

## Kvalitet og effektivitet i norsk helsevesen

En ikke-parametrisk og statistisk totrinnsanalyse av somatisk spesialisthelsetjeneste

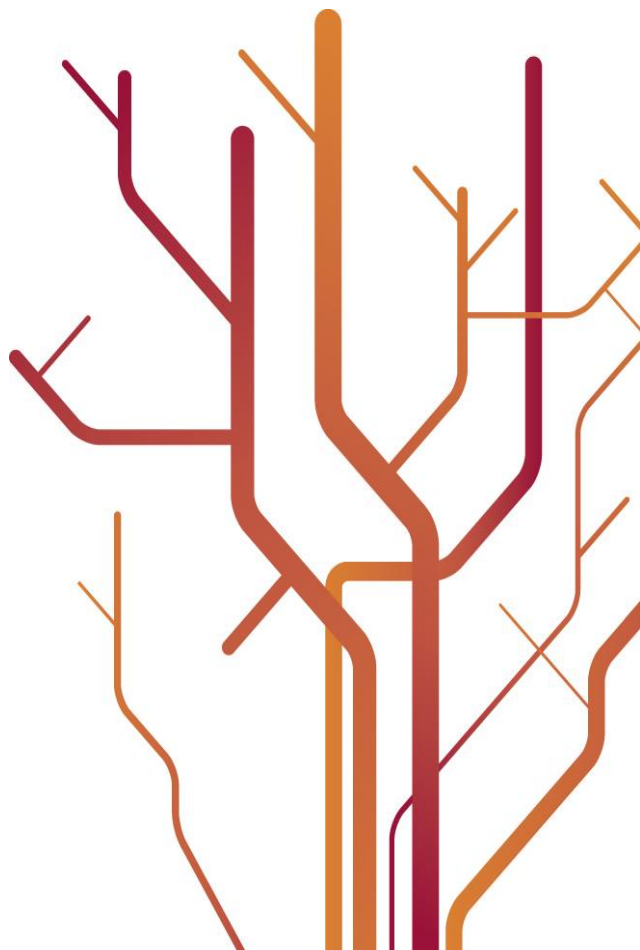


Johannes Martinussen

Ingar Reinholtsen

Masteroppgave i økonomi og administrasjon  
- studieretning bedriftsøkonomi (30 stp)

Mai 2013



## **Forord**

Denne masteroppgaven i økonomi og administrasjon avslutter et fem års langt studieløp ved Universitetet i Tromsø. Engasjementet til oppgaven er kommet både gjennom debattene i media og gjennom samtaler med Universitetssykehuset i Nord-Norge. Arbeidet har vært utfordrende, interessant og meget lærerikt.

I forbindelse med innlevering av denne oppgaven vil vi takke vår veileder, førsteamanuensis Helen Marita Sørensen Holst ved Handelshøgskolen i Tromsø, for konstruktiv kritikk, hjelp og støtte gjennom hele semesteret.

I tillegg vil vi takke alle ved Økonomi- og analysesenteret UNN som har hjulpet med informasjon og forslag til vår oppgave. Takk til Terje Vassdal, professor ved Handelshøgskolen i Tromsø, for inspirasjon og innspill. Professor Jon Magnussen, instituttleder ved institutt for samfunnsmedisin ved NTNU, takkes for vurdering og konstruktiv kritikk av våre resultater og metodebruk.

Til slutt ønsker vi å rette en stor takk til venner og familie for støtte gjennom hele studieløpet.

Tromsø, Mai 2013

---

Johannes Martinussen

---

Ingar Reinholtsen

## Sammendrag

Denne oppgaven undersøker forholdet mellom prosesskvaliteten og effektiviteten innen somatisk spesialisthelsetjeneste. Undersøkelsen er gjort på 19 sammenliknbare, norske helseforetak for perioden 2008 til 2011.

Studien er lagt opp som en totrinnsanalyse. I første trinn beregnes prosesskvaliteten, kostnadseffektiviteten og den tekniske effektiviteten for helseforetakene. Dette gjøres for hvert år ved hjelp av dataomhyllingsanalyser. For å beregne endringer mellom år er det brukt Malmquist-produktivitetsindekser.

I andre trinn brukes statistiske analysemetoder for å finne relasjoner i resultatene til dataomhyllingsanalysene og Malmquist-produktivitetsindeksene. Gruppedifferanseanalyser sammenlikner effektiviteten til helseforetak med god og dårlig prosesskvalitet, og kvaliteten til helseforetak med god og dårlig effektivitet. Korrelasjonsanalyser viser hvorvidt det eksisterer samvariasjon mellom kvaliteten og effektiviteten. Til slutt brukes regresjon for å forsøke å påvise en signifikant årsakssammenheng mellom kvalitet og effektivitet.

I modellen for prosesskvalitet inngår fire variabler: *prosentandel av epikriser sendt etter 7 dager*, *prosentandel av lårhalsbruddoperasjoner etter 48 timer*, *prosentandel strykninger av planlagt operasjonsprogram*, og *prosentandel korridorpasienter*. Modell for kostnadseffektivitet består av to variabler: *DRG-produksjon* og *DRG-kostnader*. Modell for teknisk effektivitet består av fem variabler: *DRG-produksjon*, *årsverk leger*, *årsverk behandlingspersonell*, *årsverk annet*, og *antall døgnplasser*.

Resultatene påviser ingen klar relasjon mellom kvalitet og effektivitet, men signaliserer at helseforetak som har god prosesskvalitet har gode sjanser for å ha høy effektivitet – og omvendt. Samtidig er det en svak trend til at forbedringer i prosesskvalitet relateres negativt til teknisk effektivitet.

Beregningene i oppgaven er utført ved bruk av programpakken R (versjon 2.15.0) med tilleggspakkene «Benchmarking» og «FEAR», og Microsoft Office Excel 2013 med tilleggspakkene «Analysis Toolpak» og «PHStat2».

Nøkkelord: kvalitet, prosesskvalitet, effektivitet, kostnadseffektivitet, teknisk effektivitet, helseforetak, dataomhyllingsanalyse, Malmquist-produktivitetsindeks

## Innholdsfortegnelse

1.	Innledning.....	1
1.1.	Bakgrunn .....	1
1.2.	Formål og avgrensning.....	2
1.3.	Problemstillingen.....	3
1.4.	Oppgavens struktur.....	4
2.	Det norske helsevesen .....	5
2.1.	New Public Management i helsevesenet .....	5
2.2.	Helsevesenets organisering og helsereformen 2002.....	6
2.3.	Verdier, ISF og DRG-systemet .....	6
2.4.	Benchmarking: SAMDATA og kvalitetsindikatorer.....	8
3.	Teori .....	9
3.1.	Produktivitet og effektivitet.....	9
3.2.	Kvalitet .....	10
3.3.	Litteratursammendrag .....	16
4.	Metode.....	20
4.1.	Produksjonsteknologi og distansefunksjoner .....	20
4.2.	Farrells effektivitetsbegreper .....	22
4.2.1.	Teknisk effektivitet i praktisk beregning.....	23
4.3.	Dataomhyllingsanalyse .....	25
4.3.1.	Supereffektivitet .....	29
4.3.2.	Dysons fallgruver .....	30
4.3.3.	Kritikk: substituerbare og ikke-substituerbare inputer og outputer .....	32
4.3.4.	Kritikk: Inputer og outputer i spesialisthelsetjeneste .....	33
4.4.	Malmquist-produktivitetsindeks.....	33
4.4.1.	Effektivitetsendring og teknologisk endring .....	34
4.4.2.	Malmquist med variert skalautbytte .....	35
4.4.3.	Malmquist-indeksen ved endringer i skala.....	36
4.5.	Statistiske metoder .....	37
4.5.1.	Kolmogorov-Smirnov.....	37
4.5.2.	Korrelasjonsanalyse.....	37
4.5.3.	Regresjonsanalyse .....	38
4.6.	Valg og testing av datasett.....	41
4.6.1.	Korrelasjonsanalyse og gruppetester.....	42
4.6.2.	Outlieranalyse.....	43
4.6.3.	Sensitivitetsanalyse.....	44
5.	Data .....	46
5.1.	Utvalg .....	46

5.2.	Validitet og reliabilitet.....	47
5.2.1.	Krav til dataomhyllingsanalyser.....	47
5.3.	Sekundærdata .....	48
5.4.	Konstruksjon av modeller.....	48
5.5.	Statistisk testing av variabler.....	50
5.5.1.	Korrelasjon .....	51
5.5.2.	Banker-test.....	52
5.6.	Endelige modeller.....	55
5.6.1.	Modell for kvalitet.....	55
5.6.2.	Modell for kostnadseffektivitet .....	56
5.6.3.	Modell for teknisk effektivitet.....	57
5.7.	Outliers .....	58
5.8.	Skalautbytte.....	60
6.	Resultater.....	62
6.1.	Trinn 1: DEA- og MPI-resultater .....	62
6.1.1.	Resultater for dataomhyllingsanalysene.....	62
6.1.2.	Resultater for Malmquist-produktivitetsindekser.....	67
6.1.3.	Sensitivitetsanalyse.....	73
6.2.	Trinn 2: Statistisk analyse .....	74
6.2.1.	Gruppedifferanseanalyser.....	75
6.2.2.	Korrelasjon .....	80
6.2.3.	Regresjon.....	81
7.	Diskusjon og konklusjon .....	85
7.1.	Oppsummering og viktige momenter .....	85
7.2.	Tolkning av resultater.....	87
7.3.	Sammenlikning mot tidligere litteratur .....	88
7.4.	Konklusjon .....	89
7.4.1.	Veien videre .....	90
8.	Referanser.....	91
	Vedlegg .....	95
	Vedlegg 1: Liste over helseforetak.....	95
	Vedlegg 2: Liste over tilgjengelige kvalitetsindikatorer fra Helsedirektoratet .....	96
	Vedlegg 3: Outlier-analyse.....	97
	Vedlegg 4: R-script .....	101

## Figuroversikt

Figur 1: Struktur-, prosess- og resultat kvalitet .....	13
Figur 2: Dekomponering av effektivitet .....	22
Figur 3: Farrells effektivitet .....	24
Figur 4: Eksempel på Eladder .....	45
Figur 5: Skalautbytte .....	61

## Tabelloversikt

Tabell 1: Beregning av DRG-kostnader .....	50
Tabell 2: Korrelasjonskoeffisienter for tekniske variabler .....	51
Tabell 3: Banker-test modellspekifikasjoner for kvalitet .....	53
Tabell 4: Banker-test på datasett for kvalitet (F-verdier) .....	53
Tabell 5: Banker-test modellspekifikasjoner for teknisk effektivitet .....	54
Tabell 6: Banker-test på datasett for teknisk effektivitet (F-verdier) .....	54
Tabell 7: Outlieranalyse kostnadseffektivitet .....	59
Tabell 8: Outlieranalyse teknisk effektivitet .....	59
Tabell 9: Dataomhyllingsanalyser, 2008 .....	63
Tabell 10: Dataomhyllingsanalyser, 2009 .....	64
Tabell 11: Dataomhyllingsanalyser, 2010 .....	65
Tabell 12: Dataomhyllingsanalyser, 2011 .....	66
Tabell 13: Malmquist-produktivitetssendring .....	67
Tabell 14: Malmquist-effektivitetssendring (CRS) .....	68
Tabell 15: Malmquist teknologisk endring (CRS) .....	69
Tabell 16: Malmquist-effektivitetssendring (VRS) .....	70
Tabell 17: Malmquist teknologisk endring (VRS) .....	71
Tabell 18: Malmquist-skalaendring .....	72
Tabell 19: Sensitivitetsanalyse for teknisk effektivitet .....	73
Tabell 20: Sensitivitetsanalyse for kostnadseffektivitet .....	73
Tabell 21: Sensitivitetsanalyse for kvalitet .....	74
Tabell 22: Deskriptiv statistikk for DEA-resultater .....	75
Tabell 23: Deskriptiv statistikk for MPI-resultater .....	75
Tabell 24: Gruppendifferanseanalyseresultater - kvalitet (CRS) .....	76
Tabell 25: Gruppendifferanseanalyseresultater - effektivitet (CRS) .....	76
Tabell 26: Gruppendifferanseanalyseresultater - effektivitet (VRS) .....	77
Tabell 27: Gruppendifferanseanalyseresultater - Malmquist-produktivitetssendring .....	78
Tabell 28: Gruppendifferanseanalyseresultater - Malmquist-effektivitetssendring (CRS) .....	79
Tabell 29: Gruppendifferanseanalyseresultater - Malmquist-effektivitetssendring (VRS) .....	79
Tabell 30: Korrelasjonsanalyse DEA .....	80
Tabell 31: Korrelasjonsanalyse MPI .....	80
Tabell 32: Regresjon DEA modell-1 CRS .....	81
Tabell 33: Regresjon DEA modell-1 VRS .....	82
Tabell 34: Regresjon DEA modell-2 .....	82
Tabell 35: Regresjon DEA-4 .....	83
Tabell 36: Regresjon MPI-1 .....	83
Tabell 37: Regresjon MPI-2 .....	84
Tabell 38: Regresjon MPI-3 .....	84

# 1. Innledning

## 1.1. Bakgrunn

Det norske helsevesenet er en stor og essensiell del av offentlige tjenester. I løpet av 2012 var de totale helseutgiftene på 270 milliarder, og utgjorde totalt 9,4 % av Norges BNP. Samtidig vokser helseutgiftene mer enn andre offentlige utgifter [1]. Kvalitet og effektivitet har lenge vært tema i norsk helsevesen. Opp gjennom årene har flere tiltak blitt gjort for å forbedre både kvaliteten og effektiviteten. To større tiltak har vært implementeringen av innsatsstyrt finansiering i 1997 og helsereformen i 2002.

Innsatsstyrt finansiering (ISF) ble innført i 1997 som den nye finansieringsmodellen for norske somatiske sykehus. Igjennom ISF skulle sykehusene finansieres etter mengden arbeid som ble utført i stedet for kun et rammeverk (St.meld. nr. 44 (1995-1996)). Formålet var mer smidige sykehus som ikke var bundet til forhåndsbestemte budsjettammer. Samtidig kunne ISF motivere til effektivitetstiltak internt i sykehus ettersom effektivisering betød ressurser til overs.

Helsereformen av år 2002 brakte flere organisatoriske endringer til spesialisthelsetjenesten. Sykehusvirksomheten, som tidligere var organisert som forvaltningsorganer, ble reformert til en foretaksmodell. Dette medførte at sykehusene ikke lenger var en integrert del av staten, og nå var selvstendige enheter. Reformen resulterte også i at de nyetablerte foretakene var fullstendig statseid og ikke kunne bli slått konkurs. Helsereformen var et skritt i en mer markedsorientert retning. Sykehusene fikk mer ansvar og kontroll over egen økonomi og produksjon for slik å få mest mulig ut av ressursene til rådighet. Formålet med helsereformen var et bedre helsevesen med høyere kvalitet, høyere effektivitet og bedre tilgjengelighet (Ot.prp. nr. 66 (2000-2001)).

Tiltak som ISF og helsereformen går inn under en ny ledelsesfilosofi kalt «New Public Management» (NPM), hvor offentlige instanser omorganiseres til mer markedsorienterte institusjoner med fokus på effektiv drift (NOU 2004:2). Denne mer markedsorienterte styringen av spesialisthelsetjenesten har skapt protester fra blant annet helsepersonell. Protestene bygger på påstanden om at effektivitet og kvalitet står i opposisjon til hverandre. Et fokus på effektivisering og kostnadsbesparelser vil etter deres meninger medføre lavere kvalitet i helsevesenet [2, 3, 4].

Påstandene om et negativt forhold mellom effektiviteten og kvaliteten i helsevesenet er ikke støttet av alle. En arbeidsmiljøundersøkelse fra Helse Midt-Norge skal ha vist til en positiv relasjon mellom effektiviteten og kvaliteten ved psykiatriske avdelinger [5]. Statssekretær Kåss i Helse- og omsorgsdepartementet mener ikke nødvendigvis det eksisterer en interessekonflikt mellom kvalitet og effektivitet [6]. Ifølge hans syn vil kvaliteten og effektiviteten påvirke hverandre gjensidig positivt og negativt. Han trekker også fram sykehuset i Vestfold som gikk fra verst til best kostnadseffektivitet på ti år. Sykehuset i Vestfold var det mest kostnadseffektive sykehuset i Norge i 2009, og gjorde det godt på landets kvalitetsindikatorer. Statssekretær Kåss mener også at engasjement fra ledelsen er svært viktig for god kvalitet i sykehus. Akershus universitetssykehus kan også vise til to tiltak som har bedret både kvaliteten og effektiviteten deres [7]. Ett av disse var en kampanje for forebygging mot sykehusinfeksjoner. Reduksjonen av sykehusinfeksjoner reduserte behandlingstiden til pasienter og dermed frigjorde kapasitet, ressurser og samtidig sparte universitetssykehuset mange kostnader relatert til infeksjonene.

## 1.2. Formål og avgrensning

Uenighetene har motivert oss til å undersøke forholdet mellom effektiviteten og kvaliteten i norsk helsevesen. Oppgaven vår er derfor dedikert til å undersøke hvorvidt det eksisterer en negativ eller positiv relasjon mellom kvaliteten og effektiviteten i helsevesenet, eller om det ikke eksisterer noen åpenlys relasjon.

Vi begrenser vår oppgave til å omhandle den somatiske spesialisthelsetjenesten i stedet for helsevesenet som en helhet. Dette fordi majoriteten av protestene kommer fra helsepersonell i spesialisthelsetjenesten samtidig som reformen og ISF er beregnet på spesialisthelsetjenesten. Begrensningen til den somatiske spesialisthelsetjenesten er kun for å enklere kunne sammenlikne sykehus. En annen begrensning av oppgaven er å se på kun begrepene kvalitet og effektivitet. Dette medfører at alle interne og eksterne forhold som påvirker kvaliteten og effektiviteten ikke blir målt i seg selv.



### 1.3. Problemstillingen

Med bakgrunn i oppgavens formål og avgrensning har vi formulert følgende problemstilling:

*Eksisterer det en direkte relasjon mellom kvaliteten og effektiviteten til helseforetak i somatisk spesialisthelsetjeneste?*

For å besvare problemstillingen estimerer vi kvaliteten og effektiviteten til norske helseforetak i perioden 2008-2011. Dette gjør vi både for hvert år og endring mellom årene. Estimaten vil så vurderes på forskjellige måter for å se om relasjoner kan påvises.

For å beregne ytelsen i form av kvalitet og effektivitet i hvert år brukes dataomhyllingsanalyser. Ved bruk av dataomhyllingsanalyser kan man, basert på en rekke variabler, regne ut én relativ score for hvert helseforetak. Analysens egenskaper gjør den svært anvendelig i situasjoner hvor man vurderer ytelsen til homogene enheter, eksempelvis i offentlig sektor. Resultatene fra dataomhyllingsanalysene forteller hvilke helseforetak som har den beste kvaliteten, og hvilke helseforetak som har den beste effektiviteten. Dette gir en mulighet til å se hvordan effektiviteten er ved høy og lav kvalitet. Omvendt gir det også mulighet til å se hvordan kvaliteten ligger ved høy og lav effektivitet.

Beregningen av kvalitets- og effektivitetsendringer gjøres ved bruk av Malmquist-produktivitetsindekser (MPI). Indeksen gir mulighet for både å se hvordan et helseforetaks ytelse endres og hvordan effektivitet og kvalitet endres relativt i forhold til de beste helseforetakene. Ved bruk av MPI kan vi se om helseforetakenes endringer innen kvalitet og effektivitet går i samme retning. Samtidig ønsker vi å se hvorvidt det finnes en statistisk samvariasjon mellom helseforetakenes kvalitet og effektivitet. Dette vil undersøkes ved hjelp av korrelasjonsanalyser. I tråd med påstandene fra debatten, vil vi også bruke regresjonsanalyser til å se om en årsakssammenheng mellom kvalitet og effektivitet kan bevises.

#### 1.4. Oppgavens struktur

Oppgavens kapittel 2 inneholder generell, relevant informasjon om helsevesenet som studieobjekt. I kapittelet ser vi nærmere på New Public Management (NPM) som den ledende filosofien i helsevesenet, helsevesenets organisering, verdisystemet og innsatsstyrt finansering inkludert DRG-systemet. Kapittelet inneholder også informasjon om benchmarking i helsesektoren, en del som er viktig for den empiriske undersøkelsen.

Kapittel 3 omfatter oppgavens teoretiske rammeverk. Relevante begreper presenteres her. Kapittelet inneholder også et litteratursammendrag for tidligere studier gjort på effektivitet og effektivitet og kvalitet.

Kapittel 4 redegjør for de metodiske verktøyene som brukes i oppgaven. Her forklares dataomhyllingsanalyser og Malmquist-produktivitetsindekser, samt statistiske tester som korrelasjonsanalyser, regresjonsanalyser, Banker-tester og til slutt sensitivitetsanalyser.

Kapittel 5 presenterer oppgavens datasett, samt en vurdering av de modeller og variabler som er brukt i analysen.

Kapittel 6 er samlingen av de relevante resultatene fra studien. Her presenteres de rene resultatene fra dataomhyllingsanalysene, Malmquist-produktivitetsindeksene og de statistiske testene gjort på resultatene.

Kapittel 7 dedikeres til diskusjon og konklusjon av oppgaven. Her oppsummeres resultatene i lys av metodens styrker og svakheter før de settes opp mot resultatene fra tidligere litteratur og til slutt oppgavens problemstilling.

## 2. Det norske helsevesen

Dette kapitlet redegjør for noen essensielle egenskaper ved det norske helsevesen. Kapitlet starter med en introduksjon til styring av offentlig sektor, som ligger til grunn for problemstillingen. Videre gis en forklaring av helsevesenets organisering som er med på å forklare vårt valg av analyseobjekter. Deretter gjennomgås verdier og kvalitet i helsevesenet, finansiering og DRG-systemet, og til slutt hvordan sykehus kan sammenliknes i Norge.

### 2.1. New Public Management i helsevesenet

New public management (NPM) er en doktrine eller filosofi som hadde oppstart i OECD-land rundt 1980-tallet (Hood, 1995). Doktrinen skulle hjelpe mot korrupsjon i den offentlige sektor. Politikere kunne potensielt bruke sin makt til sine egeninteresser og venners interesser. Med mye korrupsjon og egeninteresse fra politikernes side risikerte offentlig sektor å bli en dyr affære med lav produktivitet.

Datidens «progressive public administration»-doktriner (PPA) mente offentlig sektor måtte skille seg fra privat sektor. PPA innebar også et fokus på prosesser og prosedyrer som skulle sikre at offentlig sektor ikke ble korrumpert, og at politikere ikke handlet etter egeninteresse. Innføringen av NPM medførte at offentlig sektor ikke skulle skille seg vekk fra privat sektor, og i stedet fjerne de forskjellene som eksisterte mellom offentlig og privat sektor. Fokuset ble også endret til et fokus på resultater.

