in this study is not possible. The Regional Ethical Committees (REC) has approved the application of data material January 3, 2020. Work concerning the patient-level dataset have been completed under strict surveillance from Frischsenteret located at Oslo Science Park.

7.5 Concluding remarks

Increasing efficiency and quality of health services are goals that every health authority aims to accomplish and there are different tools that can be used in order to fulfill these goals. Hospital efficiency could be increased b, for instance, lowering ALOS or decrease costs. However, providing health services above an acceptable level of quality is costly, which can result in a trade-off between cost and quality. The effect of hospital mergers on productivity and quality are not in line with the assumed effect of this mechanism. In this study, hospital mergers have contributed to an increase of emergency readmission rates, where an increase in ALOS are associated with fewer emergency readmissions. Larger hospitals may have less capacity since they are covering a larger share of the population. However, increasing ALOS are assumed to increase the quality of the treatment in one way or another.

The effect of hospital mergers on productivity was positive but not significant at a 5% level. Hence, hospital mergers have not contributed to increased productivity at Norwegian hospitals. However, the effect of increased emergency readmission rates on productivity was positive, which implies that lower quality and lower costs are associated with higher productivity scores. Alternatively, higher quality and higher costs are associated with lower productivity scores. Hospital mergers is time-consuming and it takes time to establish new cultures and environments at the hospitals. Using datasets that may capture the effect of the mergers before and after the intervention is important, but it may be that the true effect of the mergers is not as significant as assumed. It will be interesting to see if the results from similar studies in the future will have different outcomes.

8 REFERENCES

Abadie, A. (2005). Semiparametric Difference-in-Differences Estimators. *The Review of Economic Studies*. 72(1). pp. 1-19.

Ahgren, B. (2008). Is it better to be big? The reconfiguration of 21st century hospitals: responses to a hospital merger in Sweden. *Health Policy*, **87** (1). pp. 92-99.

Andersen, B. M. (2012). Lite lønnsom sykehusfusjon. *Sykepleien.* (7). pp. 66-67. Published June 11, 2012. https://sykepleien.no/forskning/2012/05/lite-lonnsom-sykehusfusjon

Angrist, D. J. & Pischke, J. S. (2009). Mostly Harmless Econometrics: An Empiricist´s Companion. *Princeton University Press*. pp. 1-253.

Anthun, K. S., Kittelsen, S. A. C. & Magnussen, J. (2016) Produktivitet i spesialisthelsetjenesten. University of Oslo. Health economics research network. *Working paper 2016:7.*

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*. pp. 418-425.

Askeland, O. M. (2018). Liggetid i barsel. Ny statistikk fra Medisinsk fødselsregister. *Folkehelseinstituttet.*https://www.fhi.no/globalassets/dokumenterfiler/trykksaker/fodselsnytt-nr-1-2018.pdf

Badunenko, O. & Mozharovskyi, P. (2016). Nonpragmatic frontier analysis using Stata. *Stata Journal*. StataCorp LP, vol. 16(3), pp. 550-589.

Badunenko, O. & Tauchmann, H. (2018). Simar and Wilson two-stage efficiency analysis for Stata. *FAU Discussion Papers in Economics*. No. 08/2018. Friedrich-alexander-Universität Erlanden-Nürnberg, Institute for Economics, Erlangen. pp. 1-37.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust difference-in-differences estimates? *The Quarterly Journal of Economics.* pp. 1-23.

Breyer, J., Giacomazzi, J., Kuhmmer, R., Lima, K., Hammes, L., Ribeiro, R., Kops, N., Falavigna, M. & Wendland, E. (2019). Hospital quality indicators: a systematic review. *International Journal of Health Care Quality Assurance*. 32(1).

Christiansen, T. & Vrangbæk, K. (2018). Hospital centralization and performance in Denmark – Ten years on. *Health Policy*, **122**. pp. 321-328.

Clower, E. (2019). Introduction to Difference-in-Differences Estimation. *APTECH.* Updated April 7, 2020. https://www.aptech.com/blog/introduction-to-difference-in-differences-estimation/

Coelli, T, J., Prasado Rao, D, S., O'Donnel C, J., & Battese, G, E. (2005). An Introduction to Efficiency and Productivity Analysis. 2nd edition. *New York: Springer.* pp. 1-331.

Delgado, A. & Holland, K. (2018). Everything You Want to Know About Deep Vein Thrombosis (DVT). *Healthline*. https://www.healthline.com/health/deep-venous-thrombosis

Emons, W. (1997). Credence Goods and Fradulent Experts. *The RAND Journal of Economics*, **28**(1). pp. 107-119.

Farrell, M, J. (1957). The Measurement of Productive Efficiency. *Journal of the Roya Statistical Society*. Series A (General), Vol. **120**, No. 3. pp. 253-290.

Ferrier, G., Kerstens, K. & Vanden Eeckaut, P. (1994). Radial and Nonradial Technical Efficiency Measures on a DEA Reference Technology: A Comparison using US Banking Data. *Recherches Economiques de Louvain*. **60**. pp. 449-479.

Gok, M. S & Sezen, B. (2012) Analyzing the ambiguous relationship between efficiency, quality and patient satisfaction in healthcare services: The case of public hospital in Turkey. *Health Policy*; **111**(3): pp. 290-300.

Gutacker, N., Bojke, C., Daidone, S., Devlin, N, J., Parker, D. & Street, S. (2013). Truly inefficient or providing better quality of care? Analysing the relationship between risk-adjusted hospital costs and patients' health outcomes. *Health Economics* **22**(8). pp. 931-947.

Halsteinli, V., Kittelsen, S. A., & Magnussen, J. (2001). Scale, Efficiency and Organization in Norwegian Psychiatric Outpatient Clinics for Children. *J. Mental Health Policy Econ*. **4**. pp. 79-90.

Helsedirektoratet. (2011). Innsatsstyrt finansiering 2012. *Helsedirektoratet.* pp. 1-86.

Helsedirektoratet. (2019). DRG-systemet. Updated April 16, 2020. https://www.helsedirektoratet.no/tema/finansiering/innsatsstyrt-finansiering-og-drg-systemet/drg-systemet

Helse- og omsorgsdepartementet. (2008). En gledelig begivenhet. Om en sammenhengende svangerskaps-, fødsels- og barselomsorg. (*St. Meld. Nr. 12. (2008-2009*)). https://www.regjeringen.no/no/dokumenter/stmeld-nr-12-2008-2009-/id545600/?ch=1

Helse- og omsorgsdepartementet. (2014). Statlig finansiering av de regionale helseforetak. *Regjeringen*. Updated October 30, 2014. https://www.regjeringen.no/no/tema/helse-og-omsorg/sykehus/vurderes/statlig-finansiering-av-de-regionale-hel/id227797/

Ingebrigtsen, T. (2010) Helseøkonomiske effekter av sykehussammenslåinger. *Tidsskr Nor Legeforen nr.9, 2010;* ***130****. pp. 940-2.*

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

Ji, Y. & Lee, C. Data Envelopment Analysis in Stata. *The Stata Journal.* **10** nr.2. pp. 267-280

Johannesen, K. A., Kittelsen, S. A. C, & Hagen, T. P. (2017). Assessing physician productivity following Norwegian hospital reform: A panel and data envelopment analysis. *Social Science & Medicine.* **175**, pp. 117-126.

Kalseth, B., Anthun, K. S., Hope, Ø., Kittelsen, S. A. C & Persson, B. A. (2011). Spesialisthelsetjenesten i Norden. Sykehusstruktur, styringsstruktur og lokal arbeidsorganisering som mulig forklaring på kostnadsforskjeller mellom landene. *Rapport SINTEF Health Services Research*. pp. 1-125.

Kittelsen, S. A. C, Anthun, K. S., Goude, F., Huitfeldt, I. M., Hakkinen, U., Kruse, M., ... Ratto, H. (2015). Cost and quality at the hospital level in the Nordic countries. *Health Economics,* **24** (2), pp. 140-163.

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

Kittelsen, S. A. C, Anthun, K. S., & Huitfeldt, I. M. S. (2017). Kvalitet og produktivitet i norske sykehus. *Michael 2017*; **14**: Supplement 19, pp. 75-85.

Kjekshus, L., Bernstrøm, V. H., Dahl, E., & Lorentzen, T. (2014). The effect of hospital mergers on long-term sickness absence among hospital employees: a fixed effects multivariate regression analysis using panel data. *BMC Health Services Research 2014 14:50*. pp. 1-10.

Kjekshus, L. & Hagen, T. (2007) Do hospital mergers increase hospital efficiency? Evidence from a National Health Service country. *Journal of health services research & policy* **12**(4) pp. 230-235.

Kristensen, T., Olsen, K, R., Kilsmark, J., Lauridsen, J, T. & Pedersen, K, M. (2012). Economics of scale and scope in the Danish hospital sector prior to radical restructuring plans. *Health policy,* **106**. pp. 120-126.

Kristiansen, I, S. & Pedersen, K, M. (2000). Health care systems in the Nordic countries – more similarities than differences? *Tidskrift for Den Norske Lægeforening: tidsskrift fir praktisk medicin, ny række.* **120**. pp. 2023-9.

Kvæl, L. (2006) Beregning av kostnadsvekter til den norske versjonen av DRG-systemet 2005. *SINTEF Helse.*

Legeforeningen. (2014). Sykehus for fremtiden. Grunnlag for en nasjonal sykehusplan. *DNLF*. pp. 1-29. https://www.legeforeningen.no/contentassets/0aa80056d9d9480bb6a9c19af30b3254/sak-7-helsepolitisk-debatt-sykehus-for-fremtiden.pdf

Lindlbauer, I., Schreyogg, J. & Winter, V. (2015). Changes in technical efficiency after quality management certification: A DEA approach. Using difference-in-difference estimation with genetic matching in the hospital industry. *European Journal of Operational Research,* **250**. pp. 1026-1036.

Magnussen, J. (2012). Gir sykehussammenslåinger økt effektivitet? *Tidsskriftet. Den norske legeforening nr.7*, 2012; **132** pp. 783.