New public management viser til syv endringer. Offentlig sektor skal deles inn i mindre, foretaksliknende enheter, og hver enhet har ansvaret for et eget offentlig produkt samt kostnadene relatert til dette. Konkurransen mellom offentlige organisasjoner og mellom offentlig og private organisasjoner øker. Ledelsesmodeller oftest brukt i privat sektor er aktuelle i offentlig sektor også. Fokus øker på å finne mer kostnadseffektive og ressurseffektive løsninger. Offentlig ledelse skal bli mer synlig og ha mer direkte involvering i sine organisasjoner. Ytelsesmål for offentlig sektor blir tydeligere og med observerbare ytelsesstandarder i offentlige organisasjoner. Fokus på offentlig organisasjons resultater og output blir større, hvor kontroll og belønning fokuserer på resultat og ikke prosess.

## 2.2. Helsevesenets organisering og helsereformen 2002

Det norske helsevesen består av flere organisasjoner hvor staten har det overordnede ansvar og eierskap igjennom Helse- og omsorgsdepartementet [8]. Helsevesenets komplekse struktur medfører at det kan deles inn på flere forskjellige måter. Det første, essensielle skillet ligger mellom primærhelsetjenesten og spesialisthelsetjenesten. Primærhelsetjenesten har ansvar for landets sykehjem, legevakter og fastleger. Spesialisthelsetjenesten er ansvarlig for sykehus, legespesialister og ambulansetjeneste. I helsedirektoratets SAMDATA-rapporter deles spesialisthelsetjenesten inn i tre sektorer [9]. Dette resulterer i en somatisk spesialisthelsetjeneste, psykiatrisk spesialhelsetjeneste og en sektor for tverrfaglig spesialisert rusbehandling. For å begrense vår oppgave, fokuserer vi videre på den somatiske spesialhelsetjenesten.

Da staten igjennom Helse- og omsorgsdepartementet overtok eierskap og ansvar av spesialhelsetjenestene i Norge i 2002, medførte dette en omorganisering av spesialhelsetjenestene [8]. Norge fikk de fire regionale helseforetakene (RHF): Helse Nord RHF, Helse Midt-Norge RHF, Helse Vest (RHF) og Helse Sør-Øst RHF. Under de regionale helseforetakene finner vi videre inndelinger i helseforetak (HF) For eksempel består Helse Nord RHF av Helgelandssykehuset HF, Helse Finnmark HF, Nordlandssykehuset HF, Sykehusapoteket Nord HF og Universitetssykehuset Nord-Norge HF. Universitetssykehuset Nord-Norge HF har under seg sykehusene Universitetssykehuset Nord-Norge, Harstad sykehus (UNN Harstad), Narvik Sykehus (UNN Narvik), Longyearbyen sykehus og Åsgård sykehus (psykiatrisk avdeling). Denne typen inndeling er lik for alle helseforetak innen de forskjellige regionale helseforetakene selv om antallet og størrelsen på sykehusene varierer mellom helseforetakene. Effektene av helsereformen i 2002 er analysert av Kittelsen et al. (2007). De fant at helsereformen hadde en statistisk signifikant positiv effekt på produktiviteten i norsk helsevesen.

## 2.3. Verdier, ISF og DRG-systemet

Nasjonal helseplan er et statlig dokument, utarbeidet av Helse- og omsorgsdepartementet, som redegjør for statusen til det norske helsevesen. For oss er dokumentet interessant ettersom det gir et bilde av helsevesenets verdier og de planer som er lagt for helsevesenet. I vår redegjørelse vil vi fokusere på somatiske helseforetak innenfor helsevesenet. Nasjonal

helseplan for 2007-2010 legger opp seks fundamentale mål og verdier som skal være fellesnevnerne for alle ledd i helsevesenet (St.prp. nr. 1 (2006–2007)):

*Helhet og samhandling* skal være fellesnevnerne for alle helsevesenets tjenester. Dette medfører blant annet åpen og robust kommunikasjon mellom helsevesenets organisasjoner slik at tjenester kan leveres av flere organisasjoner i samarbeid for å øke nytten til pasienter og brukere. *Demokrati og legitimitet* medfører at befolkningen skal ha tillit til helsetjenesten. Helseorganisasjonene skal drives på en profesjonell måte. Samtidig skal helsevesenet være politisk styrt basert på åpenhet og medvirkning. *Nærhet og trygghet* betyr at helsetjenestene skal være tilgjengelige for hele befolkningen. Alt som kan desentraliseres skal desentraliseres, og kun tjenester som må være sentralisert skal forbli sentralisert. Lokale institusjoner skal også tilrettelegge for de beste mulige tilbud for sine kommunale områder ved laveste, effektive omsorgsnivå. *Sterkere brukerrolle* er et mål hvor brukerne skal ha kunnskaper og påvirkningskraft over tjenestene de bruker. Dette fordi pasientene og pasientenes pårørende ofte kjenner til deres egne situasjoner best. Dette gjør dem bedre i stand til å ta visse beslutninger enn ordinært helsepersonell. Samtidig skal den sterkere brukerrollen medvirke til høyere menneskeverd og respekt. *Faglighet og kvalitet* er påkrevd å være på et høyt nivå etter internasjonale standarder. Samtidig skal helsevesenet være åpen om feil og kvalitetsforbedring med hensyn på legitimiteten til tjenestene. Til slutt nevnes *arbeid og helse*. En god arbeidssituasjon med lønn og anseelse er viktig for helsen. Dette betyr at helsevesenet på unngå sosial utstøtning av grupper som faller ut av arbeidslivet.

Nasjonal helseplan 2007-2010 redegjør også for det todelte finansieringssystemet som brukes i spesialisthelsetjenesten i dag. I dokumentet argumenteres det for at finansieringssystemet skal fremme helsepolitiske mål. Første del er en basisfinansiering basert på historisk informasjon og hvordan neste års forbruk ser ut. Andre del består av innsatsstyrt finansiering (ISF). ISF-systemet skal sikre større frihet for spesialisthelsetjenesten ved å gi refusjoner/finansiering basert på helseforetakenes faktiske aktivitetsnivå. Den innsatsstyrte finansieringen er basert på systemet for diagnose relaterte grupper (DRG). DRG-systemet klassifiserer pasientene ved norske sykehus etter deres diagnoser og behandlingstiltak. Prinsipielt skal gjennomsnittlig ressurskrevende pasienter medføre større refusjoner enn pasienter som krever mindre ressurser. Hagen et al. (2001) fant positive relasjoner mellom innføringen av ISF i 1997 og kortsiktig økning i effektivitet i spesialisthelsetjenesten. Den kortsiktige økningen ble begrunnet i at studien strakk seg fra 1992 til 1999, og gikk derfor kun effektene i løpet av en toårsperiode. En annen studie av Biørn, et al. (2009) tok for seg

forskjellige responser av ISF-innføringen. De fant at sykehus som reagerte på ISF-innføringen hovedsakelig økte sin effektivitet. Deres hypotese om at sykehus med lavest effektivitet ville ha sterkest respons holdt likevel ikke.

Ifølge helsedirektoratets egen hjemmeside er DRG-systemet også brukbart til sammenlikning av norske sykehus og helseforetak [10]. Dette fordi DRG-systemet tillater at pasientene vektas etter hvor ressurskrevende behandlingen er.

#### 2.4. Benchmarking: SAMDATA og kvalitetsindikatorer

For vår studie er det spesielt to typer rapporteringer som er interessante med tanke på benchmarking innen spesialisthelsetjenesten: SAMDATA-rapportene og de Nasjonale kvalitetsindikatorer.

SAMDATA står for «sammenlikningsdata for spesialisthelsetjenesten» og er et prosjekt for benchmarking av norske sykehus [11]. Økonomi og analyseavdelingen til Helsedirektoratet har ansvaret for utgivelsen av de årlige SAMDATA-rapportene. Rapportene inneholder informasjon om somatisk, psykisk og rusbehandling spesialisthelsetjeneste. Presenterte data er korrigert for å gjøre sammenlikning mellom sykehus/helseforetak og sammenlikning mellom år så riktig som mulig. Rapportene gir informasjon om aktivitetsnivå, kostnader, ressursbruk med mer.

Nasjonale kvalitetsindikatorer publiseres også av Helsedirektoratet [12]. Hensikten med kvalitetsindikatorerne er å vise hvor hvert enkelt sykehus, helseforetak og regionalt helseforetak ligger på landsbasis. Beregningen av de forskjellige indikatorene er godt forklart og dokumentert for hver enkelt indikator. Helsedirektoratet er også påpasselig med hvilke indikatorer som faktisk er målt og hvilke som er beregnet uten å faktisk bli rapportert.

### 3. Teori

Denne oppgavens hovedmål er å undersøke relasjoner mellom kvalitet og effektivitet i norsk somatisk spesialisthelsetjeneste. Følgelig er det nødvendig å ha en grunnleggende forståelse for de begrepene som måles og de teoretiske konseptene som ligger til grunn. Teorikapittelet gir først en gjennomgang av de to begrepene som måles i oppgaven – kvalitet og effektivitet – som medfører en dekomponering av begrepene og ytterligere spesifiseringer rundt hva som er ønskelig å måle. Videre følger en gjennomgang av relevant litteratur, som vil gi en pekepinn på hvordan lignende studier har blitt utført og hvilke resultater de har gitt.

#### 3.1. Produktivitet og effektivitet

Med utgangspunkt i økonomisk produksjonsteori, definerer Kittelsen & Førstund (2001, s.23) *produktivitet* som «forholdet mellom produksjon og ressursbruk», og *effektivitet* som «forholdet mellom faktisk produktivitet og en norm for best mulig produktivitet gitt de begrensningene en virksomhet driver innenfor». Begrepene defineres med mye forskjellig ordlyd, men prinsipielt er innholdet det samme. Ved å benevne ressursbruken i produksjonsprosessen som *input* ( $x$ ), og produktet av prosessen som *output* ( $y$ ), kan man matematisk uttrykke begrepene som:

$$\text{Produktivitet} = \frac{y}{x} \quad (1)$$

$$\text{Effektivitet} = \frac{\text{faktisk produktivitet}}{\text{optimal produktivitet}} \quad (2)$$

I samfunnsøkonomisk forstand, deles effektivitetsbegrepet inn i «ytre» og «indre» effektivitet. *Ytre effektivitet* handler om å gjøre de rette tingene, og er et spørsmål om avveining mellom verdien av produksjon og alternativverdien av ressursene som brukes. I offentlig sektor kan avveiningen handle om prioriteringer og allokering av ressurser mellom ulike formål; eksempelvis om det skal investeres i nytt sykehus eller skole. *Indre effektivitet* handler om å gjøre ting rett, og er et spørsmål om hvorvidt produksjonen foregår med bruk av minst mulig ressurser. Oppgaven tar for seg sistnevnte type effektivitet, følgelig vil indre effektivitet videre bare omtales som *effektivitet*. (Kittelsen & Førstund, 2001; Kjekshus, 1999).

Å fastsette faktisk og optimal produktivitet, og dermed effektiviteten, blir vanskeligere hvis man har med flere produkter og innsatsfaktorer å gjøre. Enda vanskeligere blir det hvis det ikke er prisinformasjon om de ulike produktene og innsatsfaktorene. I motsetning til privat sektor hvor det etableres markedspriser, er det ikke like enkelt å oppdrive pålitelige priser innen offentlig sektor.

Kittelsen & Førstund (2001) anbefaler å bruke et sett med effektivitetsbegreper foreslått av Farrell (1957), som har etablert seg som en standard i effektivitetsberegningen. I Farrells løsning definerer man hva som maksimalt er mulig å produsere gitt en ressursbruk, eller motsatt hva som er nødvendig ressursbruk for å oppnå en gitt produksjon, og behovet for priser faller bort. Nærmere forklaring av denne løsningen, og hvordan man kan benytte seg av prinsippene for måling av effektivitet, vil gjøres rede for i metodekapittelet. Løsningen medfører at effektivitet dekomponeres til to komponenter; teknisk effektivitet og priseffektivitet. Kombinasjonen av disse utgjør total effektivitet, som tradisjonelt omtales som kostnadseffektivitet.

Denne inndeling har blitt standard i økonomiske fag, og vil være gjennomgående i oppgaven vår. Ved måling av kostnadseffektivitet brukes de totale kostnadene relatert til behandling i stedet for priser og ressurser. Teknisk effektivitet måles uten priser, men benytter seg av relative vekt for de forskjellige ressurstypene. Dette fører til at mengden og sammensetningen av ressursene kommer i fokus, og ikke prisene.

Effektivitetsberegninger kan enten være output- eller inputorientert. Med inputorientering søker man å minimere innsatsfaktorene per produksjonsenhet, mens ved outputorientering søker man å maksimere produksjonen etter de ressurser som er tilgjengelige. I oppgavens beregninger vil outputorientering benyttes. Dette valget er basert på ren skjønn ettersom modellene beregner hvor mye helseforetak kan produsere/behandle pasienter basert på ressursmengden i motsetning til hvor mange ressurser som kan bespares gitt en fast produksjon. Valget påvirker ikke rangeringen av helseforetakene.

### 3.2. Kvalitet

For å kunne si noe om hvordan kvaliteten i helsevesenet har endret seg over tid, er det nødvendig å definere «kvalitet», og se nærmere på hvordan begrepet kan operasjonaliseres og måles. Dette for å forsikre seg om at målene som knyttes til analysen er valide og reliable.



I forhold til offentlige tjenester benyttes den internasjonale standardiseringsorganisasjonen ISO sitt kvalitetsbegrep i stor grad (NOU 2012:6). En betegnelse av kvalitet basert på ISO-standard er «i hvilken grad en samling av iboende egenskaper oppfyller krav (eller behov)», og en annen også basert på ISO er at kvalitet utgjør en «helhet av egenskaper og kjennetegn et produkt/tjeneste har, som vedrører dets evne til å tilfredsstille fastsatte krav eller behov som er antydnet» (NOU 2012:6, s.234).

### **Definisjon av kvalitet i helsesektoren**

Innen helsesektoren har eksperter slitt med å bli enige om én konsis, meningsfull og generell definisjon av kvalitet, ifølge «the New England Journal of Medicine» (Blumenthal, 1996). Det forklares med at definisjonene er for komplekse og varierer for mye. Populære definisjoner er blant andre «American Medical Association» (1986) sin, som sier at kvaliteten avgjøres av graden helsetjenester øker og/eller forbedrer livskvaliteten og lengden på pasienters liv. En annen, formulert av «the Institute of Medicine» (Lohr, 1990), sier at kvalitet i helsevesenet er graden av hvor mye helsetjenester for individer og populasjoner øker sjansen for ønskede helseeffekter, og er i tråd med dagens kunnskap. Denne definisjonen er mye sitert og har fått bred anerkjennelse i fagmiljøet (Chassin, 1998).

Sosial- og helsedirektoratet (2005) skriver i en rapport at tjenester av god kvalitet:

- er virkningsfulle
- er trygge og sikre
- involverer brukere og gir dem innflytelse
- er samordnet og preget av kontinuitet
- utnytter ressursene på en god måte
- er tilgjengelig og rettferdig fordelt

Campbell et al. (2000) har gjennomgått en rekke definisjoner, og foreslår et skille mellom generiske og sammensatte kvalitetsdefinisjoner. De generiske definisjonene fokuserer på et aspekt som skal gi et fullstendig mål av kvalitet. Det kan være om behandlingen har oppfylt mål og forventninger, om behandlingen har vært feilfri, formålstjenlig eller lignende.

Generiske definisjoner er mer generaliserbare enn spesifikke og sensitive, og dermed vanskelig å operasjonalisere. De sammensatte definisjonene, på den andre side, tolker kvalitet som et kompleks og flerdimensjonalt begrep. Kvalitet defineres da i forhold til individuelle

dimensjoner eller komponenter. Hver individuelle komponent av kvalitet gir bare et delvis bilde av kvalitet når man ser på det alene, men sammensatt gir komponentene en spesifikk, kompleks definisjon av kvalitet.

Campbell et al. (2000) foreslår at kvalitetsbegrepet kan splittes opp i to deler, som til sammen dekker alle dimensjonene av kvalitet: tilgjengelighet og effekt. Til grunn for oppdeling legger de til grunn to essensielle spørsmål: Får brukere den behandlingen de trenger? Og i så fall, fungerer behandlingen de mottar? Videre deles effekten inn i to elementer – effekten den kliniske behandlingen gir, og effekten den interpersonelle omsorgen gir – og det påpekes at begge elementene må relateres til behov. På den måte defineres kvalitet av hvilken grad en har tilgang til nødvendige helsetjenester, og hvor effektiv behandlingen er.

I lys av dette, kan man si at definisjonen av «The American Medical Association» (1986) er en generisk definisjon, som assosierer kvalitet med forbedring av to attributter. En åpenbar svakhet er at kvalitetsbegrepet delvis defineres ved bruk av ordet «kvalitet». «The institute of medicine» sin definisjon inkluderer en skala, dekker alle typer helsetjenester, retter seg mot både individer og populasjon og er målorientert. Definisjonen er generiske fordi den måler kvalitet med ett generelt attributt - helseeffekt. Samtidig er den mer kompleks enn de aller enkleste, siden det er «ønsket» helseeffekt som måles (Blumenthal, 1996). På denne måten blir brukerens preferanser inkludert i kvalitetsbegrepet. Kravet til helsetjenesten er dynamisk siden det krever at de leverer i tråd med tilgjengelig teknisk, medisinske og vitenskapelige kunnskap. Noe som er en positiv egenskap ved definisjonen.

Definisjonen gitt av Sosial- og helsedirektoratet er derimot en sammensatt definisjon. Direktoratet forsøker å bygge på samfunnets føringer, oppfylle lovverkets krav og gi best mulig tjeneste til brukeren ut fra et faglig perspektiv når de definerer kvalitet. Med det trekkes alle mulige komponenter inn, og definisjonen gis et mer konkret innhold. Kontinuitet og ressursbruk blant faktorene som finnes i denne, men i få andre definisjoner. Hvorvidt man skal inkludere ressursbruk i kvalitetsvurderingen kan være avgjørende. Donabedian (1988) drøfter dette, og sier man kan enten gå for en maksimert eller optimal tilpasning. Hvis man maksimerer kvalitet, tar man ikke høyde for ressursbruk, men søker å maksimere helseeffekt. Hvis man trekker inn kostnader, forsøker man å finne en optimal tilpasning mellom helseeffekt og kostnad.

Donabedian (1988) poengterer at det finnes flere forskjellige, legitime definisjoner, avhengende av hvor i helsesystemet man er lokalisert og hvilket ansvar man har. Dette

utdypes med at medisinere, brukeren og resten av samfunnet, vil ha forskjellige perspektiv og preferanser i forhold til en kvalitetsvurdering.

### **Struktur, prosess og resultat**

En fellesnevner for kvalitetsvurderinger, både innen helsesektor og andre bransjer, er en inndeling i tre hovedfaser. Det er nødvendig å skille mellom helsetjenestens struktur, den behandlingen som leveres og konsekvensen av behandlingen (Campbell et al., 2000). Brook et al. (1996) sier spesifikt at kvalitet bør evalueres på basis av struktur, prosess og resultat. En inndeling også Helsedirektoratet bruker når de publiserer en rekke kvalitetsindikatorer. Vi skal se nærmere på hva en slik inndelingen innebærer.

Struktur beskriver de ytre forutsetninger i bred forstand. Det kan være regelverk og planer som definerer styring og danner grunnlaget for helsesystemets organisatoriske faktorer. Ressurser i form av bygninger, personell og utstyr. Faktorer som påvirker tilgjengelighet, som åpningstider, geografiske hindringer og tilgjengelig kompetanse, inngår også i struktur.

Prosess handler om virksomhetens indre aktiviteter, og relateres til aktiviteter som inngår i pasientforløpet, hovedsakelig selve behandlingen, men også aktiviteter før og etter. To prosesser trekkes ofte fram: kliniske og mellommenneskelig forhold (Campbell et al., 2000).

Kliniske forhold refererer til den faktiske medisinske behandlingen som gis. Mens mellommenneskelige forhold beskriver interaksjon mellom helsepersonell og pasient.

Mellommenneskelige forhold handler om evnen til å gjøre pasienten trygg, vise forståelse, og ha empati med pasienter. Også samspillet mellom personellet er viktig.

Resultat er konsekvensene av behandlingen, og viser således hva som skjer, eller ikke skjer, etter prosessen. Helsegevinst og tilfredshet er to hovedgrupper av resultat. Disse kan naturlig oppsplittes i spesifikke mål, eksempelvis overlevelsesgrad eller prosentvis forbedring. Ofte kan resultat være vanskelig å stadfeste, grunnet lang tidshorisont eller påvirkning fra ytre faktorer. (ebit.)



**Figur 1: Struktur-, prosess- og resultatkvalitet**

Med det har man tre stadier som bør tas med i en helhetlig vurdering av helsesystemet. Å kun måle resultat, kan virke logisk i først omgang, siden det gjenspeiler oppnådd helseeffekt. Det

er viktig å forstå hvordan resultatet påvirkes direkte og indirekte av struktur og prosess. Det er effekten av strukturen og prosessen som gjenspeiles i resultatet. Resultatet er fortsatt i fokus om man gjør oppdelingen, men man oppnår en bedre måling av de underliggende årsaker. Hvis kvalitetskriterier basert på struktur- eller prosessdata skal være troverdige, må det kunne vises at de målte kriteriene påvirker resultatet (Brook et al., 1996; Chassin et al., 1998). Det er derfor nødvendig å konstatere at dette er tilfelle før man benytter variabler fra hver del som et kvalitetsmål (Donabedian, 1988).

Kritikken mot måling av kvalitet gjennom prosessvariabler, går ut på at disse variablene ikke nødvendigvis gir utslag på resultatet i form av økt helseeffekt, og dermed tilsvarer en målefeil. På den andre siden kritiseres bruken av resultatvariabler for å være for lite detaljert. Årsaken til resultatet kan være faktorer som ligger utenfor helsevesenets kontroll. Dermed kan man ved å måle resultatet direkte gi helsevesenet et annet kvalitetsnivå enn de fortjener. Prosessdata gir vanligvis mer sensitive kvalitetsmål enn resultat, fordi feil underveis i behandlingsforløpet nødvendigvis ikke slår ut negativt på resultatet, men blir fanget opp av prosessvariablene. (Brook et al., 1996) En fordel med prosessdata, er at informasjon om den medisinske behandlingen registreres fortløpende, og er dermed tilgjengelig, mens resultatet kan ta mye tid før det kan fastslås. Resultatet gjenspeiler alle bidrag fra helsevesenet, noe som er en fordel, men klarer ikke skille ut hva som ikke bidrar positivt, som er en ulempe. (Donabedian, 1988) Det virker å være bred enighet om at verken struktur-, prosess- eller resultatmål er overlegne alene, de har forskjellige egenskaper som gjør de passende for forskjellige formål. Det er derfor nødvendig å drøfte nøye hvilke variabler man skal måle, og velge de med høyeste forklaringsgrad.