Magnussen, J., Hagen, T, P. & Kaarbøe, O, M. (2006). Centralized or Decentralized? A Case Study of Norwegian Hospital Reform. *Social science & medicine*; **64**:2129-37

Mayo Clinic Staff. (2018). Sepsis. *Mayo Clinic.* Published November 16, 2018. https://www.mayoclinic.org/diseases-conditions/sepsis/symptoms-causes/syc-20351214

McCombie, J., Pugno, M. & Soro, B. (2002). Productivity Growth and Economic Performance: Essays on Verdoorn's Law. *Palgrave Macmillan.* New York, N. Y: 294-320

McKay, N, L. & Deily, M, E. (2008). Cost inefficiency and hospital health outcomes. *Health Economics*, **17**. pp. 833-848.

Nag, O, S. (2018). Countries With The Longest Coastline. *WorldAtlas.* https://www.worldatlas.com/articles/countries-with-the-most-coastline.html

Nayar, P. & Ozcan, Y. A. (2008). Data Envelopment Analysis Comparison of Hospital Efficiency and Quality. *J Med Syst* (**32**) pp. 193-199.

Nordic Casemix Centre. (2013). Description of Nordic Casemix Work. *NordCase*. Updated June30,2015.http://www.nordcase.org/document/1/12/0135738/aboutu_595c01f_description_of_nordic_casemix_work_2013_12.pdf

NOU 1996:5. (1996). Hvem skal eie sykehusene? *Regjeringen.* https://www.regjeringen.no/contentassets/73408e2182da4f479c0c2145845ba90b/no/pdfa/nou199619960005000dddpdfa.pdf

NOU 2003:1. (2003). Behovsbasert finansiering av spesialisthelsetjenesten. *Regjeringen.* https://www.regjeringen.no/no/dokumenter/nou-2003-1/id453861/sec3

OECD. (2009). Health care quality indicators project: patient safety indicators report 2009. *OECD Health Working Papers.* pp. 1-106.

OECD Data. (2018). Length of hospital stay. *OECD*. https://data.oecd.org/healthcare/length-of-hospital-stay.htm

Oslo Universitetssykehus. (2017). Perinatal mental helse. *Rikshospitalet.* pp. 1-48.

Oslo Universitetssykehus. (2020). Kartet over veien inn i TBS. *Oslo universitetssykehus*. https://oslo-universitetssykehus.no/fag-og-forskning/nasjonale-og-regionale-tjenester/tsb/nasjonal-kompetansetjeneste-tsb/vare-prosjekter/nasjonalt-ledernettverk-i-tsb/kartet-over-veien-inn-i-tsb

Petersen, N. (1990). Data Envelopment Analysis on a Relaxed Set of Assumptions. *Management Science*, **36**(3), pp. 305-314.

Riksrevisjonen. (2017). Effektivitet i sykehus. Dokument 3:2 (2017-2018). pp. 31-36. https://riksrevisjonen.no/globalassets/rapporter/no-2017-2018/effektivitet_i_sykehus.pdf

Simar, L. & Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*. pp. 31-64

Sosial- og helsedirektoratet. (2007). *Innsatsstyrt finansiering I helsetjenesten. En vurdering av aktuelle tiltak.***15-**1479. pp. 1-58

Statistics Norway. (2020). Patient statistics. *Statistics Norway.* https://www.ssb.no/en/statbank/table/10261/tableViewLayout1/

Varkevisser, M. & Schut, F, T. (2009). Hospital merger control – an international comparison. *Erasmus Universiteit Rotterdam*. iBMG Working Paper W2009.01.

Vold, B. & Hjemås, G. (2019). 1 av 4 konsultasjoner hos private spesialister. *Statistics Norway.* Published June 24, 2019. https://www.ssb.no/helse/artikler-og-publikasjoner/1-av-4-konsultasjoner-hos-private-spesialister

Vågnes, E. (2019). Stor variasjon i kjøretiden til nærmeste fødested. *Statistics Norway*. Updated September 3, 2019. https://www.ssb.no/helse/artikler-og-publikasjoner/stor-variasjon-i-kjoretiden-til-naermeste-fodested

WHO. What is Quality of Care and why is it important? *World Health Organization*. https://www.who.int/maternal_child_adolescent/topics/quality-of-care/definition/en/

Wooldridge, J. M. (2009). Introductory econometrics: a modern approach (**4th ed**). Mason, Ohio: *South-Wester Cengage Learning.*

Wooldridge, J. M. (2016). Introductory econometrics: a modern approach (**6th ed**). Boston, Massachusetts: *Cengage Learning.*

9 APPENDIXES

A Overview of the four Norwegian health regions

Blue: Helse Nord RHF

Red: Helse Midt-Norge RHF

Green: Helse Sør-Øst RHF

Yellow: Helse Vest RHF

(Oslo Universitetssykehus, 2020)

B Overview of horizontal hospital mergers in Norway

Health Enterprises	Comment
Finnmarkssykehuset HF	Juridical fusion.
Helgelandssykehuset HF	Juridical fusion.
Nordlandssykehuset HF	Juridical fusion.
UNN (Tromsø) HF	Juridical fusion.
Helse Møre og Romsdal HF	Juridical fusion.
Helse Nord-Trøndelag HF	Juridical fusion.
St. Olavs Hospital HF	Juridical fusion.
Helse Bergen HF	Juridical fusion.
Helsea Fonna HF	Juridical fusion.
Helse Førde HF	Juridical fusion. Closed: Nærsjukehuset i Florø (2008)
Hardaldsplass Diakonale Sykehus HF	No fusion.
Helse Stavanger HF	No fusion. Juridical establishment. Only one hospital in the HF.
Akershus universitetssykehus HF	Juridical fusion. Closed: Stensby sykehus (2013)
Diakonhjemmet Sykehus AS	No fusion.
Sykehuset Innlandet HF	Juridical fusion.
Lovisenberg Diakonale Sykehus AS	No fusion.
Oslo Universitetssykehus HF	Juridical fusion.
Sørlandet Sykehus HF	Juridical fusion. Closed: Mandal hospital (2005), Farsund (2007)
Sykehuset Telemark HF	Juridical fusion. Closed: Rjukan sykehus (2014)
Sykehuset Vestfold HF	Juridical fusion. Closed: Horten hospital (2005), Sandefjord hospital (2009)
Vestre Viken HF	Juridical fusion.
Sykehuset Østfold HF	Juridical fusion.

C Omitted hospitals from patient-level dataset from 1999 until 2007

Hospital	Reasoning
Andøy Fødestue	Delivery room. Not included in hospital data.
Axess Sykehus og spesialistklinikk	Owned by Aleris Helse. Not included in hospital data.
Betanien Hospital	Part of Helse Vest RHF and not included in hospital data.
Betanien Hospital Hordaland	Part of Helse Vest RHF and not included in hospital data.
Bergen Legevakt	Emergency room and not included in hospital data.
Brønnøysund Fødestue	Delivery room. Not included in hospital data.
Feiringklinikken	Provides specialist care for cardiac diseases. Not included in hospital data.
Geilomo Barnesykehus	Provides treatment for asthma and allergy to children. Not included in hospital data.
Glittreklinikken	Treatment for lung diseases. Not included in hospital data.
Granheim Lungesykehus	A part of Sykehuset Innlandet HF. Not included in hospital data
Hallingdal Sjukestugu	A part of Ringerike Sykehus HF. Not included in hospital data
Haugesund sanitetsforenings revmatismesykehus	Provides treatment for rheumatic diseases and skin diseases. Not included in hospital data.
Helsesentret Sonjatun	Not included in hospital data
Hjertesenteret I Oslo	Center of heart diseases located at Rikshospitalet HF. - not included in hospital data.
Kragerø komb. Helseinst.	Not included in hospital data
Kysthospitalet Hagevik	Not included in hospital data
Lillehammer sanitetsforenings revmatismesykehus	Provides treatment for rheumatic diseases and skin diseases. Not included in hospital data.
Lom Helseheim	Not included in hospital data
Martina Hansens Hospital	Privat hospital. Not included in hospital data.
Medi 3	Not included in hospital data.
Nordkapp Helsesenter	Not included in hospital data
Ringvoll Klinikken	Not included in hospital data
Spes.klinikken Drammen	Not included in hospital data
Spes.sh rehab, Kristiansand	Rehabilitation center and is therefore omitted.
Spesialsykehuset for rehabilitering	Rehabilitation center and is therefore omitted.
Statens senter for epilepsi	Provides treatment of epilepsy. Not included in hospital data.
Steigen Fødestue	Delivery room. Not included in hospital data.
Sunnaas Sykehus	Part of Helse Sør-Øst RHF, is a rehabilitation center, and is therefore omitted.
Tinn Fødestue	Delivery room. Not included in hospital data.
Valdres Fødestugo	Delivery room. Not included in hospital data.
Voksentoppen	Provides treatment for asthma and allergy to children. Not included in hospital data.
Volvat Medisinske Senter	Privat hospital. Not included in hospital data.
Ørland Fødestue	Delivery room. Not included in hospital data.