### **Måling av kvalitet**

For å måle kvalitet, må struktur-, prosess- og resultatkriteriene tolkes på en måte slik at de kan kvantifiseres. Feiltolking av kriteriene vil føre til invalide tall, som fører til feilmåling av kvalitet. Ideelt bør kriteriene tolkes på grunnlag av vitenskapelig innsikt, de mest pålitelige meningene eller praktikanter praksis. Hvor graden av validitet vil være høyest med førstnevnte og synkende. Både Brook et al. (1996) og Donabedian (1988) skiller mellom eksplisitt og implisitt måling, og videre drøfting bygger på deres resonnement.

Implisitt vurdering av kvalitet skjer ved at fagfolk får informasjon om en sak, og bruker egen kunnskap og erfaring til å bedømme nivået. I en eksplisitt vurdering derimot, blir enkeltsaker

vurdert mot detaljerte kriterier som er satt til forskjellige kategorier av saker. Enten ved evaluering av bestemte prosesskriterier, eller ved bruk av eksplisitte *a priori* kriterier for å bestemme hvorvidt det observerte resultatet er i tråd med det mest sannsynlige utfallet. Begge metodene har sine fordeler og ulemper, men disse går i hovedsak ut på praktiske utfordringer.

Like viktig som valg av metode, er hvilke informasjonskilder den anvendes på. Her er det potensielt mange feilkilder som kan gjøre målingen ugyldig. En viktig kilde for prosess- og resultatdata vil være medisinske journaler, og lignende register fra sykehus. Dokumentene kan ofte være ufullstendige, feil eller forskjellig utfylt. Informasjon bearbeides i flere ledd før de havner i et offentlig register, slik at feilkildene kan være mange. Donabedian (1988) anbefaler bruk av medisinske rapporter som kilde, men understreker at informasjon bør verifiseres med å sammenligne stikkprøver av sluttrapport med kilden. Uansett er det ikke all informasjon som registreres, eksempelvis uteblir mellommenneskelige forholdet, mens det medisinske dekkes godt.

Informasjon i rapporter kan suppleres med intervju eller spørreundersøkelser med personell og pasient. Intervjuobjektene vil gjerne ha et forskjellig perspektiv når de foretar en vurdering, og vil ofte være mer nyttig for å bekrefte registrert data, enn som en objektiv måling. I en kvalitetsvurdering vil pasienter fort kunne bli påvirket emosjonelt av resultatet, mens fagpersonell kan ønske å fremstå best mulig. Derfor vil observasjoner gjort av en tredjeperson gi bedre data. Et annet poeng er at brukerne tvilsomt har tilstrekkelig kunnskap til å vurdere kvalitetskomponentene. I hvert fall ikke de tekniske komponentene.

### **Vår definisjon og forutsetninger**

Basert på den generelle definisjonen til ISO og den gjennomgåtte litteraturen, mener vi kvalitet i helsevesenet bør vurderes av helheten av de egenskaper og kjennetegn tjenestene deres har, som vedrører evnen til å tilfredsstillende fastsatte krav eller behov. Etter vår mening, bør kravene være at helsetjenester skal være tilgjengelig ved behov, ikke medføre unødvendig ubehag og øke sjansen for positiv helseeffekt.

I dette studiet undersøkes somatisk spesialist helsetjeneste, som utgjør helseforetak, og ikke hele helsevesenet. Derfor må vi forsøke å isolere hva helseforetakene bidrar med, hva deres oppgave er, og hva som kan forventes av dem. Elementer som er utenfor foretakenes makt, bør de ikke vurderes på bakgrunn av.

Således vil vi legge vekt på tekniske attributter, som retter fokus mot hovedmålet, å forbedre helsetilstanden til hver enkelt individ etter behov. Samtidig som opplevelsen til individet ikke bør neglisjeres. Prosessvariabler vil på denne måten være mest naturlig og relevant for kvalitetsmålet. Struktur og resultatvariabler, vil også tas med i vurderinger, gitt at disse har et klar kausalitet med det arbeidet helseforetaket gjør.

### 3.3. Litteratursammendrag

Først ser vi på studier som har operasjonalisert effektivitetsbegrepet i norsk helsesektor. Deretter gjennomgår vi tre studier som ser på relasjon mellom effektivitet og kvalitet i helsevesen. Metodisk bruker alle dataomhyllingsanalyser.

#### **Effektiviseringsstudier på norsk helsevesen**

Biørn, et al. (2009) undersøker 47 norske sykehus i perioden 1992 til 2001. To outputer er valgt til effektivitetsmålene. Første output er antall *utskrevne pasienter* justert etter DRG. Andre output er *polikliniske pasienter* målt etter det honoraret staten betaler for de forskjellige besøkene. Til beregning av teknisk effektivitet har de valgt ut tre forskjellige inputer. Første input er *årsverk leger*. Andre input er *andre årsverk*, og dette målet inkluderer alle årsverk utenom leger. Målet er argumentert for ettersom en mer detaljert oppdeling ikke endret resultatene. Det tredje målet er *medisinske kostnader* målt i kroner. Til beregningen av kostnadseffektiviteten har de kun brukt *driftskostnader* som input. For å gjøre årene sammenlignbare, er alle priser i beregningene deflatert til år 2000.

Hagen et al. (2001) kartlegger i en rapport effektivitetsutviklingen ved norske sykehus fra 1992 til 1999. SAMDATA-rapporter benyttes som datagrunnlag. Utvalget er begrenset til sykehus som bruker DRG-systemet. Til den tekniske analysen brukes det to typer inputs: årsverk og medisinske forbruksvarer. Årsverkene splittes inn i kategoriene leger og annen arbeidskraft. Bruken av *leger* alene som input argumenteres ved at leger er sentrale i mye av aktiviteten i sykehusene. *Annen arbeidskraft* inkluderer sykepleiere, annet pleiepersonell, administrativt personell og annet behandlingpersonell. Disse er aggregert ettersom analysen har få observasjoner, samtidig som flere inputs ikke har betydelig effekt på analysen. Den tredje inputen er *medisiner og medisinsk forbruksvarer*. Til kostnadsanalysen brukes kun en input: brutto driftsutgifter. *Brutto driftsutgifter* brukes som representant for årsverk til leger og

annen arbeidskraft, samt medisinske forbruksvarer. I både modellen for teknisk effektivitet og kostnadseffektivitet brukes de samme output: *justerte innleggelser og polikliniske refusjoner*.

SINTEF undersøker effektivitetsutviklingen for somatisk spesialisthelsetjeneste i perioden 2003 til 2006 (Pedersen, 2007). Som input benyttes *legeårsverk, årsverk øvrig personell, og driftskostnader pasientbehandling*. Hvor driftskostnadene inneholder kostnader knyttet til DRG-aktivitet og polikliniske aktivitet, og på den måte måler kostnader til output som er *DRG-poeng og polikliniske refusjoner*. Dataen er hentet fra Norsk pasientregister, men gjennomgår en rekke korreksjoner for å gjøres sammenlignbare over analyseperioden. Tilskudd til forskning og utdanning, er vanskelig å finne tall på, og korrigeres ikke for så godt som ønsket, påpekes det av forfatterne. Sammenslåinger av sykehus i perioden, tas høyde for ved å slå de sammen for hele analyseperioden, og således skape «virtuelle sykehus» som fremstår som én enhet for hele perioden.

Kittelsen et al. (2007) forklarer en analyse utført på norske sykehus gjort i sammenheng med en større nordisk studie. Ressursbruken er målt med én input som er *driftskostnader*. Kostnader til kapital, forsknings og utdannings ekskluderes. Inputen består av tre faktorer – leger, sykepleiere og andre kostnader – som er summert opp. Seks forskjellige variabler benyttes som output. *DRG-poeng* disaggregert til fem variabler for å gjøre aktivitetsmålet sammenlignbart på tvers av landene. I tillegg kommer *antall polikliniske konsultasjoner*.

### **Studier om kvalitet og effektivitet**

Morey et al. (1992) forsøker å finne en relasjon mellom kvalitet og kostnader i helsevesenet. Et datasett fra 1983, bestående av data fra 300 amerikanske sykehus, analyseres. Det hypotetiske spørsmålet blir hvordan effekt det har på kostnadene til et sykehus hvis man endrer på kvaliteten i sykehuset gitt alle andre faktorer forblir de samme (sykehusets størrelse, antall senger, omgivelsene, etc.).

For å måle kvaliteten beregner de en indeks etter sannsynligheten for dødsfall opp mot faktisk observerte dødsfall. I tillegg til kvalitetsindeksen basert på dødsfall vurderes hvert sykehus etter serviceoutput, sykehusomgivelsene og kostnadsnivået til sykehuset. Serviceoutput er målt etter en rekke variabler som beskriver antall av forskjellige behandlinger.

Sykehusomgivelsene beskrives igjennom faktorer som antall sengeposter, case-miks

alvorlighetsindeksen og forskjeller i lønningsratene. Kostnadsnivået til sykehuset beskrives med de totale kostnader delt inn i underseksjoner.

Videre gjøres det en regresjon hvor endring i kvalitetsindeksen settes opp mot kostnadene til hvert sykehus. Dette resulterer i en trade-off hvor en økning i kostnadene vil medføre en reduksjon i dødsraten relativt til den statistisk beregnede dødsraten. Samtidig som kostnadseffektiviteten står negativt relatert til kvaliteten, har kostnadseffektivitet en positiv relasjon til samtlige av serviceoutputfaktorene nevnt ovenfor. En rekke begrensninger påpekes, hovedsakelig at studiet kan ha et for ensidig kvalitetsmål, for generelle variabler for sykehusene, feilmåling pga. selvrapporing av data, og et bias til kvalitetsmålet ettersom pasienter selv kan velge sykehus etter omdømme og tidligere statistikker.

Nayar & Ozcan (2008) har utført en studie på 53 ikke-føderale akuttsykehus i Virginia, USA for året 2003. To dataomhyllingsanalyser utføres – en kun for teknisk effektivitet; og en annen som inkluderer kvalitetsoutput. Formålet med dette er å undersøke to elementer: først hvordan resultatene påvirkes hvis kvalitet inkluderes, i andre omgang hvordan teknisk effektive sykehus yte med hensyn på kvalitet. I studier gjort på andre bransjer er det vist en viss trade-off mellom teknisk effektivitet og målt kvalitet, og forfatterne ønsket å se hvorvidt dette er tilfelle for sykehus.

Teknisk effektivitet måler output med antall justerte utskrivninger, totale polikliniske pasienter og et mål for arbeidsmengde utført av studenter; mens input består av mål for sykehusets størrelse, forbruksmateriale, antall årsverk og totale eiendeler. Kvalitet inkluderes i andre omgang med tre outputs som måler hvorvidt forskjellige pasientgrupper fikk nødvendig behandling innen anbefalt tid.

Sammenlikningen av analysene viser at de sykehus som scorer høyt på teknisk effektivitet også scorer høyt når kvalitetsindikatorne inkluderes i analysen. Dette indikerer til en høy korrelasjon mellom effektiviteten og kvaliteten i sykehussektoren. Samtidig er det sykehus med lav teknisk effektivitet som scorer høyt på kvalitetsindikatorne. Studiet konkluderer med at det ikke er en uunngåelig trade-off mellom kvalitet og effektivitet, og at ledelsen derfor ikke må være redd for å forbedre sin tekniske effektivitet. Bruk av sekundærdata, mangelfull datasett, og få analyseobjekter, trekkes fram som svakheter ved analysen. Videre kritiseres dataomhyllingsanalysen for å være sensitiv til tilfeldig støy i dataene.

Valdmanis et al. (2008) tar for seg samhandlingen mellom kvalitet, kostnader og effektiviteten i urbane amerikanske sykehus. Formålet er å finne hvorvidt sykehus kan øke sin



kvalitet ved å endre på miksen av forskjellige inputer. Studiet tar for seg 1377 urbane sykehus i USA. En metningsanalysen utføres som tillegg i dataomhyllingsanalysen som forutsetter at visse output er negative for sykehuset, og brukes til å måle negative kvalitetsvariabler. Det utføres også en slakkanalyse for å finne hvilke inputer som er ineffektive for hvert sykehus. Dette gir to dataomhyllingsanalyser, hvor en inkluderer kvalitetsindikatorene. Totalt gav dette mulighet til å se på relasjonene mellom ineffektive inputer, kvaliteten og produktiviteten til hvert sykehus.

Studiet ga få, inkonsistente resultater. Høy kvalitet kunne relateres til høye totalkostnader og tilgjengelighet til høyteknologiske tjenester. Sykehus med høy kvalitet hadde også høyere gjennomsnittlig enn sykehus med lavere kvalitet. Sykehus med høy kvalitet har også gjennomsnittlig høyere effektivitet enn sykehus med lav kvalitet. Kun tre prosent av ineffektiviteten skyldes dårlig kvalitet. En reallokering av ressurser kunne øke kvaliteten i sykehusene uten å medføre høyere kostnader.

## 4. Metode

Dette kapittelet gir en kort forklaring på studiets design og valg av metode, for så gi en inngående redegjørelse av anvendte verktøy og metoder.

Det utføres en longitudinell undersøkelse, nærmere bestemt en tidsserieundersøkelse basert på sekundærdata innsamlet over en fireårsperiode. Det brukes en forklarende forskningsdesign med empirisk tottrinnsanalyse for å løse problemstillingen. I tottrinnsanalysen brukes ikke-parametriske metoder som dataomhyllingsanalyser og Malmquist-produktivitetsindekser til å beregne kvalitets- og effektivitetsscorer i første trinn. I andre trinn brukes en gruppetest for å undersøke populasjonstilhørighet, en korrelasjonsanalyse for å belyse samvariasjon, og regresjonsanalyse til å undersøke årsakssammenhenger mellom kvalitet og effektivitet.

### 4.1. Produksjonsteknologi og distansefunksjoner

Distansefunksjoner er et nyttig verktøy for måling av produktivitet og effektivitet ved hjelp av vektorer for inputer og outputer. Konseptet ble utviklet av både Shephard og Malmquist i 1953, og blir i stor grad benyttet i våre beregninger. For å måle effektivitet og produktivitet krever distansefunksjonene observerte input- og output-kombinasjoner, som kan relateres til en optimal produksjonsteknologi. Distansefunksjonene kan beregnes som enten outputorienterte eller inputorienterte. I resten av redegjørelsen fokuseres det på outputorienterte distansefunksjoner, siden dette benyttes i våre beregninger. Videre følger en nødvendig forklaring av konseptet bak produksjonsteknologi og distansefunksjoner. Denne gjennomgangen baserer seg på framstillingen til Coelli et al. (2005).

En produksjonsteknologi bestående av flere inputs og outputs, kan beskrives gjennom teknologisetten  $S$ , der  $x$  er en vektor for input, og  $y$  er vektor for output:

$$S = \{(x, y): x \text{ kan produsere } y\} \quad (3)$$

Produksjonsteknologien definert av  $S$ , kan likedan defineres ved å bruke output-settet  $P(x)$ , som representerer et sett av alle output-vektorer,  $y$ , som kan produseres ved bruk av inputvektor  $x$ , gitt produksjonsteknologien  $S$ :

$$P(x) = \{y: x \text{ kan produsere } y\} = \{y : (x, y) \in S\} \quad (4)$$

For hver  $x$ , må  $P(x)$  inneha følgende egenskaper for å være gyldig:

- i)  $0 \in P(x)$ : For et hvilket som helst sett med inputer kan man produsere null outputer.
- ii) Ved alle inputer lik null kan man ikke produsere noen outputer.
- iii) Det er mulig å produsere mindre  $y$  med gitte inputer  $P(x)$ . Hvis  $y^* \leq y$  og  $y \in P(x)$ , så vil også  $y^* \in P(x)$ .
- iv) Det er mulig å produsere samme mengde av output  $y$  gitt flere inputer. Hvis  $y$  produseres med  $x$ , kan også  $y$  produseres med  $x^*$  gitt at  $x^* \geq x$ .
- v)  $P(x)$  er lukket.
- vi)  $P(x)$  er et begrenset område. Uendelig mange  $y$  kan ikke bli produsert ved et gitt sett av  $x$ .
- vii)  $P(x)$  er konveks.

En outputorientert distansefunksjon finner den minste relative distansen mellom en outputvektor og teknologisetten  $P(x)$  gitt en fast inputvektor. Gitt at  $\delta$  er skaleringsfaktoren som skal minimeres, kan distansefunksjonen skrives som:

$$d_o(x, y) = \min \left\{ \delta : \left( \frac{y}{\delta} \right) \in P(x) \right\} \quad (5)$$

Distansefunksjonene har seks egenskaper.

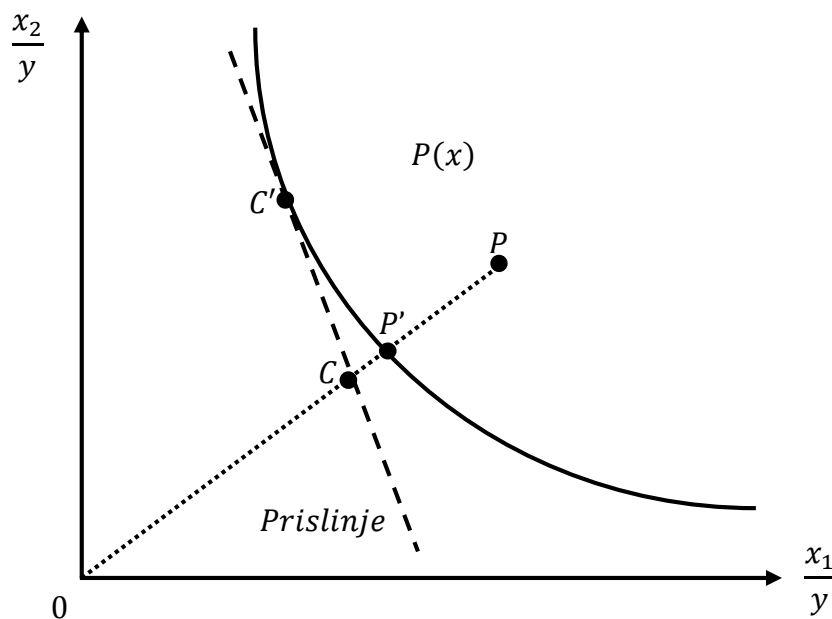
- i)  $d_o(x, 0) = 0$  i alle tilfeller hvor  $x \geq 0$ .
- ii)  $d_o(x, y)$  er ikke synkende for  $y$  og ikke økende for  $x$ .
- iii)  $d_o(x, y)$  er lineært homogen i  $y$ .
- iv)  $d_o(x, y)$  er kvasi-konveks i  $x$  og konveks i  $y$ .
- v) Hvis  $y \in P(x)$  så er  $d_o(x, y) \leq 1$ .
- vi) Hvis outputmengden  $y$  hører hjemme i fronten av produksjonsmulighetene, så er  $d_o(x, y) = 1$ .

Et viktig forhold mellom inputorienterte og outputorienterte distansefunksjoner er at  $d_o(x, y) = \frac{1}{d_i(x, y)}$ . Dette medfører at et punkt får samme effektivitetsscore uansett hvilken orientering som brukes.

## 4.2. Farrells effektivitetsbegreper

Som tidligere nevnt har Farrell (1957) splittet effektivitetsbegrepet inn i *teknisk effektivitet*, *priseffektivitet* og *total effektivitet*.

I et tenkt tilfelle produserer en DMU<sup>1</sup> én output basert på to inputs  $x_1$  og  $x_2$ . Outputen kan produseres igjennom forskjellige kombinasjoner av  $x_1$  og  $x_2$ . Forutsetter vi at skalautbytte er konstant for uansett størrelse på produksjonen, kan kombinasjonene presenteres i et diagram som vist under:



**Figur 2: Dekomponering av effektivitet**

Punktet P viser en gitt kombinasjon av  $x_1$  og  $x_2$  som er brukt til å produsere outputen  $y$ . Dette er presentert igjennom produktivitetmålene  $x_2/y$  og  $x_1/y$ . Forutsetter vi også at vi kjenner til de kombinasjonene av  $x_1$  og  $x_2$  som gir optimalt effektivitet, kan vi konstruere fronten til produksjonsmulighetsområdet. Gitt at  $x_1$  og  $x_2$  er faste, må DMU-en øke produksjonen av  $y$  for å komme nærmere optimal front. Den optimale tilpasningen for DMU-en, gitt kombinasjonen av  $x_1$  og  $x_2$ , representeres av punktet  $P'$ . DMU-en får to vektorer  $0P$  og  $0P'$  som representerer distansen mellom faktisk produksjon, optimal produksjon og origo. Forholdet mellom disse distansene utgjør Farrells (1957) *tekniske effektivitet*:

<sup>1</sup> DMU (Decision Making Unit) er en benevnelse for objektene som produserer output basert på gitte input. Videre i oppgaven vil DMU-er brukes om helseforetak for enkelthets skyld.

$$\text{Teknisk effektivitet (TE)} = \frac{OP'}{OP} \quad (6)$$

Vi kan nå tenke oss at inputene  $x_1$  og  $x_2$  har prisene  $p_1$  og  $p_2$ . Prisforholdet kan konstruere ei prislinje med helningsgraden  $-p_2/p_1$  som tangerer optimal front. Krysningepunktet mellom  $P0$  og prislinjen representerer den priseffektive tilpasningen (punkt  $C$ ) ved samme kombinasjon av inputs som tidligere. Dette punktet er utenfor produksjonsmulighetsområdet, og dermed uopnåelig. Distansen kan benyttes til å regne ut det Farrell kaller priseffektivitet:

$$\text{Priseffektivitet (PE)} = \frac{OC}{OP'} \quad (7)$$

Skal DMU-en være priseffektiv, må den tilpasse seg i punkt  $C'$  som er innenfor produksjonsmulighetsområdet. I denne tilpasningen er man både teknisk effektiv og priseffektiv, en tilpasning Farrell kaller total effektivitet. Punktet representer optimal kostnadsbruk, og kalles i ettertid kostnadseffektivitet.

$$\text{Kostnadseffektivitet (CE)} = \frac{OC}{OP} = TE * PE \quad (8)$$

Formlene til Farrell (1957) har mange likhetstrekk med distansefunksjonene, men formelene er ikke helt like. Vi kan se at:

$$TE_o = \frac{1}{d_o(x, y)} \quad (9)$$

#### 4.2.1. Teknisk effektivitet i praktisk beregning

Ved veldig kompliserte produksjonsprosesser kan beregningen av produksjonsmulighetskurven være svært vanskelig. Kompliserte produksjonsprosesser skaper usikkerhet. Farrell (1957) foreslår to mulige løsninger. Ingeniører kan konstruere en teoretisk funksjon for kurven ved å observere beste praksis. Dette gir rom for menneskelige feil. Jo mer kompleks en produksjonsprosess er, desto høyere er sannsynligheten for feil i den teoretiske funksjonen. Andre løsning består i å la beste praksis selv representere funksjonen. Optimal front konstrueres etter de beste observerte kombinasjoner av input og output slik at de beste observerte DMU-er settes til 100 % teknisk effektivitet. Svakheten med denne løsningen er at kurven ikke representerer de optimale kombinasjoner, men i stedet de optimale observerte

kombinasjoner. Til de videre beregninger brukes løsning nummer to. Farrell (1957) presenterer den algebraiske definisjonen slik:

$$x_{k1} = \gamma x_{i1} + \delta x_{j1} \quad (10)$$

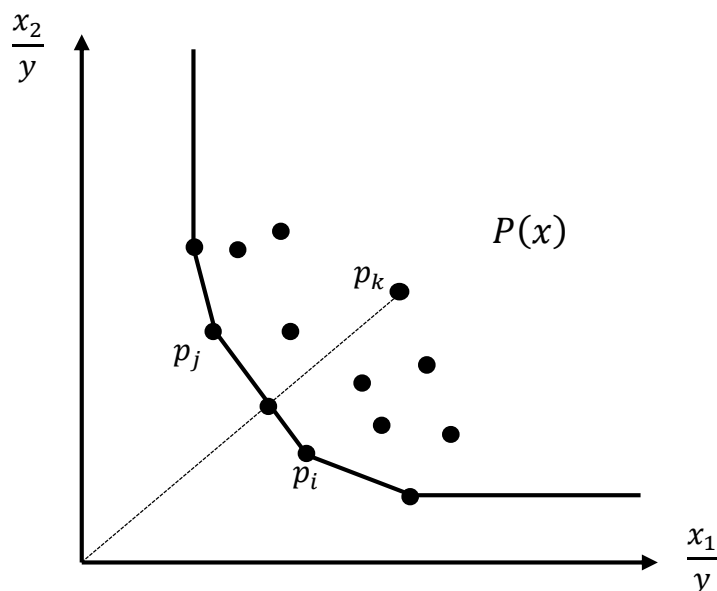
$$x_{k2} = \gamma x_{i2} + \delta x_{j2} \quad (11)$$

Formelen viser punktene  $P_k, P_i$  og  $P_j$  i produksjonsdiagrammet. Her er punktene  $P_j$  og  $P_i$  en del av optimal front hvis  $\gamma_{ijk} + \delta_{ijk} \geq 1$  for alle  $P_k$  i produksjonsdiagrammet.  $P_k$  representerer et hvilken som helst punkt i diagrammet innenfor produksjonsmulighetsområdet  $P(x)$ .  $\gamma$  og  $\delta$  representerer vektene for input  $x_1$  og  $x_2$ .