D Omitted hospitals from patient-level dataset from 2008 until 2014

Hospital	Reasoning
Akademiklinikken Oslo AS	Not included in hospital data
Aleris Helse AS Bergen	Not included in hospital data.
Aleris Helse AS Oslo	Not included in hospital data.
Aleris Helse AS Stavanger	Not included in hospital data.
Aleris Helse AS Tromsø	Not included in hospital data.
Aleris Helse AS Trondheim	Not included in hospital data.
Bergen Legevakt	Emergency room and not included in hospital data.
Betanien Hospital	Part of Helse Vest RHF and not included in hospital data.
Betanien spesialistpoliklinikk	Part of Helse Vest RHF. Polyclinic with specialist care, not included in hospital data.
Friskvernklinikken AS	Rehabilitation center and therefore omitted.
FysMed-Klinikken AS	Privat polyclinic not included in hospital data.
Granheim lungesykehus	Not included in hospital data
Haugesund sanitetsforenings revmatismesykehus	Provides treatment for rheumatic diseases and skin diseases. Not included in hospital data.
Helse More og Romsdal HF klinikk for Rehabilitering	Rehabilitation center and therefore omitted.
Hjertesenteret i Oslo	Center of heart diseases - not included in hospital data.
IbsenSykehuset AS	Not included in hospital data
Kolibri Medical AS	Not included in hospital data.
LHL-klinikkene Feiring	Provides clinical treatment and rehabilitation. Not included in hospital data
LHL-klinikkene Glittre	Provides clinical treatment and rehabilitation. Not included in hospital data.
Martina Hansens Hospital	Privat hospital. Not included in hospital data.
Medi 3 AS	Not included in hospital data.
Mjøskirurgene Lege og Spesialistsenter	Privat hospital not included in hospital data
NIMI AS avd. Mini Ullevål	Rehabilitation center. Not included in hospital data.
Oslo kommunale legevakt	Emergency room. Not included in hospital data
OUS HF Olafiaklinikken	Treats sexually transmittable diseases. Not included in hospital data.
Privathospitalet AS	Not included in hospital data.
Privatsykehuset Haugesund AS	Not included in hospital data.
Revmatismesykehuset AS	Provides treatment for rheumatic diseases. Not included in hospital data.
Ringvoll klinikken AS	Not included in hospital data.
Sunnaas Sykehus	Part of Helse Sør-Øst RHF and is a rehabilitation center and is therefore omitted
SVC Norge AS	Provides treatment for varicose veins. Not included in hospital data.
Teres Bergen	Provides plastic surgery. Not included in hospital data.
Teres Colosseum Stavanger	Provides plastic surgery. Not included in hospital data.
Teres Klinikken Bodø	Provides plastic surgery. Not included in hospital data.
Teres Rosenborg	Provides plastic surgery. Not included in hospital data.
Teres Stokkan	Provides plastic surgery. Not included in hospital data.
Teres Sørlandsparken	Provides plastic surgery. Not included in hospital data.
Teres Tromsø	Provides plastic surgery. Not included in hospital data.

E Descriptive statistics of patient characteristics, treatment variables and travel time.

	AHUS HF	**Diakonhjem.**	**Finnma. HF**	**Haraldsplass**	**Helgela. HF**	**H. Bergen HF**
Patient characteristics						
Male	46.37 %	40.75 %	44.44 %	51.43 %	45.54 %	46.14 %
Agegrp0	3.28 %	0.002 %	1.51 %	0.001 %	1.62 %	2.73 %
Agegrp1_9	8.10 %	0.009 %	5.68 %	0.11 %	4.96 %	8.08 %
Agegrp10_19	7.85 %	0.83 %	7.13 %	2.53 %	6.70 %	6.95 %
Agegrp20_29	10.12 %	6.03 %	9.46 %	7.89 %	9.10 %	11.21 %
Agegrp30_39	13.81 %	10.26 %	12.83 %	8.64 %	10.82 %	12.40 %
Agegrp40_49	11.90 %	12.37 %	13.46 %	10.83 %	12.14 %	10.71 %
Agegrp50_59	12.07 %	17.75 %	15.01 %	13.78 %	13.98 %	12.93 %
Agegrp60_69	14.45 %	21.22 %	15.44 %	17.65 %	16.03 %	15.05 %
Agegrp70_79	11.12 %	16.30 %	12.38 %	19.20 %	14.27 %	11.85 %
Agegrp80_89	6.38 %	12.40 %	6.71 %	16.21 %	9.01 %	6.75 %
Agegrp90	0.83 %	2.76 %	0.82 %	3.26 %	1.43 %	1.08 %
Treatment variables						
Transinownhospital	2.69 %	7.38 %	2.32 %	0.39 %	2.21 %	3.41 %
Transoutownhospital	2.69 %	7.38 %	2.31 %	0.37 %	2.20 %	3.40 %
Transinotherhospital	0.15 %	0.51 %	0.23 %	1.45 %	0.07 %	0.11 %
Transoutotherhospital	0.17 %	0.36 %	0.21 %	0.44 %	0.13 %	0.14 %
Average Traveltime (in hours)	0.45	0.37	1.73	0.48	0.77	0.85