Noen forutsetninger gjør seg gjeldende. Produksjonsmulighetsområdet er alltid positivt. Alle observasjoner ligger mellom origo og front for produksjonsmulighetsområdet. Når forutsetningene holder kan den tekniske effektiviteten for punkt  $P_k$  uttrykkes som:

$$TE_{P_k} = \frac{1}{\gamma_{ijk} + \delta_{ijk}} \geq 1 \quad (12)$$

Dette fordi vektene  $\gamma_{ijk} + \delta_{ijk} = 1$  bare når punktet  $P_k$  ligger på kurven, og er dermed 100 % teknisk effektiv. Hvis  $P_k$  er ineffektiv vil  $0 < \gamma_{ijk} + \delta_{ijk} < 1$ .



**Figur 3: Farrells effektivitet**

Beregningene av teknisk effektivitet i kombinasjon med distansefunksjoner og optimale fronter gir grunnlaget for beregning av effektivitetsscorene i en dataomhyllingsanalyse.

### 4.3. Dataomhyllingsanalyse

Dataomhyllingsanalysen (DEA<sup>2</sup>) fikk sitt navn av Charnes, Cooper og Rhode (1978).

Analysemetoden baserer seg på Farrells matematiske definisjon av teknisk effektivitet, og viser til hvordan dette praktisk kan gjøres gjennom lineær programmering.

Dataomhyllingsanalyser er ikke-parametriske av natur og har hovedsakelig to styrker som er viktige for oppgaven. For det første behøves ingen avansert produksjon for helseforetakene. I stedet benchmarkes alle helseforetakene opp mot hverandre og får en score ut ifra dette. Hvorvidt et helseforetak får en god score blir derfor avgjort av både helseforetakets egen prestasjon og prestasjonen til de beste helseforetakene. For det andre kan DEA håndtere mange inputer og outputer med forskjellige måleenheter hvor forhåndsbestemte priser og vektorer ikke er et krav. Vi anser derfor dataomhyllingsanalyser som godt egnet for beregning av kvalitets- og effektivitetsscorer på så komplekse DMU-er som helseforetakene er. DEA er allerede anvendt mye i helsesektor (se kapittel 3.3). Dataomhyllingsanalysen kan beskrives igjennom en primal-modell og en dual-modell, kalt multiplikatormodellen og omhyllingsmodellen.

*Multiplikatormodellen* søker å finne et mål for effektiviteten til DMU-ene igjennom å sette forskjellige vektorer på outputene og inputene til de forskjellige DMU-ene. Vektene bestemmes på en slik måte at effektiviteten maksimeres for hver DMU, slik at alle får et mål så nærme 1 som mulig – lik Farrells definisjon av effektivitet. Alle DMU-er med en effektivitet under 1 vil derfor bli regnet som ineffektiv. Multiplikatormodellen med hensyn på output kan skrives slik:

$$\min_{u,v} \sum_{i=1}^m v_i x_{i0} = e_0 \quad (13)$$

Når:

$$\sum_{r=1}^s u_r y_{r0} = 1$$

$$\sum_{i=1}^m v_{ij} x_{ij} - \sum_{r=1}^s u_{rj} y_{rj} \geq 0$$

---

<sup>2</sup> Data Envelopment Analysis

$$\text{Her er: } j = 1, \dots, n, \quad u_r, v_i \geq 0, \quad r = 1, \dots, s; \quad i = 1, \dots, m.$$

I denne framstillingen av multiplikatormodellen er  $e_o$  effektivitetstallet som skal minimeres, og beregnes igjennom den angitte formelen som begrenses til å ha minimumsverdien 1. Her er  $y$  selve verdien til et spesifikt output og  $x$  er verdien til et spesifikt input. DMU-ene representeres av bokstaven  $j$  rangert fra nummer 1 til  $n$ . Outputene er rangert fra  $r = 1$  til  $s$ , og inputene rangeres fra  $i = 1$  til  $m$ . Effektivitetstallet som beregnes igjennom denne modellen blir et relativt tall etter hvor den bestemte DMU-en ligger i forhold til de beste DMU-ene. Dette betyr at selv med optimaliserte vektorer er ikke DMU-ene sikret å få effektivitetstallet 1. For det andre betyr dette at effektivitetstallet vil endre seg hvis DMU-en selv endrer produksjonsnivå, men også hvis de beste endrer produksjonsnivå.  $y_{rj}$  er output verdien til output  $r$  for DMU  $j$  med tilhørende vekt  $u_{rj}$  til output  $r$  for DMU  $j$ . På samme vis er  $x_{ij}$  inputverdien for input  $i$  for DMU  $j$  med vekten  $v_{ij}$  til input  $i$  for DMU  $j$ .

*Omhyllingsmodellen* er en omformulering av multiplikatormodellen hvor de forskjellige vektene til inputene og outputene ikke er like synlige framstilt. Charnes et al. (1978) presenterer omhyllingsmodellen som en omformulering av multiplikatormodellen.

Omhyllingsmodellen med hensyn på output kan skrives som:

$$\max z_0 \tag{14}$$

Når:

$$z_0 y_{r0} \leq \sum_{j=1}^n y_{rj} \lambda_j$$

$$x_{ij} \geq \sum_{j=1}^n x_{ij} \lambda_j$$

$$\lambda_j \geq 0$$

$$r = 1, \dots, s, \quad i = 1, \dots, m; \quad j = 1, \dots, n.$$

Som i multiplikatormodellen vil indikatorene  $y$  og  $x$  representere de observerte output og input mens  $j$ ,  $r$  og  $i$  viser til hvilken DMU og hvilke output og input som skal settes inn med  $s$  antall output,  $m$  antall input og  $n$  antall DMU-er. Modellenes forskjell ligger i bruken av



vektoren  $\lambda$  i stedet for vektene til hver enkelt input og output. Vektoren har som funksjon å skape en imaginær observasjon basert på de beste DMU-enes produksjonsmiks som skal representere det beste mulige produksjonspunktet for den DMU-en man skal beregne effektivitetstallet på. Dette kan skrives som:

$$z_o = \frac{y_{r0}^*}{y_{r0}} \quad (15)$$

Når:

$$y_{r0}^* = \sum_{j=1}^n \lambda_j y_{rj}$$

Her vil  $y_{r0}^*$  representere det teoretisk optimale punktet for DMU 0 relativt til de optimale DMU-enes output. Denne beregningen gjøres for alle input og output til alle DMU-ene inntil man har et fullstendig matematisk punkt for optimal produksjon for DMU 0 basert på de andre DMU-ene i datasettet. Her er det viktig å notere seg at punktet kun beregnes på bakgrunn av de mest effektive DMU-ene, ettersom alle ineffektive DMU-er vil få sin  $\lambda_j$  satt lik null. Størrelsen til  $\lambda_j$  vil være størst for de DMU-ene det relative punktet  $y_{r0}^*$  ligger nærmest. Hvis, for eksempel, punktet  $y_{r0}^*$  ligger nøyaktig oppå en effektiv DMU, vil denne DMU-ens  $\lambda_j$  settes lik 1 og alle andre DMU-ers  $\lambda_j$  settes lik null. Vi kan også bruke vektoren  $\lambda_j$  til å beregne det optimale punktets skala i forhold til de effektive DMU-ene. Vi regner da ut  $\sum \lambda_j$ . Hvis dette tallet ligger over 1, vil den optimale produksjonen for DMU 0 ha høyere skala enn de effektive DMU-ene. Ligger tallet under 1, vil den optimale produksjonen for DMU 0 ligge under de effektive DMU-enes skala.

Multiplikatormodellen og omhyllingsmodellen ovenfor kalles også CCR-modellene etter sine skapere, og begge modellene forutsetter konstant skalautbytte (CRS<sup>3</sup>). Dette medfører at samme mulige effektivitet forventes uansett størrelsen på produksjonen til en DMU.

Banker, Charnes og Cooper (1984) har videreutviklet CCR-modellen ved å innføre variabelt skalautbytte (VRS<sup>4</sup>). I praksis medfører variabelt skalautbytte at flere DMU-er kan oppnå effektivitetsscoren 1 gjennom å representere beste praksis for høy-skala DMU-er og lav-skala DMU-er. I CCR-modellen forutsettes det at:

---

<sup>3</sup> Constant Return to Scale

<sup>4</sup> Variable Return to Scale

$$\lambda_j \geq 0 \quad (16)$$

$$j = 1, \dots, n$$

Dette medfører at variabelen  $\lambda$  ikke er begrenset på annen måte en for negative tall. Resultatet er at effektiviteten beregnes uten hensyn til skala, og kun de beste DMU-ene får en høy effektivitetsscore. For variabelt skalautbytte, også kalt BCC-modellen etter sine skapere, inkluderes det et ekstra krav:

$$\sum_{j=1}^n \lambda_j = 1 \quad (17)$$

$$j = 1, \dots, n$$

Denne begrensningen gjør at alle DMU-er vil få en effektivitetsscore relativ til de beste DMU-ene av samme skala og ikke de beste DMU-ene totalt. I praksis vil bruken av BCC-modellen føre til flere effektive og generelt høyere effektivitetstall for flere av DMU-ene enn ved en vanlig CCR-modell. Ved å beregne begge modellene kan man også finne en beregnet skalaeffektivitet ved å se på effektivitetsscoren fra CCR-modellen opp mot effektivitetsscoren til BCC-modellen for en spesifikk DMU.

Differansen mellom effektivitetsscoren forutsatt konstant og variabelt skalautbytte indikerer om en DMU er betydelig større eller mindre i forhold til de mest effektive DMU-ene gitt konstant skalautbytte. Denne differansen kalles skalaeffektivitet, og beregnes igjennom formelen:

$$\text{Skalaeffektivitet} = \frac{\text{Effektivitet}_{CRS}}{\text{Effektivitet}_{VRS}} \quad (18)$$

Dette kan skrives om ved bruken av distansefunksjoner, og vi får:

$$\text{Skalaeffektivitet} = \frac{d_o^{VRS}(x, y)}{d_o^{CRS}(x, y)} \quad (19)$$

Ved en skalaeffektivitet tilnærmet lik 1 vil DMU-en ha et skalautbytte av samme størrelse som de mest effektive DMU-ene gitt konstant skalautbytte. Hvis skalaeffektiviteten er lav indikerer dette at DMU-en har en for stor eller for liten produksjon sett opp imot de beste DMU-ene gitt konstant skalautbytte.

### 4.3.1. Supereffektivitet

I dataomhyllingsanalysemodellene vist ovenfor skiller vi mellom effektive og ineffektive DMU-er. Vi ser på hvorvidt effektivitetsscoren ligger på 1 eller om effektivitetsscoren er lavere enn 1. Dette medfører at vi ikke kan rangere de effektive DMU-ene på noen måte ettersom alle har samme effektivitetsscore. En løsning på dette problemet, introdusert av Andersen og Petersen (1993), er å innføre supereffektivitet i modellen.

Supereffektivitetsscoren beregnes nærmest helt likt med effektivitetsscoren til omhyllingsmodellen. Forskjellen ligger i at når en DMU får sin effektivitetsscore beregnet, så er ikke DMU-en tillatt å være en del av den effektive fronten. Resultatet av dette er at DMU-en har en mulighet til å legge seg mellom den effektive fronten og origo ved inputorientering, og dermed få en effektivitetsscore over 1. Et viktig poeng er at alle DMU-er som allerede er beregnet å være ineffektiv vil få samme effektivitetsscore ved beregning av supereffektivitet som ved bruken av omhyllingsmodellen. De eneste DMU-ene som påvirkes er DMU-ene med effektivitetsscoren 1, ettersom disse tillates å få en score over 1. Det lineære problemet til supereffektiviteten kan skrives som følgende:

$$\max z_0 \tag{20}$$

Når:

$$z_0 y_{r0} \leq \sum_{\substack{j=1 \\ j \neq 0}}^n y_{rj} \lambda_j$$

$$x_{ij} \geq \sum_{\substack{j=1 \\ j \neq 0}}^n x_{ij} \lambda_j$$

$$\lambda_j \geq 0$$

$$r = 1, \dots, s, \quad i = 1, \dots, m; \quad j = 1, \dots, n.$$

Ovenfor ser vi en CRS-modell for supereffektivitet. Eneste forskjell fra tidligere vist omhyllingsmodell er forutsetningen  $j \neq 0$  når vi beregner effektiviteten til DMU 0. I praksis betyr dette at DMU-enes effektivitetsscore ikke kan beregnes med referanse i seg selv. I en

vanlig omhyllingsmodell er optimal effektivitetsscore lik 1 ettersom de mest effektive DMU-ene kan bruke seg selv som referanse. I supereffektivitetsmodellen beregnes alltid den effektive fronten uten den DMU-en som skal beregne sin effektivitetsscore. I praksis kan dette medføre at tidligere ineffektive DMU-er blir en del av den effektive fronten når effektivitetsscoren til en effektiv DMU skal beregnes.

Selv om har valgt å bruke supereffektivitet i vår studie, er det viktig å poengtere tre svakheter med denne metoden, som er forklart av Adler et al. (2002). Første svakhet ligger i å kalle det en rangering av de supereffektive DMU-ene når hver DMU har forskjellige vektorer for sine inputer og outputer. For det andre risikerer man at spesialiserte enheter tildeles en svært høy supereffektivitetsscore fordi den effektive fronten blir liggende langt bak DMU-ens produksjonsmiks av inputer og outputer. Siste svakhet ligger i bruken av supereffektivitet ved noen typer omhyllingsmodeller utenom ved konstant skalautbytte. I praksis vil man ikke kunne beregne supereffektiviteten til DMU-er som ligger i ytterkantene av fronten når man ikke har konstant skalautbytte. Dette fordi disse ikke blir liggende mellom den effektive fronten og origo når de ikke er en del av den effektive fronten.

#### 4.3.2. Dysons fallgruver

Når dataomhyllingsanalysen beregner sine effektivitetsscorer, gjøres dette under visse forutsetninger. Dyson et al. (2001) redegjort for forskjellige typer feil som bør unngås ved bruk av dataomhyllingsanalyser, og kommer med forslag for å unngå eller løse utfordringer knyttet til forutsetningene. Slike feil og fallgruver er det viktig at vi er bevisste på når vi velger utvalg og input og output til modellene.

*Homogene utvalg* er en nødvendighet om man skal utføre en dataomhyllingsanalyse på disse. Alle DMU-er må være like igjennom å produsere samme type produkt ved bruken av samme type inputer og ha en – relativt – lik produksjonsprosess. I heterogene utvalg risikerer man at noen enheter urettmessig stemples som mindre eller mer effektive enn andre.

*Homogene omgivelser til DMU-ene er også et krav* ved sammenlikning. Dette fordi ulike omgivelser kan gi urettferdige fordeler eller ulemper til de forskjellige DMU-ene. Resultatet blir at noen får urettmessig høy eller lav effektivitetsscore på grunn av omgivelsesfaktorer i stedet for prestasjonen internt i DMU-en.

*Skalaøkonomi* kan også spille inn på homogeniteten til forskjellige DMU-er. Størrelsen på produksjonen til en DMU kan påvirke effektiviteten igjennom fordeler og ulemper ved å ha høy eller lav skala på produksjonen. Resultatet kan også her bli urettferdig sammenlikning på grunnlag av noen enheter som har fordeler andre enheter ikke har.

*Antallet input og output* er viktige for kvaliteten til analysen. Selv om antallet input og output skal dekke spekteret av produkter og ressurser brukt i prosessen, kan en udiskriminerende inkludering av alle inputer og outputer medføre et urealistisk antall effektive enheter. Dette fordi inkludering av hver ny input eller output gir muligheten for DMU-er å vise seg mer effektiv på disse områdene. Det er anbefalt å ha  $n$  antall DMU-er, hvor  $n = 2 * r * i$ , og hvor  $r * i$  er produktet av antall inputer og outputer.

*Korrelasjon mellom forskjellige input* kan oppstå. I slike tilfeller kan man ofte bruke en input som representant for en gruppe med korrelerte inputer. Det er likevel en mulighet for at denne nedkorting har en høy effekt på effektivitetstallene. I så fall bør man sjekke resultatene med og uten alle faktorene, for å se om forskjellene i effektivitetsscorene er betydelige.

*Miks av mengde og indeksverdier* bør ikke oppstå i samme dataomhyllingsanalyse. Den lineære programmeringen kan ikke skille mellom mengdeverdier som øker igjennom skala og indeksverdier som ikke er skalasensitive. I en slik miks risikerer man at DMU-er med høy skala måles til urettferdig høy effektivitet igjennom indekstallene ettersom disse ikke øker med skala.

*Prosjenter og normaliserte data* kan brukes i en dataomhyllingsanalyse, men kun hvis alle inputer og outputer er av denne typen verdier. Hvis ikke risikerer man samme type feilmåling av effektivitetsscoren som ved miksing av mengde og indeksverdier.

*Kvalitative data* kan medføre problemer ved innføring i en dataomhyllingsanalyse. For det første er kvantifiseringen av kvalitative data en prosess hvor man vanskelig kan få noe mer enn ordinale verdier. Dermed vet man ikke hvor mye bedre eller verre en DMU gjør det opp mot en annen selv med høyere eller lavere ordinale verdier. For det andre er kvalitative data ofte subjektive og inkluderer en form for bias som kan forskyve dataverdiene og gi gale effektivitetsscorer.

*Uønskede output og input* omhandler blant annet negative output som forurensning og produksjonshemmende input som antall konkurrenter. Inkluderes disse i dataomhyllingsanalysen uendret vil resultatene bli feil, ettersom forurensning beregnes som et

positivt produkt DMU-en bør maksimere, og høyere produksjon betyr økning i antall konkurrenter. Som huskeregel kan man sette alt som skal maksimeres som output og alt som skal minimeres som input. Eventuelt kan man snu brøken slik at forurensningsoutputen blir 1/forurensning. På denne måten kan man få et positivt forurensningsmål hvor verdien blir større jo mindre DMU-en forurenses. Dette kan også gjøres for input, hvor en negativ input kan snus om i stedet for å settes som output.

*Eksogene og begrensede faktorer* innebærer faktorer i omgivelsene, fysiske begrensninger eller begrensninger igjennom lov. Slike faktorer bør ikke brukes som input eller output. Eksogene faktorer vil ikke en DMU ha noen kontroll over. Begrensede faktorer vil ikke overgå et visst punkt, noe som vil gi feilmålinger mellom store og små DMU-er av samme type som miksing av indekser, prosenter og mengdemål.

Dyson et al. (2001) viser også til fallgruver med hensyn til vektrestriksjoner. Disse er ikke relevante for denne studien ettersom vi ikke tar i bruk dataomhyllingsanalysenes vektrestriksjoner i beregning av effektivitetsscorer. Dette betyr at DMU-ene risikerer å få spesielle vekter som ikke nødvendigvis realistiske.

#### 4.3.3. Kritikk: substituerbare og ikke-substituerbare inputer og outputer

Dataomhyllingsanalyser forutsetter normalt at inputer og outputer er substituerbare med hverandre, selv om i praksis så er ikke dette helt riktig. Man risikerer derfor feilberegning av effektivitetsscorer til DMU-ene om inputer og outputer som ikke er substituerbare forutsettes å være det.

Omhyllingsmodellens effektive front formes etter forutsetningen om at inputer er i stand til å erstatte hverandre. Ved å øke en input, kan man redusere en annen input og fortsatt produsere like mye i output. I praksis er ikke dette alltid helt riktig, og Barnum et al. (2009) har undersøkt effektene på effektivitetsscorene til forskjellige sykehus ved bruk av forskjellige modeller hvor noen tar hensyn til ikke-substituerbare inputer og outputer. I stedet for bruk av ordinære omhyllingsmodeller, foreslår de å bruke ikke-substituerbare versjoner av DEA for å beregne effektivitetsscoren.