	H. Fonna HF	**H. Førde HF**	**Lovisenberg**	**MøreRo. HF**	**Nordland. HF**	**Nordmøre HF**
Patient characteristics						
Male	45.01 %	45.50 %	48.18 %	46.80 %	45.30 %	46.18 %
Agegrp0	2.16 %	1.42 %	0.007 %	1.50 %	1.50 %	1.16 %
Agegrp1_9	7.71 %	7.19 %	2.71 %	6.09 %	5.70 %	6.19 %
Agegrp10_19	7.79 %	7.56 %	1.92 %	6.81 %	6.93 %	7.49 %
Agegrp20_29	10.95 %	8.13 %	7.81 %	8.12 %	9.19 %	8.12 %
Agegrp30_39	13.09 %	10.67 %	11.93 %	9.98 %	11.01 %	10.33 %
Agegrp40_49	10.22 %	11.73 %	13.81 %	10.90 %	12.07 %	10.77 %
Agegrp50_59	11.57 %	14.20 %	17.00 %	13.64 %	14.58 %	13.89 %
Agegrp60_69	13.54 %	15.49 %	19.14 %	17.57 %	16.09 %	15.48 %
Agegrp70_79	12.20 %	13.10 %	13.72 %	14.25 %	13.73 %	14.20 %
Agegrp80_89	9.15 %	9.18 %	9.74 %	9.37 %	8.18 %	11.05 %
Agegrp90	1.64 %	1.52 %	2.27 %	1.77 %	1.19 %	1.76 %
Treatment variables						
Transinownhospital	2.20 %	3.21 %	1.58 %	3.59 %	2.37 %	2.35 %
Transoutownhospital	2.19 %	3.20 %	1.57 %	3.59 %	2.35 %	2.34 %
Transinotherhospital	0.11 %	0.05 %	0.43 %	0.04 %	0.08 %	0.03 %
Transoutotherhospital	0.16 %	0.09 %	0.39 %	0.08 %	0.11 %	0.05 %
Average Traveltime (in hours)	0.72	1.39	0.42	0.76	1.37	0.80

	Nord-Tr. HF	OUS HF	St. Olavs HF	Stavanger HF	Sunnmøre HF	Innlandet HF
Patient characteristics						
Male	44.39 %	46.48 %	44.24 %	45.66 %	45.92 %	44.83 %
Agegrp0	1.62 %	1.66 %	1.90 %	2.93 %	2.00 %	1.33 %
Agegrp1_9	6.72 %	6.00 %	5.42 %	8.11 %	8.48 %	5.35 %
Agegrp10_19	8.02 %	5.71 %	6.41 %	6.61 %	7.34 %	6.65 %
Agegrp20_29	9.17 %	10.92 %	11.24 %	10.79 %	8.81 %	8.22 %
Agegrp30_39	10.28 %	15.26 %	13.03 %	12.96 %	11.37 %	9.98 %
Agegrp40_49	10.76 %	12.15 %	11.38 %	11.38 %	10.52 %	11.13 %
Agegrp50_59	13.43 %	13.71 %	14.05 %	12.98 %	13.41 %	13.88 %
Agegrp60_69	15.91 %	16.97 %	16.27 %	13.90 %	13.98 %	17.70 %
Agegrp70_79	13.88 %	11.68 %	12.59 %	11.28 %	13.46 %	14.62 %
Agegrp80_89	8.99 %	6.28 %	7.10 %	7.62 %	9.25 %	10.14 %
Agegrp90	1.26 %	1.11 %	0.91 %	1.25 %	1.33 %	1.63 %
Treatment variables						
Transinownhospital	2.73 %	6.63 %	3.99 %	2.22 %	2.31 %	1.71 %
Transoutownhospital	2.73 %	6.62 %	3.99 %	2.22 %	2.31 %	1.70 %
Transinotherhospital	0.05 %	0.42 %	0.07 %	0.06 %	0.03 %	0.10 %
Transoutotherhospital	0.09 %	0.30 %	0.04 %	0.05 %	0.09 %	0.20 %
Average Traveltime (in hours)	0.81	0.68	0.86	0.47	0.77	0.52

	Telemark HF	Sy. Vestfold HF	Sy. Østfold HF	Sørlandet sy. HF	UNN Tromsø HF	Vest.Viken HF
Patient characteristics						
Male	45.48 %	46.66 %	45.57 %	46.32 %	46.18 %	43.70 %
Agegrp0	1.72 %	1.78 %	2.57 %	2.05 %	1.30 %	2.30 %
Agegrp1_9	5.42 %	5.59 %	5.75 %	7.05 %	5.30 %	6.05 %
Agegrp10_19	6.08 %	5.89 %	5.56 %	6.76 %	6.42 %	5.84 %
Agegrp20_29	8.67 %	7.44 %	8.34 %	8.80 %	9.62 %	8.32 %
Agegrp30_39	11.99 %	10.34 %	11.05 %	10.87 %	12.29 %	12.19 %
Agegrp40_49	11.12 %	11.66 %	11.18 %	10.94 %	12.93 %	11.21 %
Agegrp50_59	13.58 %	14.50 %	13.82 %	13.42 %	15.18 %	13.12 %
Agegrp60_69	16.01 %	17.08 %	17.00 %	16.43 %	16.89 %	16.35 %
Agegrp70_79	13.56 %	14.50 %	14.68 %	13.88 %	12.87 %	13.80 %
Agegrp80_89	10.24 %	9.24 %	8.68 %	8.36 %	6.69 %	9.17 %
Agegrp90	1.63 %	1.41 %	1.23 %	1.33 %	0.86 %	1.68 %
Treatment variables						
Transinownhospital	1.84 %	2.05 %	2.08 %	2.27 %	3.55 %	2.29 %
Transoutownhospital	1.83 %	2.05 %	2.07 %	2.27 %	3.55 %	2.29 %
Transinotherhospital	0.09 %	0.10 %	0.08 %	0.06 %	0.15 %	0.16 %
Transoutotherhospital	0.15 %	0.09 %	0.14 %	0.07 %	0.10 %	0.24 %
Average Traveltime (in hours)	0.58	0.43	0.51	0.67	2.73	0.63

The table is based on outputs retrieved from STATA 16.