En modell for effektivitetsberegninger gitt ikke-substituerbare inputer og outputer er «fixed proportion additive»-modellen (FPA). I stedet for å beregne effektivitetsscoren via en segmentert isokvantfront, beregnes effektivitetsscoren etter ett effektivt punkt som

representerer hele fronten. Dette medfører en L-formet front i stedet for den delvis konvekse linjen i vanlige omhyllingsmodeller med substituerbare inputer og outputer. Modellen kan skrives som:

$$FPA_k = \sum_{m=1}^M \left[ \left| \left( \frac{x_{km}}{y_{k1}} \right) - \min_j \left( \frac{k_{jm}}{y_{j1}} \right) \right| \right] \quad (21)$$

$j = 1, 2, \dots, J, m = 1, 2, \dots, M$

Her er DMU k sin effektivitetsscore beregnet igjennom forholdet til produktiviteten for hver input opp mot den laveste produktiviteten for inputen i hele datasettet. Dette gjøres for alle inputer til man har fått sin FPA-indeks for DMU-en. Denne modellen gir mål med absolutte tall i stedet for en indeks, og det anbefales derfor å bruke en indeksvariant av formelen. For å finne en effektivitetsscore brukes derfor:

$$ef f_{kmn} = \left[ \frac{\left( \frac{y_{kn}}{x_{km}} \right)}{\max_j \left( \frac{y_{jn}}{x_{jm}} \right)} \right] \quad (22)$$

#### 4.3.4. Kritikk: Inputer og outputer i spesialisthelsetjeneste

I en studie av Magnussen (1996) kommer det fram at dataomhyllingsanalyser gjort på sykehus er veldig sensitive til hvilke outputer som brukes. I studien satte han inputene til leger, sykepleiere, annet personell og senger. Outputene varierte mellom antall pasienter, pasientdager og pasientinnleggelser og polikliniske besøk. Hvilke sykehus som kom ut som de mest effektive varierte etter hvilke output som ble brukt. Variasjon i output hadde også effekt på skalaegenskapene til de forskjellige sykehusene i undersøkelsen. Dette medfører at tolkningen av dataomhyllingsanalysene må sees i lys av de outputene og inputene som brukes, og ikke som en fullstendig analyse av den helhetlige effektiviteten og skalaen til de forskjellige sykehusene.

#### 4.4. Malmquist-produktivtetsindeks

Malmquist-produktivtetsindeks er et ikke-parametrisk og relativt mål for å sammenlikne to enheter i forskjellige økonomier. Indeksen kan brukes til både sammenlikning av to forskjellige enheter og som et mål for endring over tid ved å se på produktiviteten mellom to forskjellige tidspunkt. I likhet med DEA krever ikke Malmquist-produktivtetsindeks fastsatte

priser til de forskjellige inputer og outputer for å beregne produktivitetsendringen. Indeksen baserer seg også på distansefunksjoner, og er i stand til å håndtere mange inputer og outputer. Produktivitetsindeksen baserer seg på både distansefunksjoner og produksjonsfunksjoner. I vår studie vil vi bruke prinsippet til Malmquist produktivitetsindeks ved å se på den relative produktiviteten mellom en DMU og den optimale fronten for høyest produktivitet. Färe et al. (1994) viser til Caves et al. (1982) sine Malmquist-produktivitetsindekser:

$$M_{CCD}^t = \frac{d_o^t(x_o^{t+1}, y_o^{t+1})}{d_o^t(x_o^t, y_o^t)} \quad (23)$$

$$M_{CCD}^{t+1} = \frac{d_o^{t+1}(x_o^{t+1}, y_o^{t+1})}{d_o^{t+1}(x_o^t, y_o^t)} \quad (24)$$

Her representerer  $t$  og  $t + 1$  de forskjellige tidspunktene til DMU-en.  $M_o^t$  og  $M_o^{t+1}$  estimerer begge Malmquist-produktivitetsindeks, men de skiller seg ved å bruke optimal front ved år  $t$  og ved år  $t + 1$ . Den brukte Malmquist-produktivitetsindeksen vi bruker er av det geometriske gjennomsnittet av disse to produktivitetsindeksene.

$$M_o(x^{t+1}, y^{t+1}, x^t, y^t) = [M_{CCD}^t * M_{CCD}^{t+1}]^{\frac{1}{2}} \quad (25)$$

Fullt utskrevet med distansefunksjoner blir dette:

$$M_o(x^{t+1}, y^{t+1}, x^t, y^t) = \left[ \left( \frac{d_o^t(x_o^{t+1}, y_o^{t+1})}{d_o^t(x_o^t, y_o^t)} \right) \left( \frac{d_o^{t+1}(x_o^{t+1}, y_o^{t+1})}{d_o^{t+1}(x_o^t, y_o^t)} \right) \right]^{\frac{1}{2}} \quad (26)$$

Malmquist-indeksen er en relativ indeks med likevekten på tallet 1. Det vil si at en Malmquist-indeks lik 1 medfører ingen forskjell i produktivitet mellom tidspunkt  $t$  og  $t + 1$ . For outputorienterte Malmquist-indekser vil alle tall over 1 indikere en økning i produktiviteten, mens alle tall under 1 vil indikere en nedgang i produktivitet.

#### 4.4.1. Effektivitetsendring og teknologisk endring

Färe et al. (1994) har videre dekomponert Malmquist-produktivitetsindeksen.

Dekomponeringen resulterer i en todeling av indeksen som gir mulighet for dypere forståelse av produktivitetsendringene. Dette kan skrives som:



$$M_o(x^{t+1}, y^{t+1}, x^t, y^t) = \frac{d_o^{t+1}(x_o^{t+1}, y_o^{t+1})}{d_o^t(x_o^t, y_o^t)} * \left[ \left( \frac{d_o^t(x_o^{t+1}, y_o^{t+1})}{d_o^{t+1}(x_o^{t+1}, y_o^{t+1})} \right) \left( \frac{d_o^t(x_o^t, y_o^t)}{d_o^{t+1}(x_o^t, y_o^t)} \right) \right]^{\frac{1}{2}} \quad (27)$$

Brøken utenfor kvadratrotten representerer endringen i effektivitet. Dette målet viser hvorvidt DMU 0 har økt sin effektivitet mellom år  $t$  og år  $t + 1$  med hensyn til sin posisjon til produksjonsfunksjonen. Brøkene i kvadratrotten representerer den teknologiske endringen. Dette tallet vil fortelle hvordan produksjonsfunksjonen selv har endret seg mellom periode  $t$  og periode  $t + 1$  fra ståstedet til DMU 0. Dekomponeringen viser hvordan DMU 0 endrer sin effektivitet i forhold til produksjonslinjen og endringen til produksjonslinjen alene, hvor produktet av disse gir den faktiske produktivitetsendringen til DMU 0. Med bakgrunn i dette kan vi se hvorvidt produktivitetsendringen til en DMU skyldes at DMU-en selv har blitt bedre eller verre, eller om det skyldes endring i teknologi på et større plan.

$$\text{Effektivitetsendring} = \frac{d_o^{t+1}(x_o^{t+1}, y_o^{t+1})}{d_o^t(x_o^t, y_o^t)} \quad (28)$$

$$\text{Teknologisk endring} = \left[ \left( \frac{d_o^t(x_o^{t+1}, y_o^{t+1})}{d_o^{t+1}(x_o^{t+1}, y_o^{t+1})} \right) \left( \frac{d_o^t(x_o^t, y_o^t)}{d_o^{t+1}(x_o^t, y_o^t)} \right) \right]^{\frac{1}{2}} \quad (29)$$

Verdiene for effektivitetsendringen og teknologisk endring kan tolkes på samme måte som Malmquist-indeksen. Verdien 1 forteller oss at DMU-en ikke har endret sin effektivitet i forhold til produksjonslinjen. En verdi over 1 viser til positiv effektivitetsendring, som vil si at DMU-en har nærmet seg den optimale produktiviteten, og en verdi under 1 viser til negativ effektivitetsendring. Hvis teknologisk endring settes til 1, betyr dette at optimal produktivitet ikke har endret seg fra DMU-ens ståsted fra tid  $t$  til  $t + 1$ . Hvis teknologisk endring beregnes til over 1, har det optimale produksjonsnivået forbedret seg over tid. En verdi under 1 viser til nedgang i optimal produktivitet.

#### 4.4.2. Malmquist med variert skalautbytte

Ray & Desli (1997) har videreutviklet Malmquist-indeksen, og implementert forutsetningen for variert skalautbytte. Resultatet er en dekomponering av indeksen som viser til både

effektivitetsendring, endring i front for variabelt skalautbytte og endring i skalaeffektivitet.

Malmquist-indeksen kan skrives som:

$$M_o = \text{Effektivitetsendring} * \text{Teknologisk endring} * \text{Skalaeffektivitet} \quad (30)$$

Her er:

$$\text{Effektivitetsendring} = \frac{d_{vrs}^{t+1}(x^{t+1}, y^{t+1})}{d_{vrs}^t(x^t, y^t)} \quad (31)$$

$$\text{Teknologisk endring} = \left[ \frac{d_{vrs}^t(x^t, y^t)}{d_{vrs}^{t+1}(x^t, y^t)} * \frac{d_{vrs}^t(x^{t+1}, y^{t+1})}{d_{vrs}^{t+1}(x^{t+1}, y^{t+1})} \right]^{\frac{1}{2}} \quad (32)$$

$$\text{Skalaeffektivitetsendring} = \left[ \frac{SE^t(x^{t+1}, y^{t+1})}{SE^t(x^t, y^t)} * \frac{SE^{t+1}(x^{t+1}, y^{t+1})}{SE^{t+1}(x^t, y^t)} \right]^{\frac{1}{2}} \quad (33)$$

Skrevet med distansefunksjoner:

$$\text{Skalaeffektivitetsendring} = \left[ \frac{\frac{d_{crs}^t(x^{t+1}, y^{t+1})}{d_{vrs}^t(x^{t+1}, y^{t+1})}}{\frac{d_{crs}^t(x^t, y^t)}{d_{vrs}^t(x^t, y^t)}} * \frac{\frac{d_{crs}^{t+1}(x^{t+1}, y^{t+1})}{d_{vrs}^{t+1}(x^{t+1}, y^{t+1})}}{\frac{d_{crs}^{t+1}(x^t, y^t)}{d_{vrs}^{t+1}(x^t, y^t)}} \right]^{\frac{1}{2}} \quad (34)$$

De dekomponerte enhetene ovenfor viser distansefunksjoner som baserer seg på VRS-fronten i stedet for CRS-fronten. Ettersom denne dekomponeringen bygger på front for variabelt i stedet for konstant skalautbytte, får vi en tredje indeks for skalaeffektivitet som er den relative endringen mellom distansefunksjonene basert på konstant og variabelt skalautbytte.

#### 4.4.3. Malmquist-indeksen ved endringer i skala

Grifell-Tatjé & Lovell (1995) viser til en svakhet ved Malmquist-indeksen når det gjelder større endringer i skala. Igjennom flere forsøk har de funnet at Malmquist-indeksen ikke beregner endring i produktivitet riktig når DMU-en endrer sin skala betydelig. En økning i skala medfører at beregnet Malmquist-indeks kan vise en betydelig høyere produktivitetsendring enn det som er virkelig, og i verste fall skjule en produktivitetsnedgang

mellom perioden  $t$  og  $t + 1$ . Denne svakheten vil ikke påvirke våre resultater i betydelig grad ettersom ingen av DMU-ene i vår studie går igjennom store endringer i skala.

#### 4.5. Statistiske metoder

Videre presenteres tre statistiske metoder som benyttes for å lete etter statistiske relasjoner.

##### 4.5.1. Kolmogorov-Smirnov

«Kolmogorov-Smirnov»-testen er en ikke-parametrisk test for å sammenligne to grupper. Banker (1993) nevner denne testen som god i forhold til sammenlikning av dataomhyllingsanalyser. Testen forsøker å finne hvorvidt to datasett kommer fra samme populasjon, men behøver ingen forutsetninger for dataenes spredninger. Lilliefors (1967) viser til formelen for «Kolmogorov-Smirnov»-testen:

$$D = \max_x |F^*(X) - S_N(X)| \quad (35)$$

Antall observasjoner er lik  $N$ .  $S_N(X)$  er funksjonen for utvalgets kumulative distribusjon, og  $F^*(X)$  er funksjonen for den kumulative normale distribusjonen.  $D$  er da den maksimale differansen i distribusjon mellom funksjonene. Nullhypotesen forkastes hvis  $D$  overstiger en gitt kritisk verdi. Den kritiske verdien kommer man fram til igjennom en Monte Carlo kalkulasjon.

##### 4.5.2. Korrelasjonsanalyse

Korrelasjonsanalyse benyttes for å analysere forholdet mellom variabler. Det er i mange sammenhenger interessant å vite noe om styrken og retningen på den lineære avhengigheten mellom to variabler. Den mest brukte indikatoren for korrelasjons er Pearsons produkt-moment korrelasjonskoeffisient. For variablene  $x$  og  $y$ , kan koeffisienten beregnes som følgende for et utvalg (Lind et al., 2010):

$$R = \frac{\sum(x - \bar{x})(y - \bar{y})}{(n - 1)s_x s_y} \quad (36)$$

Hvor  $n$  er antall observasjoner, og standardavviket til hvert variabelsett representeres av  $s_x$  og  $s_y$ . Pearsons korrelasjonskoeffisienten,  $R$ , kan variere mellom perfekt positiv korrelasjon,  $+1,0$ , og perfekt negativ korrelasjon,  $-1,0$ . En koeffisient nær null viser at det ikke er noen

korrelasjon. Korrelasjonen forteller bare noe om styrken til forholdet mellom variablene, og sier ikke om variasjonen i den ene er grunn til variasjon i den andre (Lind et al., 2010).

#### 4.5.3. Regresjonsanalyse

Regresjonsanalyse er en statistisk teknikk som undersøker sammenhengen mellom variabler. Teknikken benyttes for å lage prognoser og undersøke årsakssammenhenger. Metoden definerer en variabel som den avhengige som blir forklart av variasjonen i en annen variabel. Slik at man får en lineær sammenheng mellom variablene. I en enkel regresjonsmodell har man to variabler:  $y$ , den avhengige variabelen, og  $x$ , den uavhengige, eller forklarende, variabelen. I en multipl regressjonsanalyse tilføres ytterligere uavhengige variabler til modellen, på denne måten vil analysen kunne forklare mer av variasjon i den avhengige variabelen. Bak regresjonsanalysen ligger en matematisk ligning som beskriver sammenhengen mellom den avhengige og uavhengige variabler. En generell multipl regressjonsmodell, er den avhengige variabel  $y$  relatert til en rekke forklarende variabler  $x_1, x_2, \dots, x_k$  gjennom en lineær ligning som kan uttrykkes på følgende måte (Hill et al., 2012):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (37)$$

Parameteren  $\beta_0$  representerer konstantleddet til regresjonen, og tilsvare verdien av  $y$  når alle  $x$ -ene er lik null, eller gjennomsnittet av  $y$ . Regresjonskoeffisientene  $\beta_1, \beta_2, \dots, \beta_k$  korresponderer med de avhengige variablene  $x_1, x_2, \dots, x_k$ . En bestemt koeffisient, eksempelvis  $\beta_k$ , forteller hvor mye  $y$  endres når  $x_k$  endres med én enhet, og de andre  $x$ -ene holdes konstante. Den uforklarte variansen, forskjellen mellom regresjonslinjen og faktiske observasjoner, såkalt målefeil eller feilledd, representeres av parameteren  $\varepsilon$ . Antall avhengige variabler gis av  $k$ .

Modellen benytter seg av det minste kvadraters metode (OLS), som er den vanligste formen for lineær regresjon. McDonald (2009) argumenterer for at OLS er en god metode å bruke i sammenheng med en tottrinnsanalyse hvor dataomhyllingsanalyser er i første trinn. Metoden forsøker å trekke en rett linje gjennom de faktiske observasjonene. Linjen beste tilpasning beregnes ved å minimere summen av de kvadrerte avvikene mellom observerte og beregnede verdier av den uavhengige variabel  $y$ . Minste kvadrats metode finner slik verdien for hver regresjonskoeffisientene som minimerer den uforklarte variansen.

Bogetoft & Otto (2011) kritiserer bruk av OLS på DEA-resultater. Dette fordi DEA-resultatene begrenses til intervallet  $[0,1]$ , noe som kan hindre en normalfordeling av data. Denne begrensningen unngår vi i vår studie ved å bruke supereffektivitet.

Den multiple regresjonsmodellen har en rekke forutsetninger som må oppfylles. Fire forutsetninger må holde for restleddet i regresjonsfunksjonen. Siden hver observasjon av den avhengige variabelen avhenger av det tilfeldige feilleddet  $\varepsilon$ , medfører det at  $y$  også er en tilfeldig variabel. Dermed følger  $y$  de samme statistiske egenskapene som  $\varepsilon$ . I tillegg til forutsetningene til feilleddet og avhengig variabel, er det to forutsetninger for de forklarende variablene. Forutsetningene er som følgende (Hill et al., 2012; Lind et al., 2010):

- 1)  $\varepsilon \sim N(0, \sigma^2)$ . Feilleddet  $\varepsilon$  er normalfordelt.
- 2)  $E(\varepsilon) = 0$ . Hvert feilledd har en sannsynlighetsfordeling med gjennomsnitt lik null.
- 3)  $var(\varepsilon) = \sigma^2$ . Homoskedastisitet: feilleddet har lik varians for alle nivå av  $x$ .
- 4)  $cov(\varepsilon_i, \varepsilon_j) = 0$ . Kovariansen mellom to tilfeldige feilledd er lik null.
- 5) De uavhengige variabler korrelerer ikke.
- 6) Forholdet mellom den avhengige og de uavhengige variablene må være lineært.

Gauss-Markov teoremet sier at hvis forutsetning 1-5 holder, gir bruk av det minste kvadraters metode de beste estimatene av parameterne. De ønskede egenskapene oppsummeres populært som «the best linear unbiased estimators» (BLUE).

### **Tolkning av resultatene**

Forklaringsgraden til regresjonen,  $R^2$ , måler hvor stor andel av den totale variasjonen modellen forklarer. I en multippel regresjonsanalyse, vil den vise hvor mye de uavhengige variablene, til sammen, forklarer den avhengige variabelens variasjon. I en enkel regresjonsmodell finnes det bare en regresjonskoeffisient, denne kan enkelt kvadreres for å finne forklaringsgraden  $R^2$ . I multippel regresjonsanalyse vil den beregnes etter følgende matematiske sammenheng (Groebner et al., 2008):

$$R^2 = \frac{SSR}{SST} \quad (38)$$

SST<sup>5</sup> representerer total variasjon, uavhengig av gruppetilhørighet. Mens SSR<sup>6</sup> er variasjon på grunn av gruppetilhørighet, og SSE<sup>7</sup> uforklart variasjon.  $R^2$  varierer mellom 1 og 0, og tolkes som et prosentvis uttrykk. Med en verdi på 1, forklarer modellen all variasjon i den avhengige variabelen, og omvendt for verdi på 0. Forklaringsgraden kan korrigeres for hvor mange variabler som er inkludert i modellen. Dette kan være ønskelig fordi flere variabler fører til at  $R^2$  øker, selv om de nye variablene ikke bidrar til å forklare variansen. Ved å justere for dette, får man alltid et lavere, mer presist mål. Den justerte forklaringsgraden benevnes videre som  $R_A^2$  («adjusted»  $R^2$ ). Tolkningen er den samme for begge.

Hill et al. (2012) advarer mot å evaluere kvaliteten til regresjonsmodellen basert på forklaringsgraden. Det er viktig å vurdere omfanget til estimatene, statistisk signifikans, estimatenes presisjon og økonomisk signifikans. Eksempelvis estimerer modellen best rundt gjennomsnittet til utvalgsdataen, ekstrapolering av resultatene til ekstreme verdier er generelt en dårlig idé. I første omgang vil vi sjekke om modellen er signifikant. Hvis den er signifikant, vil det være nødvendig å sjekke de individuelle variablenes signifikans, om standardavviket til modellens feilledd er for stort til å gi meningsfulle resultater, og om multikollinearitet<sup>8</sup> er et problem for modellen, før man trekker en konklusjon om årsakssammenheng ut av modellen (Groebner et al., 2008).

### Modellens signifikans

Lind et al. (2010) forklarer hvordan man kan teste de uavhengige variablene  $x_1, x_2, \dots, x_k$  sin evne til å predikere den avhengige variabel  $y$ . Dette gjør man ved å undersøke hvorvidt det er mulig at alle uavhengige variablene har regresjonskoeffisienter lik null. Dette gjøres ved å formulere og teste følgende hypoteser:

$$H_0 = \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1 = \text{minst en } \beta \text{ er ikke lik null}$$

Hvis null-hypotesen er sann, er alle regresjonskoeffisientene lik null, og kan dermed ikke benyttes til å predikere den avhengige variabelen. For å teste dette utføres en F-test. Hvis den

---

<sup>5</sup>  $SST = \text{Sum of squares total} = SSR - SSE$

<sup>6</sup>  $SSR = \text{Sum of squares regression} = \sum(\hat{y} - \bar{y})^2$

<sup>7</sup>  $SSE = \text{Sum of squares error} = \sum(y - \hat{y})^2$

<sup>8</sup> Multikollinearitet inntreffer i tilfeller hvor uavhengige variabler bidrar med redundant informasjon til modellen, og dermed medfører en ugunstig effekt på resultatene (Groebner et al., 2008).

beregnete F-verdien er mindre enn kritisk F-verdi, forkastes ikke nullhypotesen. Kritisk F-verdi avhenger av signifikansnivå og antall frihetsgrader. Med  $n$  enheter og  $k$  uavhengige variabler, beregnes F-verdi:

$$F = \frac{\frac{SSR}{k}}{\frac{SSE}{(n - k - 1)}} \quad (39)$$

Modellens signifikans kan også vurderes på grunnlag av signifikanssannsynligheten. Denne fremgangsmåten rapporterer sannsynligheten for å få et resultat som er minst like ekstremt som det observerte, antatt at nullhypotesen er sann. I metoden sammenlignes signifikanssannsynligheten, kalt p-verdi<sup>9</sup>, med signifikansnivået. Hvis p-verdien er mindre enn signifikansnivået, forkastes null-hypotesen. Hvis den er større enn signifikansnivået, forkastes ikke nullhypotesen. Hvis man benytter et 95% konfidensintervall, vil det si at hypotesen godtas som sann hvis det er minst 95% sannsynlig at  $H_0$  er korrekt. En p-verdi på under 0,05 vil dermed si at nullhypotesen forkastes. I tillegg til å basere beslutninger på p-verdien, er en god indikator på styrken til konklusjonen.

#### 4.6. Valg og testing av datasett

Til nå har vi sett på studiets analysemetoder for å besvare vår problemstilling. Den siste delen av kapittel 4 er dedikert til de framgangsmåter, metoder og analyser vi har gjennomgått for å sikre et så godt datasett som mulig. Selv om studiet praktisk kan utføres uten noen av disse, har vi valgt å inkludere dem ettersom de hjelper å sikre kvaliteten til variablene og dermed studiets resultater.

Datasettene som inkluderes i våre modeller må bestå to typer krav. Første type krav er de som stilles til input og output i modellene selv, og disse er allerede tatt opp tidligere i Dysons fallgruver (kapittel 4.3.2). Andre typen krav er forskningsbaserte, og består av at datasettene skal være både reliable og valide med hensyn på det de skal måle, dette tas hensyn til under utforming av hver enkel modell.

---

<sup>9</sup> Fra engelsk «probability value»

#### 4.6.1. Korrelasjonsanalyse og gruppetester

Fra Dyson et al. (2001) husker vi at valget av variabler – alle inputene og outputene – er viktige for kvaliteten til hver modell. Modellene må ha mange nok variabler til å utfylle det de skal måle, og samtidig være få nok til at ikke unødvendig mange får høy effektivitet. For å sikre så gode modeller som mulig har vi valgt å stegvis utføre tester på potensielle variabler og modeller.

Første steg er korrelasjonsanalyser på forskjellige variabler. Fra teori om produktivitet vet vi at en økning i input bør medføre en økning i output. Dette er en forutsetning for at inputen faktisk er en ressurs som brukes i produksjonen av output. Det bør derfor være en sterk positiv korrelasjon mellom output og input. Hvis en input har svak eller negativ korrelasjon med outputen, bidrar inputen sannsynligvis ikke med produksjonen av outputen og bør derfor ikke være med.

Korrelasjonsanalysen kan også brukes mellom de forskjellige inputene i en mulig modell. Sterk korrelasjon mellom to eller flere inputer kan signalisere at én av inputene kan representere de andre. Dette er ingen garanti, og bør undersøkes nærmere før en slik forkorting av modellen utføres. Resultatet av de utførte korrelasjonsanalysene er alle de outputene og inputene som potensielt kan brukes i våre videre analyser, samt variasjoner av de forskjellige modellene alt etter hvilke output og input som inkluderes.

Neste steg er å finne hvilke modeller som er optimale for vår videre analyse. Til dette har vi valgt å bruke hypotesetestene til Banker (1993). Testene er kompatible med dataomhyllingsanalysene og kan finne modeller med signifikant forskjellige resultater. Vi sammenlikner slik et detaljert datasett med mange variabler opp mot et datasett hvor en eller flere variabler er fjernet. Kan en forskjell bevises mellom datasettene, bruker vi det mest detaljerte datasettet.

Effektivitetstallene forutsettes å kunne fordele seg enten som en eksponentielt fordelt funksjon eller som en halv-normalfordelt funksjon. Derfor velger vi å kjøre to hypotesetester. Den første forutsetter eksponentialfordeling og den andre forutsetter halv-normalfordeling. Formlene for disse hypotesetestene kan skrives som følger:

$$T_{\text{Eksponential}} = \frac{\sum_{j \in M_1} \frac{(1 - z_j)}{m_1}}{\sum_{j \in M_2} \frac{(1 - z_j)}{m_2}}, \quad \text{Frihetsgrader } (m_1, m_2) \quad (40)$$



$$T_{Halv-normal} = \frac{\sum_{j \in M_1} \frac{(1 - z_j)^2}{m_1}}{\sum_{j \in M_2} \frac{(1 - z_j)^2}{m_2}}, \quad \text{Frihetsgrader } (2 * m_1, 2 * m_2) \quad (41)$$

$$j = 1, 2, \dots, m_1 \text{ for alle } j \in M_1, \quad j = 1, 2, \dots, m_2 \text{ for alle } j \in M_2$$

I begge disse testene har vi to datasett  $M_1$  og  $M_2$  hvor et datasett mangler eller har sammensatt noen variabler. Datasettene inneholder  $m_1$  og  $m_2$  antall DMU-er med effektivitetstallet  $z_j$  for hver DMU  $j$ .

Både eksponentialtesten og halv-normaltesten har nullhypotesen at  $M_1 = M_2$ . Hypotesen testes med en F-test, metoden for hypotesetesting gjennomgått i sammenheng med regresjonsanalyser. Hvis en testverdi  $T_* > F_{Kritisk}$  så er avvises nullhypotesen.

Selv om hypotesetestene ikke viser noen statistisk signifikant forskjell mellom gruppene, er ikke dette en garanti for en god modell. Etter å ha luket ut de dårligste modellene igjennom hypotesetester, vil de resterende testene evalueres etter resultatene de gir. Noen DMU-er kan få betydelig dårligere effektivitetsscorer ved reduksjon av en variabel uten at dette slår ut på hypotesetestene.

#### 4.6.2. Outlieranalyse

Et viktig moment i vår studie er kjennskap til de såkalte outlierne. En outlier er en DMU hvor kombinasjonen av input og output ligger langt unna verdiene til resten av datasettet. Dette kan forekomme på grunn av registreringsfeil som medfører mye høyere eller lavere verdier enn det som er riktig. En DMU med verdier svært forskjellig fra andre DMU-er i datasettet kan også være et tegn på at DMU-en ikke hører hjemme i datasettet, og at den ikke er sammenliknbar med resten av datasettet. I så fall er det eneste riktige å fjerne outlieren. Til slutt kan det hende at DMU-en bare er generelt bedre eller verre enn de andre DMU-ene. For å finne outlierne har vi valgt å bruke metoden til Andrews & Pregibon (1978) som er videreutviklet av Wilson (1993). Videreutviklingen innebærer at man kan inkludere flere enn en output i analysen. Vi kan ta utgangspunkt i en vanlig lineær modell som:

$$y = \beta * X + \varepsilon \quad (42)$$

Hvor  $y$  er vektoren til den avhengige variabelen mens  $X$  er en matrise av uavhengige variabler. Modellen er sammenlignbar med regresjonsformelen fra kapittel 4.3 med koeffisient  $\beta$  og støyvariabelen  $\varepsilon$ .

For å finne outlierne fjernes DMU-ene systematisk fra datasettet med den hensikt å redusere støyen i modellen så mye som mulig. Logikken ligger i at en DMU som skaper mye støy i modellen også har sterk påvirkning på modellen igjennom ekstreme verdier. Metoden er likevel ikke garantert å finne alle outlierne hvis det finnes grupper med outlierne som ligger nært hverandre. Ved å fjerne en outlier vil fortsatt støyen være høy ettersom de andre outlierne forblir i samme område som den fjernede DMU-en. Det finnes heller ingen bestemt grense for når en DMU skal karakteriseres som outlier. Wilson (1993) anbefaler at metoden brukes til å prioritere hvilke DMU-er som bør undersøkes nærmere.

Metoden tar ikke hensyn til om en DMU er effektiv eller ikke. Det vil si at modellen forsøker å finne DMU-ene som skiller seg fra gruppen i stedet for å finne DMU-ene med urimelige effektivitetstall. Dette betyr at DMU-er som kan vise seg å være outlierne ifølge metoden til Wilson (1993) og Andrews & Pregibon (1978) ikke nødvendigvis har betydning for effektivitetsscorene til andre DMU-er i datasettet.

Outlier-analysen har vi gjennomført ved hjelp av Wilson-metodikken. Resultatet fra analysen presenteres grafisk gjennom et plott og gjennom en matrise hvor de mest ekstreme DMU-ene fjernes. Prioriteringen av DMU-er skjer lettest igjennom å lese plottet og matrise som fremkommer av analysen.

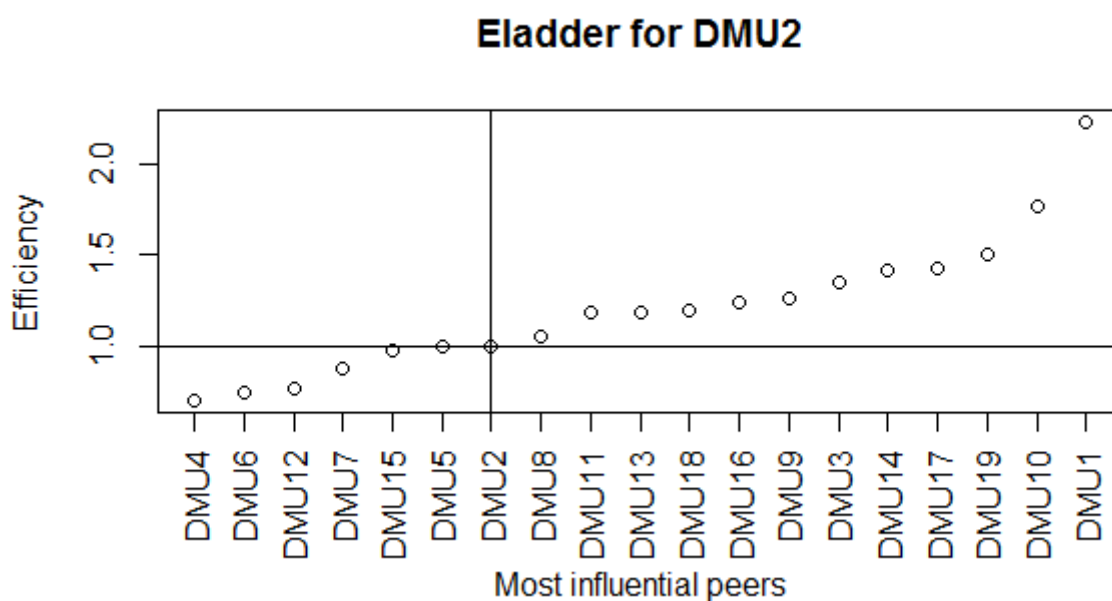
#### 4.6.3. Sensitivitetsanalyse

Det er ønskelig å undersøke hvor robuste effektivitetsestimaterne fra DEA-beregningene er. Forutsetningene for DEA-modellen er strenge, og oppfylles ikke alltid i virkeligheten. Brudd på forutsetningene fører til at estimatene forstyrres av feil relatert til utvalg, spesifisering og måling. Derfor undersøker man hvor sensitiv DEA-resultatene er til eventuelle feil. Dette kan hovedsakelig gjøres på to forskjellige måter; enten ved å måle avvik blant observasjonene ved hjelp av en modifisere LP-modellen (se Zhu, 1996), eller ved å ekskludere én og én observasjon fra datasettet (se Edvardsen, 2004).

Sistnevnte metodetype vil benyttes i denne oppgaven, nærmere bestemt effektivitetsstige-metoden presentert av Edvardsen (2004). Metoden går ut på å observere hvordan

effektivitetsestimater, til én utvalgt DMU, påvirkes når de mest innflytelsesrike DMU-ene ekskluderes sekvensielt fra analysen. Høy effektivitetsendringer antyder at estimatet er sensitivt for feilmålinger i referanseenheterne.

I praksis utgjør sensitivitetsanalysen en iterativ prosess med tre steg. Man velger ut én DMU som skal analyseres. I første steg utføres en dataomhyllingsanalyse med alle observasjoner. I andre ledd lokaliseres DMU-ens mest innflytelsesrike referanseenhet. I tredje steg utføres dataomhyllingsanalysen på nytt uten den referanseenheten som ble funnet i steg to. Andre og tredje ledd repeteres helt til analyseenheten selv er den mest effektive.



**Figur 4: Eksempel på Eladder**

Figuren viser resultatene av en effektivitetsstige-analyse. Plottene representerer analyseenhetens effektivitetsestimater for hver utførte dataomhyllingsanalyse. Hvis vi ser på eksempelet i figuren, så er den iterative prosessen gjentatt seks ganger før analyseenheten står igjen som den mest effektive. Stigningen virker å være lineær, noe som tilsier at effektivitetsendringen og sensitiviteten er relativt lav. Selv om stigen i seg selv gir et godt bilde på robustheten til analysen, er det mer praktisk å undersøke de nye effektivitetstallene som analysen presenterer. Ved å notere oss de nye effektivitetstallene for hvert steg i stigen, kan vi regne ut prosentvise effektivitetsendringer for analyseenheten. Dette vil bli gjort for alle DMU-ene, slik at vi kan få et bedre bilde av robustheten til hver enkelt enhet og av helheten.

## 5. Data

I løpet av dette kapittelet presenteres utvalget og de modeller og variabler som vi har valgt å bruke i vår studie. Først redegjøres det for utvalget med hensyn på de enhetene som vi ønsker å analysere. Deretter gjennomgår vi krav og tester som har resultert i sluttproduktet i form av de tre modellene for produksjon, kostnader og kvalitet.

### 5.1. Utvalg

Helse- og omsorgsdepartementet oppgir at det er 26 offentlige helseforetak i Norge per 01.01.2013. Helsedirektoratet arbeider med en liste på 27 helseforetak når de rapporterer data for somatisk spesialisthelsetjeneste for 2011. For å få et sammenlignbart utvalg, velger vi å inkludere kun offentlige foretak som leverer de samme funksjonene. Utgangspunktet for å inkludere enheter i analysen er at helseforetaket skal ha akuttfunksjon, medisinsk avdeling, kirurgisk avdeling og fødeavdeling, i tråd med Pedersen (2007) anbefaling og praksis. Dette ekskluderer en rekke spesialsykehus og private sykehus, og vi sitter igjen med 19 helseforetak. Listen over de 19 DMU-ene som utgjør utvalget vårt er listet opp i vedlegg 1.

Av de 19 helseforetakene i datasettet er fire sammenslått i perioden. Disse fire enhetene bestod før sammenslåingen av 11 forskjellige helseforetak. Sammenslåingene blir gjort ved forskjellige tidspunkter i analyseperioden, og er av forskjellig karakter. I noen tilfeller er det snakk om administrative endringer hvor strukturen beholdes, i andre tilfeller ny funksjonsfordeling og nedleggelse av sykehus. For å få samme enheter for hele perioden, har vi i datamaterialet slått sammen de aktuelle enhetene for hele perioden. Det betyr at vi har skapt «virtuelle» helseforetak for hvert enkelt år inntil sammenslåingen fant sted. Dette er normal praksis i effektivitetsanalyser av sykehus og medføre få svakheter så lenge variablene kan behandles additivt (Hagen & Kjekshus, 2003).

Tidsperioden som betraktes er fra 2008 til 2011. Bak dette valget ligger flere forklaringer. For det første trenger vi flere år for å kunne analysere utviklingen over tid. Helsedirektoratet tok over utarbeidelsen av SAMDATA fra SINTEF, og starter sin rapportering med tall fra 2008. Overtagelsen medførte en del endringen i rapporteringen og utarbeidelsen av tall. På bakgrunn av dette, velger vi å forholde oss til tall fra Helsedirektoratet. Analyseperioden strekker seg til år 2011, siden det er siste tilgjengelige SAMDATA-rapport på nåværende tidspunkt.

Totalt sikrer dette at de 19 DMU-ene i vår studie er homogene nok til å kunne sammenliknes igjennom en dataomhyllingsanalyse. Utvalget er ikke gjort med hensyn på homogene omgivelser. Dette medfører at resultatene av dataomhyllingsanalysene vil være påvirket av eksterne faktorer. Dette er ikke en betydelig svakhet i vår studie ettersom formålet ikke er å finne de faktiske ytelsesmålene, men i stedet hvordan kvalitet, effektivitet og kostnader faktisk korrelerer.

## 5.2. Validitet og reliabilitet

Ved valg av variabler til modellene er validiteten og reliabiliteten viktig for sluttresultatenes styrke (Blumberg et al., 2011). Validiteten og reliabiliteten, som forklart under, er derfor tatt sterkt hensyn til under vurderingen av inputene og outputene brukt i analysene.

Validiteten til studien er i hvilken grad målingene er konsistente med det man ønsker å måle. Validitetsbegrepet kan deles inn i validitet av innhold, kriteriebasert validitet og validitet av konsept. Validitet av innhold reflekterer hvorvidt bredden av målinger er stort nok til å skape et helhetlig bilde. Kriteriebasert validitet viser til hvor godt indikatorene korrelerer med det begrepet eller konseptet de skal måle. Validiteten av et konsept viser til hvor godt operasjonaliseringen av konseptet er opp mot teorien bak konseptet.

Reliabiliteten handler om hvor pålitelige, nøyaktige og konsistente resultatene er.

Reliabiliteten kan forklares igjennom stabilitet, ekvivalens og intern konsistens. Studien har konsistens hvis resultatene blir de samme hvis man utfører samme studie igjen på det samme utvalget med samme bruk av metoden. Ekvivalens viser til at måling av alle enheter i utvalget vurderes ut ifra et likt grunnlag. Intern konsistens omhandler hvorvidt indikatorene innenfor et begrep korrelerer eller ikke.

### 5.2.1. Krav til dataomhyllingsanalyser

For at resultatene fra dataomhyllingsanalysene skal være valide og reliable, må en rekke krav oppfylles. Fra kapittel 4.3.2 husker vi Dysons fallgruver. Disse er relevante for alle våre modeller. Hver modell må begrenses med hensyn på antall variabler. Gjøres ikke dette risikerer vi å få ubrukelige resultater med for mange effektive enheter. Samtidig må modellene ha nok variabler til å måle det de skal måle, og hver input må kunne settes i sammenheng med minst en output. Kvantumsbaserte mål må heller ikke mikses med indekser

og prosenter, slik at resultatene ikke konsekvent avgjøres av størrelsene på DMU-ene. Samtidig må vi vurdere om de potensielle variablene er såkalte uønskede variabler eller eksogene variabler. Av praktiske krav til variablene har vi satt at disse må være tilgjengelige for analyseperioden 2008 til 2011. Med hensyn til reliabiliteten, må variablene også være målt i forhold til en standard som gjør dataene sammenliknbare.

### 5.3. Sekundærdata

Til analysen har vi valgt å bruke sekundærdata i stedet for innhenting av primærdata. Dette fordi primærdata krever mye tid og ressurser for å få et komplett datasett, og vår studie er begrenset i både tidsrom og ressurser. Bruken av sekundærdata er likevel ikke optimalt. Dataene som brukes vil være innsamlet til andre formål enn dette studiet spesifikt, noe som betyr de ikke er helt valide i forhold til måten de blir brukt på. Samtidig har vi ingen kontroll på selve rapporteringen av tallmaterialet. Feilrapporteringer og bias i tallmaterialet kan ha skjevheter som i verste fall påvirker analyseresultatene i betydelig grad. For å sikre oss mot dette har vi gjort grundige vurderinger av alle kilder, innsamlingsmetodene og de mulige variablene. Vi har derfor bare valgt ut variabler som viser seg å være godt dokumentert og forklart.

De data vi har valgt å bruke er hentet fra tre kilder: Helsedirektoratets kvalitetsdata [12] og SAMDATA-rapporter [13], og Statistisk sentralbyrås statistikkbank for spesialhelsetjenesten [14]. Majoriteten av tallene som publiseres av disse, har igjen blitt innhentet fra Norsk pasientregister (NPR) [15]. Mens noen få har blitt tatt ut av andre register. Dette vil bli redegjort nærmere for i forhold til hver kategori av variabler. Samtidig som vi forklarer nærmere hva variablene vi benytter måler, og hvorfor de utgjør relevante mål.

### 5.4. Konstruksjon av modeller

Første steg har vært å finne alle variabler vi har sett på som potensielle for vår analyse. Deretter har vi filtrert ut de variablene som ikke er relevante for studiet, inkludert alle uønskede variabler og eksogene variabler. Variabler som ikke er godt dokumentert eller forklart vil bli utelatt.

Kvalitet i helsevesenet vurderes av helheten av de egenskaper og kjennetegn tjenestene deres har, som vedrører evnen til å forbedre helsetilstanden til hvert enkelt individ etter behov. I

Helsedirektoratets offentlige registre finner man en rekke «kvalitetsindikatorer» for somatisk sektor, i tillegg til en rekke målverdier for ventetid og fristbrudd. Vedlegg 2 inneholder en komplett liste av kvalitetsindikatorer. Mange av indikatorene viste seg imidlertid å være mangelfulle. Et gjengående problem med indikatorene er at de er rapportert for få år, eller mangler data for en periode. Disse er merket som «ufullstendig» i vedlegget. Andre indikatorer rapporteres etter helseregion, og ikke helseforetak som ønskelig. Disse er merket som «uegnet» i vedlegget. Validiteten og reliabiliteten vurderes utfra definisjonsarkene til hver kvalitetsindikatorer, som er å finne på Helsedirektoratets sider [12]. Metoden bak indikatoren har vi vurdert opp mot anbefalingene gjennomgått i teorikapittelet, og selv vurdert sammenligningsgraden basert på hvor standardiserte indikatorene virker å være.

Med bakgrunn i drøftelsene i teori-kapittelet, mener vi enkelte av indikatorer ikke er passende mål for kvalitet i denne sammenheng. Dette gjelder blant annet resultatindikatorer som teller overlevelsesrate etter forskjellige inngrep. Overlevelse vil i stor grad på virkes av faktorer utenfor selve behandlingen, og er heller ikke justert for sannsynlighet for vanskelighetsgrad, slik at indikatoren ikke utgjør et sammenlignbart mål. Helsedirektoratet advarer selv om lignende problemstillinger rundt indikatorer [15]. Variabler som baserer seg på ventelister er også utelatt. Dette fordi vi verken finner målet godt nok relatert til vår kvalitetsdefinisjon, basert på skjønn og at det metodisk dårlig utført [16]. Følgende variabler er komplett for vår periode, godt dokumentert innsamlet, sammenlignbare og i tråd med kvalitetsbegrepet vi ønsker å måle:

- $s_1$ : Strykninger fra planlagt operasjonsprogram
- $s_2$ : Lårhalsbruddsoperasjoner innen 48 timer
- $s_3$ : Korridorpasienter
- $s_4$ : Epikriser sendt innen 7 dager

Kostnadseffektiviteten beregnes teoretisk i forhold til den tekniske effektiviteten og prisene til de forskjellige inputene i beregningen. Til vår oppgave har vi ikke priser til hver input, og vi bruker i stedet ett kostnads mål for de totale kostnadene. Utfordringen i denne analysen er å finne gode sammenliknbare indikatorer for både produksjon og kostnadene, og til dette har vi brukt DRG-vektene. Data for både produksjon og kostnader er hentet fra de årlige SAMDATA-rapportene til Helsedirektoratet. Dette medfører at prisforskjeller allerede er korrigert for slik at dataene blir så sammenliknbare som mulig. Antall DRG-poeng ( $p_1$ )

produsert benyttes som produkt, mens kostnader til DRG-produksjon ( $f_2$ ) er innsatsfaktoren. Sistnevnte er beregnet på følgende vis av Helsedirektoratet:

**Tabell 1: Beregning av DRG-kostnader**

Totale driftskostnader (A)
Polikliniske inntekter (B)
Andre inntekter (C)
Ekstern virksomhet (netto) (D)
Kjøp fra andre private institusjoner uten driftsavtale (E)
Forskning og undervisning (F)
Pasientskade-erstatning (G)
Kostnader til DRG produksjon ( $H = A - B - C - D - E - F - G$ )

Teknisk effektivitet skal beregne hvor produktiv hver enkelt DMU/helseforetak er i forhold til de beste helseforetakene i Norge gitt ressursbruken. For å måle dette trenger vi variabler som klart definerer kvanta av både innsatsfaktorer og produkter. Vårt valg av variabler for teknisk effektivitet bygger på Hagen et al. (2001), Hagen & Magnussen (2009) og Nayar & Ozcan (2008) sine liknende målinger av teknisk effektivitet. Fra kildene Helsedirektoratet og Statistisk Sentralbyrå, finner vi mål for produksjon, og ressursbruk i form av årsverk og kapasitet. Ressursbruk i form av vareforbruk, har vi ikke funnet reliable, sammenlignbare tall for. Følgende variabler tilfredsstiller våre krav, og vil bli vurdert videre for å måle den tekniske effektivitet:

- $p_1$ : DRG-poeng
- $k_1$ : Døgnplasser
- $k_2$ : Årsverk leger
- $k_3$ : Årsverk sykepleiere
- $k_4$ : Årsverk hjelpepleiere
- $k_5$ : Årsverk annet behandlingspersonell
- $k_6$ : Årsverk annet personell

### 5.5. Statistisk testing av variabler

I vurderingen av de resterende variablene har vi utført korrelasjons- og Banker-tester, som forklart i metode-kapittelet. Resultatene av disse testene har så blitt brukt til sluttvurderingen av modellene og variablene.