F Descriptive statistics of costs, number of patients, and capital for each HF.

HOSPITAL	VARIABLE	MEAN	STD. DEV	MIN	MAX
AHUS HF	cost	3253690	850145.5	2370505	4405038
	nupatients	256595.8	63895.63	173537	355703
	capital	329881.4	111468.1	133157	462702
Diakonhjemmet sykehus	cost	824177.9	79048.7	703573	956081
	nupatients	59760.92	13740.03	41732	85096
	capital	58574.67	7039.256	49246	73186
Finnmarksykehuset HF	cost	855636.2	37932.02	796317	915071
	nupatients	70478.15	3505.647	65326	77531
	capital	64846.69	24007.69	0	103154
Haraldsplass Diakonale sykehus	cost	590462.8	54819.59	499878	676895
	nupatients	31934.58	3778.902	25498	39482
	capital	41373.17	3437.751	37394	48200
Helgelandssykehuset HF	cost	865756.4	51122.73	773423	947090
	nupatients	88223.58	12224.86	65914	101896
	capital	63817.58	15131.3	42736	92197
Helse Bergen HF	cost	5532782	520778.7	4513451	6273603
	nupatients	440543.1	85113.4	338557	582255
	capital	468646.9	67742.51	394985	603424
Helse Fonna HF	cost	1604193	100994.6	1451955	1765462
	nupatients	143223.3	18091.16	116597	165567
	capital	123665.6	37138.22	88497	203942
Helse Førde HF	cost	1404053	72019.34	1299348	1520468
	nupatients	130892.7	11292.96	111639	149203
	capital	107640.2	48222.83	0	170757
Lovisenberg Diakonale sykehus	cost	737247.2	147329.3	571809	1054034
	nupatients	59736.33	13964.15	41426	77861
	capital	49529.33	6497.758	32286	59516
Møre og Romsdal HF	cost	3286881	41488.46	3249304	3334367
	nupatients	328052.3	9466.519	346687	369400
	capital	157551.8	16484.31	133415	169773
Nordlandssykehuset HF	cost	1872237	218928.7	1528167	2156993
	nupatients	146100.7	24029.04	111968	174316
	capital	131933.1	48347.17	0	211750
Nordmøre HF	cost	1244184	50013.62	1176996	1320476
	nupatients	122606.5	8408.507	111818	138446
	capital	93231.38	48806.11	0	157202
Nord-Trøndelag HF	cost	1484753	90835.5	1354898	1627756
	nupatients	128767	12947.77	111414	151577
	capital	89042.42	15244.25	71553	109123
OUS HF	cost	13500000	696278.7	12800000	24300000
	nupatients	982171.2	25229.5	940165	1007428
	capital	947802.2	190327.7	776654	1219277

St. Olavs hospital HF	cost	5251961	348255.6	4435770	5707000
	nupatients	392319	64032.93	324735	494293
	capital	451791.3	78027.81	305450	583891
Stavanger Universitetssykehus HF	cost	3280561	288690	2845950	3645483
	nupatients	301213.6	52847.23	210584	367305
	capital	199556.6	66478.81	7097	251114
Sunnmøre HF	cost	1863225	553691.6	1594076	3218586
	nupatients	179352.8	61027.63	140393	327258
	capital	124661.3	24374.55	85867	150022
Sykehuset Innlandet HF	cost	4127047	733658.7	1944283	4563943
	nupatients	376735.9	85256.93	136461	449815
	capital	285412.6	118396.4	0	420563
Sykehuset Telemark HF	cost	1817415	260913.1	1485770	2148874
	nupatients	182646.8	27935.34	147911	218921
	capital	140665.5	38826.76	96065	209545
Sykehuset Vestfold HF	cost	2289447	151402.5	2093829	2519180
	nupatients	230753	36090.92	183803	278703
	capital	194452.2	37431.84	146505	255428
Sykehuset Østfold HF	cost	2740750	99758.96	2624676	2944241
	nupatients	238625.8	30942.99	188611	277513
	capital	211439.1	47096.53	151838	275661
Sørlandet sykehus HF	cost	3085499	263713	2745858	3453880
	nupatients	300375.6	52346.5	228355	369590
	capital	231781	72134.89	161654	380470
UNN Tromsø HF	cost	3908377	126584.2	3748717	4130440
	nupatients	274581.6	17848.71	253136	293357
	capital	244511.1	30214.15	211724	291743
Vestre Viken HF	cost	4212375	118589.6	4093190	4395022
	nupatients	398072.7	24783.78	368934	434873
	capital	210670.7	30610.31	177768	242427

G Year dummies for the dependent variables in the DiD estimations on quality

Dependent variable: Performance_all (sum of the two quality indicators and the five patient safety indicators)