### 5.5.1. Korrelasjon

Korrelasjonsanalyse utføres for å undersøke om input-variabelen faktisk påvirker output-variabelen, og om input-variabler korrelerer med hverandre og potensielt representerer redundant informasjon. Først testes output,  $p_1$ , mot hver enkelt input,  $k$ , deretter testes hver enkelt av inputene mot hverandre. Siden output for kvalitet er en konstant, analyseres bare input-variablene mot hverandre. Tabellene som følger viser resultatene i form av Pearsons korrelasjonskoeffisient R, for hver effektivitetstype, år og variabelpar.

**Tabell 2: Korrelasjonskoeffisienter for tekniske variabler**

	$p_1-k_1$	$p_1-k_2$	$p_1-k_3$	$p_1-k_4$	$p_1-k_5$	$p_1-k_6$	$k_1-k_2$	$k_1-k_2$	$k_1-k_4$	$k_1-k_5$	$k_1-k_6$
2008(09)	0,99	0,99	0,99	0,90	0,95	0,98	0,99	0,99	0,91	0,95	0,97
2009(09)	0,99	0,98	0,99	0,95	0,97	0,97	0,97	0,98	0,95	0,96	0,96
2009(10)	0,99	0,97	0,99	0,94	0,97	0,97	0,97	0,98	0,95	0,96	0,96
2010(10)	0,98	0,98	0,99	0,95	0,97	0,97	0,94	0,96	0,95	0,93	0,93
2010(11)	0,98	0,98	0,99	0,95	0,97	0,98	0,94	0,96	0,95	0,93	0,93
2011(11)	0,98	0,98	0,99	0,97	0,97	0,97	0,95	0,97	0,96	0,93	0,94
	$k_2-k_3$	$k_2-k_4$	$k_2-k_5$	$k_2-k_6$	$k_3-k_4$	$k_3-k_5$	$k_3-k_6$	$k_4-k_5$	$k_4-k_6$	$k_5-k_6$	
2008(09)	0,99	0,88	0,98	0,97	0,89	0,96	0,97	0,83	0,86	0,92	
2009(09)	0,99	0,92	0,99	1,00	0,93	0,98	0,98	0,90	0,91	0,99	
2009(10)	0,99	0,92	0,99	1,00	0,93	0,98	0,98	0,90	0,91	0,99	
2010(10)	0,99	0,94	1,00	1,00	0,94	0,99	0,98	0,93	0,94	0,99	
2010(11)	0,99	0,94	1,00	1,00	0,94	0,99	0,98	0,93	0,94	0,99	
2011(11)	0,99	0,96	0,99	0,99	0,96	0,98	0,98	0,93	0,94	0,99	

**Tabell 2: Korrelasjonskoeffisienter kostnadsvariabler**

	$f_2-p_1$
2008(09)	0,99
2009(09)	0,99
2009(10)	0,98
2010(10)	0,98
2010(11)	0,99
2011(11)	0,98

Korrelasjonsanalysene viser at det er tilnærmet perfekt positiv korrelasjon mellom alle variablene for teknisk effektivitet og kostnadseffektivitet. At output korrelerer med input, er et

godt signal på at de er relatert til hverandre og dermed bør benyttes videre i analysen. Korrelasjon mellom de uavhengige variablene for teknisk kvalitet, gir derimot et signal om at variablene potensielt har redundant informasjon. Dette er i tråd med hva man kan forvente. DRG-produksjonen ( $p_1$ ) bestemmer direkte kostnadene, og økt produksjon avhenger av økt antall årsverk og kapasitet i form av døgnplasser. At de uavhengige variablene for teknisk effektivitet også korrelerer, er heller ingen overraskelse siden fire av fem variabler er årsverk. De forskjellige typer årsverk øker gjerne proporsjonalt med hverandre. Flere operative døgnplasser, gir likedan større behov for personell og øker produksjonen.

**Tabell 3: Korrelasjonskoeffisienter for kvalitetsvariabler**

	$s_1-s_2$	$s_1-s_3$	$s_1-s_4$	$s_2-s_3$	$s_2-s_4$	$s_3-s_4$
2008	-0,06	0,20	0,26	-0,09	-0,08	-0,01
2009	-0,08	-0,01	0,65	0,02	-0,19	0,13
2010	-0,13	-0,04	0,36	0,17	-0,33	0,20
2011	0,19	-0,01	0,56	0,11	0,07	0,19

Analysen viser at det er varierende korrelasjon mellom kvalitetsvariablene; strykninger ( $s_1$ ), operasjoner ( $s_2$ ), korridordøgn ( $s_3$ ) og epikrisetid ( $s_4$ ). Generelt ser man at det er en svak korrelasjon, men at den både er positiv og negativ. Kvalitetsvariablene er av forskjellig art, og måler ting som ikke noe direkte forhold til hverandre, så dette kan sies å være som forventet. Unntaksvis ser man styrkninger har en middels sterk positiv korrelasjon med epikrisetid.

Følger vi Dyson et al. (2001) sine forholdsregler, trenger vi videre testing og vurdering av de forskjellige variablene før disse potensielt kan tas ut av modellen. Vi bestemmer oss derfor for å kjøre Banker-tester på både modellene for teknisk effektivitet og modellene for kvalitet.

### 5.5.2. Banker-test

En opprinnelig modell,  $M_1$ , testes mot mindre detaljerte modellversjoner,  $M_n$ . For teknisk effektivitet kan vi slå sammen årsverk for å redusere antall variabler. For kvalitet undersøkes endringen som oppstår hvis én variabel ekskluderes, siden det ikke er mulig å summere opp de forskjellige relative målene. Nedenfor følger en tabell som viser de forskjellige modellspesifikasjonene som undersøkes og resultatet i form av F-verdier fra hypotesetester.

## Kvalitet:

**Tabell 3: Banker-test modellspekifikasjoner for kvalitet**

Modell	Output	Input
$M_1$	1	Strykninger, operasjoner, korridordøgn, epikrisetid
$M_2$	1	Strykninger, operasjoner, korridordøgn
$M_3$	1	Strykninger, operasjoner, epikrisetid
$M_4$	1	Strykninger, korridordøgn, epikrisetid
$M_5$	1	Operasjoner, korridordøgn, epikrisetid

**Tabell 4: Banker-test på datasett for kvalitet (F-verdier)**

Periode / $H_0$	Ekspensialfordelt			
	$M_1 = M_2$	$M_1 = M_3$	$M_1 = M_4$	$M_1 = M_5$
2008	2,049	1,479	1,427	1,066
2009	1,541	1,276	1,312	1,264
2010	1,541	1,276	1,312	1,264
2011	1,186	1,319	1,377	1,203
Periode / $H_0$	Halv-normalfordelt			
	$M_1 = M_2$	$M_1 = M_3$	$M_1 = M_4$	$M_1 = M_5$
2008	3,401	1,812	1,805	1,182
2009	2,079	1,585	1,451	1,527
2010	2,079	1,585	1,451	1,527
2011	1,309	1,481	1,623	1,408

Forutsatt en eksponensialfordeling av verdiene med signifikansnivå 0,05 og 38 frihetsgrader får vi en kritisk F-verdi på 1,717. Forutsetter vi en halv-normalfordeling av verdiene med signifikansnivå 0,05 og med 19 frihetsgradene, får vi en kritisk F-verdi på 2,168. I enkelte tilfeller går F-verdiene over den kritiske F-verdien, slik at  $H_0$  kan forkastes.

## Teknisk effektivitet:

**Tabell 5: Banker-test modellspekifikasjoner for teknisk effektivitet**

Modell	Output	Input (antall variabler)
$M_1$	DRG-poeng	Døgnplasser, årsverk leger, sykepleiere, hjelpepleiere, annet behandlingspersonell, annet personell (6)
$M_2$	DRG-poeng	Døgnplasser, årsverk (2)
$M_3$	DRG-poeng	Døgnplasser, årsverk leger, alle andre årsverk (3)
$M_4$	DRG-poeng	Døgnplasser, årsverk leger, årsverk «pleie/behandlings»-personell, årsverk annet personell (4)
$M_5$	DRG-poeng	Døgnplasser, årsverk leger, pleiere, annet behandlingspersonell, annet personell (5)

**Tabell 6: Benker-test på datasett for tekniske effektivitet (F-verdier)**

	Ekspensialfordelt			
	$M_1 = M_2$	$M_1 = M_3$	$M_1 = M_4$	$M_1 = M_5$
2008(09)	1,905	1,776	1,385	1,287
2009(09)	1,563	1,443	1,339	1,229
2009(10)	1,629	1,517	1,404	1,244
2010(10)	1,563	1,495	1,344	1,161
2010(11)	1,704	1,572	1,387	1,160
2011(11)	1,582	1,520	1,314	1,055
	Halv-normalfordelt			
	$M_1 = M_2$	$M_1 = M_3$	$M_1 = M_4$	$M_1 = M_5$
2008(09)	2,392	2,137	1,462	1,287
2009(09)	1,642	1,529	1,319	1,229
2009(10)	1,684	1,572	1,359	1,244
2010(10)	1,590	1,541	1,285	1,161
2010(11)	1,776	1,667	1,351	1,160
2011(11)	1,761	1,698	1,335	1,055

Forutsatt en eksponensialfordeling av verdiene med signifikansnivå 0,05 og 38 frihetsgrader får vi en kritisk F-verdi på 1,717. Forutsetter vi en halv-normalfordeling av verdiene med signifikansnivå 0,05 og med 19 frihetsgradene, får vi en kritisk F-verdi på 2,168. I noen tilfeller holder F-verdiene seg under den kritiske F-verdien, og dermed kan  $H_0$  forkastes.

## 5.6. Endelige modeller

Korrelasjonsanalysene og Banker-testene viser oss hvilke modeller og variabler som potensielt kan inkluderes i de modellene vi bestemmer oss for å bruke. Samtidig tar vi hensyn til at Banker-testene har en helhetlig synsvinkel. Dette betyr at større endringer for enkelte DMU-er ikke nødvendigvis fører til stort utslag for hele modellen. Med bakgrunn i dette har vi vurdert endringer i hver enkelt helseforetaks effektivitetsscore når vi bestemmer oss for hvilke modeller vi skal bruke. Får en reduksjon av variabler stor betydning for enkelte DMU-ers effektivitetsscore, velger vi å beholde variabelen. Sluttevalueringene resulterer i de tre modellene vi vil bruke videre i vår studie.

### 5.6.1. Modell for kvalitet

Til kvalitetsanalysen har vi valgt et oppsett hvor output settes som konstanten 1 opp mot input beregnet som relative mål på kvalitet. Datagrunnlaget for variablene er bearbeidet fra absoluttverdier til prosentmål for å passe til undersøkelsen. I tillegg har målene blitt bearbeidet slik at prosenten skal minimeres. Med bakgrunn i drøfting av målenes egenskaper og testene ovenfor, har vi valgt følgende fire variabler til vår analyse.

*Epikriser etter 7 dager i prosent* måler andelen epikriser som ikke sendes til neste behandlingsinstitusjon i tide. Målet er et uttrykk for kommunikasjon i behandlingsforløpet, og lang epikrisetid er med på å svekke pasientens oppfølging. Svekket oppfølging kan påvirke helseeffekteten negativt, og relateres på denne måte opp mot resultatet av behandlingen.

*Lårhalsbruddoperasjoner etter 48 timer i prosent* måler andelen lårhalsbrudd på pasienter over 65 år som ikke opereres innen 48 timer etter innleggelse. Dette er en prosessindikator valgt ut av Helsedirektoratet ettersom lang ventetid på lårhalsbruddoperasjoner har blitt relatert til dårligere prestasjoner.

*Strykninger fra planlagt operasjonsprogram i prosent* måler andelen operasjoner som ikke blir utført til planlagt tidspunkt. Strykninger sier noe om planlegging og utførelse i institusjonen, noe som gjør dette til en prosessindikator. Samtidig vil en strykning gi forlenget lidelse til pasienten som må vente lengre på behandling enn først planlagt.

*Korridorpasienter i prosent* måler andelen innlagte pasienter som ikke ligger på eget rom i avdelingen. Slike pasienter kan være plassert på korridor, bad, skyllerom, dagligstue eller liknende. Pasienter plassert på korridor får ikke ivaretatt sin integritet og verdighet på

tilstrekkelig vis, og plasseringen svekker pleie og behandling av pasienten. Antall korridorpasienter måles kl. 0700, og får med de pasienter som har overnattet på gangen før disse tas videre til behandling.

### **Vurdering av kvalitetsmodell**

Hele kvalitetsmodellen baserer seg på relative mål, og avgjøres helt uten hensyn til skala. På denne måten kan vi bruke prosentmål på antall korridorpasienter uten å falle ned i Dysons fallgruver.

Struktur- og resultatorientert kvalitet blir ikke målt ettersom godt dokumenterte og sammenlignbare indikatorer ikke er tilgjengelige. I henhold til teorien vil et prosessorientert kvalitetsmål være hensiktsmessig for å måle helseforetakenes indre aktivitet. Selv om de fire kvalitetsindikatorne er valide i forhold til kvalitetsbegrepet, kan vi ikke si at kvalitetsbegrepet i sin helhet blir målt. For å opprettholde validiteten begrenser vi kvalitetsbegrepet til kun prosesskvalitet.

Kvalitetsindikatorne kan også sies å være reliable ettersom prosessen for måling og registrering av de forskjellige variablene er godt dokumentert og standardisert. Samtidig kommer alle mål fra pasientjournalene, og er ikke basert på noen typer skjønn som kan gi systematiske eller usystematiske skjevheter i dataene. Kvalitetsmodellen er også valid ettersom alle variablene er prosessorienterte og vi konsekvent måler prosesskvaliteten. For videre studier er det likevel anbefalt å bruke kvalitetsindikatorer for både struktur, prosess og resultater, gitt at disse er tilgjengelige.

#### 5.6.2. Modell for kostnadseffektivitet

*DRG-produksjon* setter vi som output til hvert enkelt sykehus. Denne målingen representerer den totale produksjonen utført for hvert helseforetak i studiet. DRG-produksjonen måles, som vist i kapittel 2, igjennom å vektlegge hver pasient med en viss poengmengde etter hvor høy ressursbruken er beregnet å være for behandlingen. Ettersom hver pasient vektlegges med en mengde DRG-poeng, blir de forskjellige helseforetakene sammenliknbare med hensyn på total produksjonsmengde i form av pasienter. DRG-vektene brukes både for innlagte pasienter og polikliniske pasientbesøk.

*DRG-kostnader* blir eneste input i modellen for kostnadseffektivitet. Korrelasjonsanalysen viser en sterk korrelasjon mellom DRG-kostnadene som input og DRG-produksjonen som output. DRG-kostnader er et mål for de totale kostnader påført helseforetaket som kan relateres DRG-vektede pasienter. Derfor kan denne inputen relateres direkte til behandlingen av pasientene uten å inkludere andre mål.

### **Vurdering av kostnadsmodell**

En potensiell svakhet er endringer i DRG-vektene som skjer årlig. Dette kan medføre en forbedring eller forverring i kostnadseffektivitetsmålet til helseforetakene alt etter hvordan foretakets pasientmiks er. Dette tar vi hensyn til ved å forholde oss til korrigerte tall fra kun SAMDATA-rapport om gangen. For eksempel vil vi kun bruke SAMDATA-rapport for 2010 når vi beregner endringer mellom 2009 og 2010. Da vil tall for 2009 være korrigert for å bli mer sammenliknbare med tall for 2010.

#### 5.6.3. Modell for teknisk effektivitet

Etter å ha vurdert de statistiske testene gjennomført på de utvalgte variabler for teknisk effektivitet, og drøftet variablenes betydning, har vi kommet fram til en modell bestående av fem variabler. Årsverk for sykepleiere, hjelpepleiere og annet behandlingspersonell, er slått sammen til en variabel «årsverk behandlingspersonell». Dette fordi de utgjør en tilsvarende funksjon i sykehuset, sammenlignet med de andre årsverkene, som vi har valgt å holde avskilt fra hverandre. Begrunnelsen for dette er at en sammenslåing vil gå hardt utover detaljnivået i modellen. Lignende studier, gjennomgått i teorien, påpeker også at det kan være høvelig å holde årsverk for leger avskilt, siden de utgjør en fundamental rolle i prosessen. En nærmere beskrivelse av variablene benyttet i modellen følger.

*DRG-produksjon* representerer helseforetakenes totale output, og er lik outputen brukt i kostnadseffektivitetsmodellen.

*Årsverk leger* viser til den arbeidsmengden i antall årsverk leger ved hvert helseforetak har lagt ned for hvert år. Målet for årsverk er et totalmål, og begrenser seg ikke til kun de årsverk brukt på behandling av pasienter.

*Årsverk behandlingspersonell* er et akkumulert mål for alt behandlingspersonell utenom leger. Målet omfatter sykepleiere, hjelpepleiere og annet behandlingspersonell målt i antall årsverk

fra SAMDATA-rapportene. Dette er også et totalmål som ikke begrenser seg til DRG-produksjonen.

*Årsverk annet personell* er en oppsummering av årsverk som ikke inkluderer helsepersonell i form av leger, sykepleiere og behandlingspersonell generelt. Dette er akkumulerte årsverk for ledere, teknisk personell og administrativt personell i helseforetaket.

*Antall døgnplasser* er et mål for størrelsen på sykehuset. Høyere antall døgnplasser medfører en høyere kapasitet med hensyn til innleggelse av pasienter, og kostnader for å holde døgnplassene gående.

### **Vurdering av teknisk modell**

Noen momenter ved den tekniske effektiviteten bør nevnes. For det første har vi ikke funnet noe reliabelt mål for bruk av medisiner og medisinsk utstyr. Dette medfører potensielt til at noen sykehus som er effektive med hensyn på dette målet får en mindre effektivitetsverdi enn de fortjener. Samtidig har vi heller ingen reliable mål på produksjon av forskning og studenter ved helseforetakene. Dette medfører et bias på noen sykehus ettersom årsverkene ikke går direkte til behandling av pasienter (DRG-produksjon), og disse får en lavere effektivitetsscore enn de fortjener ettersom den fulle produksjonen ikke blir belyst. Kostnader til forskning og utvikling utgjør rundt én prosent av totale driftskostnader, hvis produksjonen er i samme størrelsesorden, er dette biaset tilnærmet ubetydelig. Samtidig vil foretakene få arbeidskapasitet tilbake fra studenter som utfører oppgaver i undervisningssammenheng, slik at effekten på årsverk vil utjevnes.

#### 5.7. Outliers

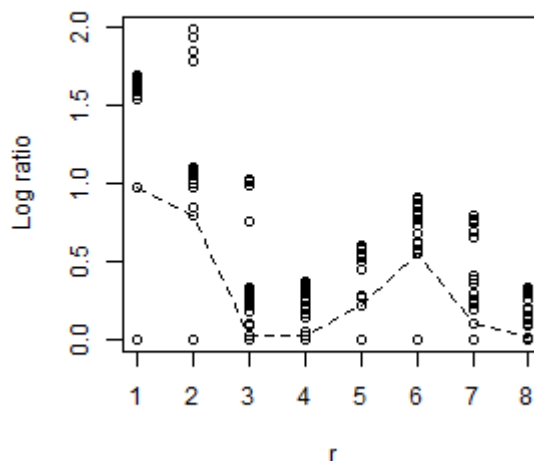
Wilson-rutinen, som beskrevet i metodekapittel, anvendes for å finne potensielle outliere i datasettet. Resultatet fremkommer grafisk gjennom et plott og en matrise som viser de mest ekstreme DMU-ene som fjernes. I plottet viser horisontal akse antallet DMU-er som er fjernet ( $r$ ), og den vertikale aksene viser en logaritmisk rate som indikerer betydningen av å fjerne variablene. Hvis den logaritmiske raten øker når variablene fjernes, er dette et tegn på at variablene er outliere. Matrisen viser hvilke DMU-er som er fjernet og logaritmeratioen  $R^r$ , som rutinen søker å minimere. Outlieranalyse for kvalitetsmodellen lar seg ikke gjøre siden outputene er konstant.



Modell for kostnadseffektivitet 2011(11)

**Tabell 7: Outlieranalyse kostnadseffektivitet**

<i>r</i>								$R^r$	
1	3							0,18	
2	5	3						0,05	
3	11	5	3					0,03	
4	14	11	5	3				0,02	
5	2	14	18	5	3			0,01	
6	2	14	18	11	5	3		0,00	
7	2	14	18	4	11	5	3	0,00	
8	2	14	18	17	4	11	5	3	0,00

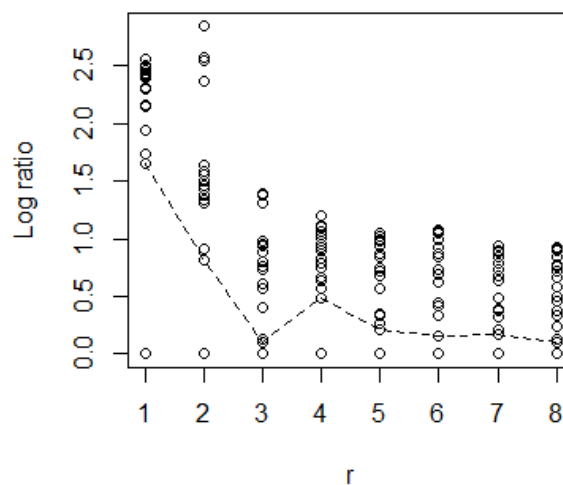


For kostnadseffektivitet ser vi at DMU 3 skiller seg ut. Forklaringen bak dette mener vi er den samme for begge modellene. DMU 3 skiller seg spesielt ut grunnet dens massive størrelsen på DMU-en opp mot andre DMU-er, og ikke på grunn av feilmålinger. Vi velger ikke å fjerne DMU-en fra datasettet. DMU 3 har ingen påvirkning på andre helseforetaks tekniske effektivitetsscore gitt konstant skalautbytte. Samtidig har ikke DMU 3 en teknisk effektivitetsscore som er urimelig sammenliknet med andre helseforetak.

Modell for teknisk effektivitet 2011(11).

**Tabell 8: Outlieranalyse teknisk effektivitet**

<i>r</i>								$R^r$	
1	3							0,07	
2	11	3						0,01	
3	14	11	3					0,00	
4	4	14	11	3				0,00	
5	7	4	14	11	3			0,00	
6	7	1	4	14	11	3		0,00	
7	7	1	4	5	14	11	3	0,00	
8	7	1	4	5	6	14	11	3	0,00



I plottet ser vi en topp ved fjerning av en DMU, og i tabellen ser vi at DMU 3 prioriteres. Neste topp er ved fjerning av fire DMU-er, og da er det DMU 4, 14, 11 og 3 som prioriteres. Samtidig ser man utfra logaritmeratioen i matrisen at verdien er minimert allerede etter å ha fjernet én DMU.