Explanatory Variable	Coefficient	Standard Error	P-value	95% Conf. Interval	
Year					
2000	.024	.039	0.544	-.053	.101
2001	-.016	.039	0.688	-.093	.061
2002	-.131*	.048	0.007	-.225	-.036
2003	-.161*	.048	0.001	-.256	-.066
2004	-.182*	.049	0.000	-.278	-.087
2005	-.198*	.049	0.000	-.295	-.101
2006	-.192*	.050	0.000	-.289	-.094
2007	-.202*	.050	0.000	-.300	-.104
2008	.010	.050	0.836	-.088	.109
2009	.004	.055	0.945	-.104	.112
2010	-.087	.056	0.122	-.199	.024
2011	-.070	.057	0.219	-.182	.042
2012	-.106	.057	0.065	-.219	.007
2013	-.093	.057	0.106	-.206	.020
2014	-.109	.058	0.060	-.223	.005

Significant coefficients at a 5% significance level are marked with *.

Dependent variable: Performance_main (sum of the to two quality indicators)

Explanatory Variable	Coefficient	Standard Error	P-value	95% Conf. Interval	
Year					
2000	-.057	.031	0.066	-.118	.004
2001	-.118*	.031	0.000	-.179	-.057
2002	-.236*	.038	0.000	-.311	-.162
2003	-.286*	.038	0.000	-.361	-.211
2004	-.310*	.038	0.000	-.385	-.235
2005	-.309*	.039	0.000	-.386	-.233
2006	-.354*	.039	0.000	-.430	-.277
2007	-.266*	.039	0.000	-.343	-.188
2008	-.280*	.039	0.000	-.357	-.202
2009	-.286*	.043	0.000	-.371	-.201
2010	-.343*	.044	0.000	-.431	-.255
2011	-.388*	.045	0.000	-.476	-.299
2012	-.406*	.045	0.000	-.495	-.317
2013	-.417*	.045	0.000	-.506	-.328
2014	-.438*	.045	0.000	-.527	-.348

Significant coefficients at a 5% significance level are marked with *.

Dependent variable: perf_u2_readm30_emgc

Explanatory Variable	Coefficient	Standard Error	P-value	95% Conf. Interval	
Year					
2000	-.004	.039	0.906	-.081	.072
2001	-.002	.039	0.959	-.078	.074
2002	-.154*	.048	0.001	-.247	-.060
2003	-.089	.048	0.065	-.183	.005
2004	-.086	.048	0.075	-.181	.009
2005	-.035	.049	0.470	-.132	.061
2006	-.040	.049	0.412	-.137	.056
2007	.144*	.049	0.004	.047	.241
2008	-.023	.050	0.642	-.121	.074
2009	-.027	.055	0.618	-.135	.080
2010	-.056	.056	0.316	-.166	.054
2011	-.109	.056	0.053	-.221	.001
2012	-.095	.057	0.094	-.207	.016
2013	-.111	.057	0.052	-.223	.001
2014	-.109	.057	0.057	-.222	.003

Significant coefficients at a 5% significance level are marked with *.

Dependent variable: perf_u2_inhosp_mortality_1

Explanatory Variable	Coefficient	Standard Error	P-value	95% Conf. Interval	
Year					
2000	-.109*	.044	0.013	-.195	-.023
2001	-.234*	.044	0.000	-.321	-.147
2002	-.319*	.054	0.000	-.425	-.213
2003	-.484*	.054	0.000	-.590	-.377
2004	-.534*	.054	0.000	-.641	-.427
2005	-.583*	.055	0.000	-.692	-.474
2006	-.667*	.055	0.000	-.776	-.558
2007	-.675*	.056	0.000	-.785	-.565
2008	-.536*	.056	0.000	-.646	-.425
2009	-.545*	.062	0.000	-.666	-.423
2010	-.630*	.063	0.000	-.754	-.505
2011	-.666*	.064	0.000	-.791	-.540
2012	-.717*	064	0.000	-.843	-.590
2013	-.723*	.064	0.000	-.849	-.596
2014	-.766*	.065	0.000	-.894	-.639

Significant coefficients at a 5% significance level are marked with *.

Dependent variable: average_traveltime

Explanatory Variable	Coefficient	Standard Error	P-value	95% Conf. Interval	
Year					
2000	.018	.087	0.833	-.153	.190
2001	.029	.088	0.744	-.144	.201
2002	-.165	.107	0.124	-.375	.045
2003	-.099	.108	0.356	-.312	.112
2004	-.092	.108	0.398	-.305	.121
2005	-.099	.110	0.365	-.317	.117
2006	-.080	.110	0.468	-.297	.137
2007	.012	.111	0.911	-.206	.231
2008	.052	.112	0.644	-.168	.271
2009	.098	.123	0.426	-.143	.339
2010	.068	.126	0.589	-.180	.316
2011	.093	.127	0.464	-.156	.343
2012	.102	.128	0.424	-.149	.354
2013	.093	.128	0.469	-.159	.345
2014	.102	.129	0.431	-.152	.355

Significant coefficients at a 5% significance level are marked with *.