Resultatene av analysen for de resterende datasettene, er å finne i vedlegg 4. Resultatet viser det samme som de gjennomgått ovenfor, dermed er vår beslutning å beholde hele datasettet.

### 5.8. Skalautbytte

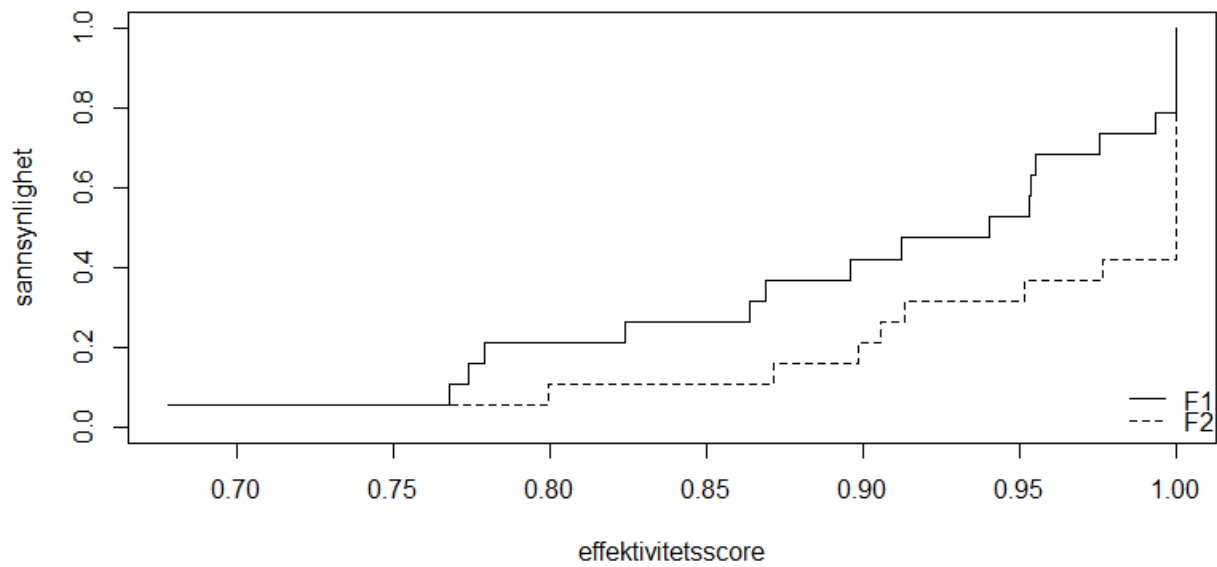
Hvorvidt dataomhyllingsanalysene skal beregnes med forutsetning om konstant eller variabelt skalautbytte, er et spørsmål man må drøfte med basis i kunnskap og erfaringer relatert til sektoren som analyseres (Cooper et al., 2011). I henhold til norsk helsesektoren har vi ingen kilder som slår fast skalautbytte. Studier utført på amerikanske sykehus gir skiftende resultater. Resultatene til Wilson & Carey (2004) finner bevis for økende skalautbytte blant sykehus med over middels størrelse. Det påpekes i studiet at mange sykehus fusjonerer for å oppnå stordriftsfordeler. Sammenslåingen finner også sted i Norge, det er ikke umulig å se for seg at stordriftsfordeler ligger til grunn også her.

Intuitivt bør det gå med like mye ressurser til hver pasient som behandles, slik at økningen er proporsjonal, og dermed taler for konstant skalautbytte. Samtidig er organiseringen av våre analyseobjekter slik at de fleste helseforetak består av flere mindre sykehus. Hvis de mindre sykehusene ikke får utnyttet kapasiteten maksimalt på grunn av for lav etterspørsel, vil de være mer ressurskrevende per produksjonsenhet enn større sykehus med høy etterspørsel.

Basert på datasettet, kan man undersøke om valg av skalautbytte har stor påvirkning på beregningene. Figur 5 viser beregnede effektivitetsscorer for teknisk effektivitet for 2011, hvor  $F_1$  er beregnet med CRS, og  $F_2$  med VRS. Plottet viser at valg av skalautbytte utgjør en viss forskjell for beregnede effektivitetsscorer. Nærmest dobbelt så mange av helseforetakene beregnes til å være effektive ved variabelt skalautbytte opp mot konstant skalautbytte.

Samtidig har helseforetakene viser plottet av noen helseforetak har en betydelig økning i effektivitetsscorene om variabelt skalautbytte forutsettes. Banker-test på eksempelåret viser at

det er signifikant forskjell mellom  $F_1$  og  $F_2$ . Det samme finner vi for andre år og modeller.



**Figur 5: Skalautbytte**

For å være på den sikre siden, har vi ikke tatt noen beslutning angående skalautbytte, og beregner videre estimater for både VRS og CRS.

## 6. Resultater

Dette kapittelet er dedikert til presentasjon av studiets resultater og kommentarer til resultatene. Først presenteres og tolkes resultatene fra dataomhyllingsanalysene og Malmquist-indeksene. Deretter presenteres en sensitivitetsanalyse utført på estimatene. Til slutt går vi over til å lete etter relasjoner mellom kvalitet og effektivitet ved hjelp av diverse statistiske metoder.

### 6.1. Trinn 1: DEA- og MPI-resultater

Resultatene fra dataomhyllingsanalysene og Malmquist-produktivitetsindeksene er vist i tabeller basert på modellene for prosesskvalitet (kvalitet), kostnadseffektivitet (kostnad) og teknisk effektivitet (teknisk). Resultatene for kostnadseffektivitet og teknisk effektivitet er beregnet både ved forutsetning om konstant (CRS) og variabelt (VRS) skalautbytte. Alle effektivitetsscorene er basert på prinsippet om supereffektivitet. De fire første sifrene representerer året dataene er samlet inn i. De to sifrene i parentes representerer rapportåret som dataene er korrigert etter. Resultatene fra dataomhyllingsanalysene er korrigert slik at all scorer over 1 er effektive. Malmquist-produktivitetsindeksene er korrigert slik at en score over 1 betyr en kvalitets- eller produktivitetsforbedring. Alle scorer og indekser lik 1 er skravert grønt.

#### 6.1.1. Resultater for dataomhyllingsanalysene

I tabellene som følger (tabell 9-12) er dataomhyllingsanalysen representert. Før de enkelte tabellene kommenteres gjennomgår vi noen generelle trekk ved resultatene.

For det første må effektivitetsscorene ved variabelt skalautbytte beregnes via helseforetak i egen skala. Dermed vil store og små helseforetak få høyere effektivitet enn ved konstant skalautbytte. Et spesielt tilfelle her er DMU 3 som er ineffektiv gitt CRS, men mest kostnadseffektiv og teknisk effektiv gitt VRS. Dette kommer av at helseforetaket er betydelig større enn de andre.

For det andre påvirker variabelbruken effektivitetsscorene. Kvalitetsmodellen med konstant som output og prosentmål som inputs har større variasjon i maks og minimumsverdi og større standardavvik. Kostnadsmodellen med få variabler gir få effektive enheter, og dermed gjennomsnittlig lavere effektivitetsscore.

Til slutt er det noen VRS-verdier som ikke er tatt med i beregningene. Helseforetakene dette gjelder er effektive DMU-er som ligger i ytterkantene av utvalget i kostnads- eller teknisk modell. Når supereffektivitetsscoren skal beregnes, vil ikke den effektive DMU-en komme mellom origo og VRS-fronten, noe som gjør det umulig å beregne en score. Disse har blitt merket med en asterisk og gitt verdien 1,000 siden de er effektive.

**Tabell 9: Dataomhyllingsanalyser, 2008**

DMU	2008(09)				
	CRS			VRS	
	Kvalitet	Kostnad	Teknisk	Kostnad	Teknisk
DMU1	0,668	0,862	0,934	0,876	0,992
DMU2	0,771	0,815	0,878	0,832	0,892
DMU3	0,735	0,764	0,972	2,221	2,221
DMU4	1,725	0,937	0,976	0,996	1,055
DMU5	1,032	0,960	1,063	1,017	1,124
DMU6	0,778	0,986	1,007	0,990	1,116
DMU7	1,366	1,014	1,120	1,023	1,178
DMU8	1,215	0,979	1,057	1,005	1,081
DMU9	0,966	0,858	0,873	0,883	0,894
DMU10	0,771	0,896	0,917	0,925	0,991
DMU11	0,770	0,867	0,914	0,981	0,966
DMU12	1,077	0,733	0,840	0,768	1,002
DMU13	0,846	0,908	0,882	0,928	0,939
DMU14	0,652	0,774	0,893	0,867	0,929
DMU15	0,917	0,862	0,824	0,903	0,910
DMU16	0,844	0,828	0,836	*1,000	*1,000
DMU17	0,809	0,795	1,013	0,799	1,066
DMU18	0,659	0,722	0,851	0,750	0,898
DMU19	1,417	0,643	0,698	0,768	*1,000
Gjennomsnitt	0,948	0,853	0,924	0,975	1,066
Standardavvik	0,286	0,097	0,098	0,307	0,284
Maks	1,725	1,014	1,120	2,221	2,221
Min	0,652	0,643	0,698	0,750	0,892

I kvalitetsmodellen er fronten konstruert via DMU 4, 5, 7, 8, 12 og 19. Kun DMU 7 er kostnadseffektiv med hensyn på konstant skalautbytte. Inkluderer vi variabelt skalautbytte, vil også DMU 3, 5 og 8 bli kostnadseffektive. For teknisk modell er det DMU 5, 6, 7, 8 og 17 som er mest effektive med konstant skalautbytte. DMU 3, 4 og 12 blir også teknisk effektiv forutsatt variabelt skalautbytte.

Samtidig ser vi at DMU 4, 5, 7, 8, 12 og 19 ligger både på kvalitetsfronten og på den tekniske fronten gitt variabelt skalautbytte. Samtidig er også DMU 5, 7 og 8 kostnadseffektive. Forutsettes konstant skalautbytte, er det kun DMU 5, 7 og 8 som er teknisk effektive og på kvalitetsfronten. Da er kun DMU 7 både på kvalitetsfronten og kostnadsfronten.

**Tabell 10: Dataomhyllingsanalyser, 2009**

DMU	2009(09)				
	CRS			VRS	
	Kvalitet	Kostnad	Teknisk	Kostnad	Teknisk
DMU1	0,647	0,903	0,995	0,906	1,038
DMU2	0,719	0,778	0,985	0,784	1,001
DMU3	0,641	0,762	1,074	2,145	2,145
DMU4	1,204	0,986	1,023	1,026	1,075
DMU5	0,903	0,972	0,984	0,985	1,040
DMU6	0,900	1,005	1,044	1,007	1,061
DMU7	1,150	0,995	1,104	1,021	1,183
DMU8	1,694	0,989	1,063	0,997	1,067
DMU9	0,877	0,887	0,893	0,895	0,895
DMU10	0,694	0,916	0,937	0,960	1,023
DMU11	0,925	0,875	0,920	0,966	0,920
DMU12	0,826	0,719	0,841	0,768	0,961
DMU13	1,690	0,861	0,848	0,868	0,849
DMU14	0,737	0,777	0,978	0,849	0,986
DMU15	0,923	0,832	0,921	0,888	1,126
DMU16	0,781	0,775	0,780	1,186	*1,000
DMU17	0,603	0,774	0,835	0,782	0,879
DMU18	0,725	0,672	0,830	0,681	0,843
DMU19	0,896	0,646	0,681	*1,000	*1,000
Gjennomsnitt	0,923	0,849	0,934	0,985	1,057
Standardavvik	0,305	0,109	0,108	0,296	0,271
Maks	1,694	1,005	1,104	2,145	2,145
Min	0,603	0,646	0,681	0,681	0,843

For 2009 er kvalitetsbildet litt annerledes. Fire DMU-er (4, 7, 8 og 13) utgjør nå den effektive kvalitetsfronten. For konstant skalautbytte har DMU 6 blitt den mest kostnadseffektive, og DMU 3, 4, 6, 7 og 8 er de teknisk effektive. Ser vi på variabelt skalautbytte, blir ytterligere fem DMU-er kostnadseffektive og åtte teknisk effektive.

Resultatene viser et mindre klart forhold mellom kvalitet og effektivitet. Av de seks DMU-ene som er både kostnadseffektiv og teknisk effektiv gitt variabel skalautbytte, er bare to på

kvalitetsfronten. Gitt variabelt skalautbytte, er kun DMU 6 effektiv på begge forklarende variabler, men den scorer ikke høyt på kvalitet. Forutsatt konstant skalautbytte ligger kun tre DMU-er (4, 7 og 8) på både kvalitetsfront og teknisk front. Ved variabelt skalautbytte er disse tre også teknisk effektive. De to høystscorende på kvalitet har middels effektivitetsscorer.

**Tabell 11: Dataomhyllingsanalyser, 2010**

DMU	2010(10)				
	CRS			VRS	
	Kvalitet	Kostnad	Teknisk	Kostnad	Teknisk
DMU1	1,125	0,887	1,006	0,903	1,048
DMU2	0,811	0,751	0,896	0,789	0,897
DMU3	0,678	0,687	1,177	2,120	2,120
DMU4	1,166	0,924	1,018	1,056	1,104
DMU5	0,696	0,898	0,952	0,964	1,030
DMU6	1,092	1,028	1,124	1,040	1,138
DMU7	1,152	0,972	1,117	1,007	1,130
DMU8	2,769	0,921	1,074	0,964	1,075
DMU9	0,871	0,862	0,918	0,906	0,933
DMU10	0,710	0,863	0,974	0,929	1,068
DMU11	0,596	0,811	0,867	0,963	0,878
DMU12	1,110	0,692	0,848	0,773	0,951
DMU13	1,468	0,837	0,886	0,874	0,886
DMU14	0,769	0,739	0,951	0,860	0,983
DMU15	0,824	0,776	0,933	0,861	1,128
DMU16	0,733	0,684	0,749	*1,000	*1,000
DMU17	0,579	0,658	0,765	0,665	0,788
DMU18	0,944	0,625	0,802	0,673	0,815
DMU19	0,721	0,549	0,687	0,766	*1,000
Gjennomsnitt	0,990	0,798	0,934	0,953	1,051
Standardavvik	0,478	0,125	0,130	0,296	0,272
Maks	2,769	1,028	1,177	2,120	2,120
Min	0,579	0,549	0,687	0,665	0,788

Syv enheter utgjør kvalitetsfronten i 2010, fire av disse (DMU 4, 7, 8 og 13) utgjorde fronten forrige periode. Gitt konstant skalautbytte er fortsatt DMU 6 den eneste kostnadseffektive, og for teknisk kvalitet er seks effektive, av disse er kun DMU 1 ny på fronten. Forutsetter vi variabelt skalautbytte, er fem kostnadseffektive, fire av de samme fra forrige periode, og hele 11 DMU-er er teknisk effektive, hvor bare DMU 2 har mistet sin posisjon på fronten fra forrige år.

Igjen ser vi vi at flere av DMU-ene på kvalitetsfronten også er teknisk effektive og/eller kostnadseffektive. DMU 1, 4, 6, 7 og 8 ligger både på kvalitetsfronten og på teknisk front forutsatt konstant eller variabelt skalautbytte. Forutsettes variabelt skalautbytte, er DMU 4, 8 og 7 både på kvalitetsfronten og kostnadsfronten. DMU 12 og 13 er også på kvalitetsfronten, men er ikke kostnads- eller teknisk effektiv under konstant eller variabelt skalautbytte.

**Tabell 12: Dataomhyllingsanalyser, 2011**

DMU	2011(11)				
	CRS			VRS	
	Kvalitet	Kostnad	Teknisk	Kostnad	Teknisk
DMU1	0,616	0,880	0,940	0,892	0,977
DMU2	0,603	0,763	0,912	0,826	0,913
DMU3	0,749	0,693	1,173	1,867	1,867
DMU4	1,260	0,923	0,975	1,000	1,037
DMU5	0,640	0,894	0,955	0,963	1,043
DMU6	1,200	1,012	1,155	1,026	1,172
DMU7	1,415	0,988	1,058	1,020	1,099
DMU8	1,499	0,958	1,086	1,006	1,088
DMU9	0,715	0,886	0,896	0,930	0,906
DMU10	0,504	0,858	0,864	0,892	0,899
DMU11	0,556	0,893	0,954	1,067	1,017
DMU12	1,092	0,774	0,824	0,850	0,951
DMU13	0,921	0,840	0,869	0,874	0,871
DMU14	1,630	0,790	0,994	0,879	1,020
DMU15	0,722	0,846	0,952	0,904	1,177
DMU16	0,644	0,771	0,779	1,190	*1,000
DMU17	0,509	0,732	0,774	0,738	0,799
DMU18	0,862	0,717	0,768	0,765	0,768
DMU19	0,771	0,631	0,678	*1,000	*1,000
Gjennomsnitt	0,890	0,834	0,927	0,984	1,032
Standardavvik	0,344	0,100	0,129	0,233	0,224
Maks	1,630	1,012	1,173	1,867	1,867
Min	0,504	0,631	0,678	0,738	0,768

I 2011 er det DMU 4, 6, 7, 8, 12 og 14 som ligger på kvalitetsfonten. Ved konstant skalautbytte er DMU 6 eneste kostnadseffektive, og sammen med DMU 3, 7 og 8 utgjør disse fire helseforetakene den tekniske effektivitetsfronten. Ved variabelt skalautbytte er DMU 3, 4, 7, 8, 11, 16 og 19 også beregnet som kostnadseffektive DMU-er og DMU 11, 14 og 15 blir teknisk effektive.



Igjen ser vi at DMU 4, 6, 7, 8 og 14 både ligger på kvalitetsfronten og er teknisk effektive gitt variabelt skalautbytte. DMU 4, 6, 7 og 8 er både på kvalitetsfronten og kostnadsfronten gitt variabelt skalautbytte. Kun DMU 12 er på kvalitetsfronten uten å være enten kostnadseffektiv eller teknisk effektiv.

### 6.1.2. Resultater for Malmquist-produktivitetsindekser

Tabell 13-18 viser resultater for Malmquist-produktivitetsindekser og dekomponeringene. For hver modell kan vi tolke alle verdier over 1 som en forbedring i ytelse mens alle tall under 1 er reduksjon i ytelse. Resultatene fra Malmquist-beregningene kan ikke sammenliknes direkte med resultatene fra dataomhyllingsanalysene. Dette fordi datagrunnlaget til Malmquist-beregningene er justert for å gjøre årene sammenliknbare.

**Tabell 13: Malmquist-produktivitetsendring**

DMU	2008-2009(09)			2009-2010(10)			2010-2011(11)		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
DMU1	0,934	1,012	0,981	1,733	1,005	0,924	0,944	0,954	1,056
DMU2	1,135	0,921	0,932	1,274	0,964	1,059	0,951	1,010	0,967
DMU3	1,060	0,963	0,972	1,188	0,953	0,862	1,299	0,938	1,036
DMU4	0,943	1,016	1,000	1,105	0,956	1,005	1,242	0,950	1,027
DMU5	1,064	0,978	1,028	0,937	0,974	0,989	1,121	0,957	0,956
DMU6	1,161	0,980	1,014	1,194	1,003	0,993	1,501	0,990	0,982
DMU7	0,991	0,961	0,959	1,219	0,930	1,083	1,057	0,980	1,018
DMU8	1,176	0,976	1,005	1,360	0,937	1,015	0,948	0,970	0,991
DMU9	1,103	0,998	0,957	1,062	1,014	0,949	0,777	0,957	1,042
DMU10	0,994	0,987	0,970	1,077	0,955	0,962	0,924	0,931	1,126
DMU11	1,078	0,974	1,027	0,757	0,950	0,991	1,127	1,020	0,934
DMU12	0,902	0,948	1,022	1,432	0,974	0,989	1,334	1,049	1,007
DMU13	1,480	0,916	1,040	1,093	0,968	0,961	0,701	0,969	0,999
DMU14	1,188	0,969	0,957	1,178	0,987	0,959	1,601	1,008	0,949
DMU15	1,053	0,932	0,903	1,073	0,960	1,009	1,357	1,013	0,984
DMU16	0,983	0,903	1,079	1,094	0,988	1,024	1,080	0,983	0,953
DMU17	0,842	0,940	1,116	1,175	0,934	1,072	0,963	0,964	0,989
DMU18	1,171	0,899	1,012	1,450	1,013	0,959	1,124	1,005	1,049
DMU19	0,787	0,971	1,056	1,001	0,960	1,014	1,137	0,975	1,023
Gjennomsnitt	1,055	0,960	1,002	1,179	0,970	0,990	1,115	0,980	1,005
Standardavvik	0,150	0,033	0,050	0,207	0,025	0,051	0,227	0,030	0,045
Maks	1,480	1,016	1,116	1,733	1,014	1,083	1,601	1,049	1,126
Min	0,787	0,899	0,903	0,757	0,930	0,862	0,701	0,931	0,934

Tabell 13 viser den totale produktivetsendringen til hvert helseforetak gjennom fireårsperioden. For perioden 2008-2009 har ingen en ytelsesforbedring i alle tre modellene. Det samme gjelder for 2009-2010. To helseforetak har derimot i 2010-2011 forbedring i alle tre modellene. Dette kan forklares med at de fleste foretakene reduserer sin kostnadseffektivitet fra et år til et annet.

Fem helseforetak (DMU 3, 6 14, 15, 18) øker kvaliteten i alle tre periodene. Ingen helseforetak hatt en jevn økning i kostnadseffektivitet. Bare DMU 4 og 19 øker tekniske effektiviteten gjennom hele perioden. Samtidig er det ingen synlige relasjoner mellom tallene og de forskjellige DMU-ene. Det er dermed vanskelig å finne en klar relasjon.

**Tabell 14: Malmquist-effektivitetsendring (CRS)**

DMU	2008-2009(09)			2009-2010(10)			2010-2010(11)		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
DMU1	0,970	1,048	0,938	1,545	1,051	0,988	0,616	0,963	1,059
DMU2	0,933	0,954	0,891	1,127	1,008	1,082	0,743	1,020	0,967
DMU3	0,872	0,997	0,972	1,059	0,996	1,000	1,105	0,947	1,000
DMU4	1,000	1,052	0,976	1,000	0,999	1,000	1,000	0,959	1,025
DMU5	0,903	1,013	1,016	0,771	1,018	0,983	0,919	0,967	0,967
DMU6	1,156	1,014	1,000	1,111	1,048	1,000	1,000	1,000	1,000
DMU7	1,000	0,995	1,000	1,000	0,972	1,000	1,000	0,990	1,000
DMU8	1,000	1,010	1,000	1,000	0,980	1,000	1,000	0,979	1,000
DMU9	0,908	1,034	0,978	0,993	1,060	0,947	0,820	0,967	1,037
DMU10	0,900	1,022	0,979	1,023	0,998	0,967	0,709	0,940	1,124
DMU11	1,202	1,009	0,994	0,644	0,993	1,062	0,933	1,030	0,926
DMU12	0,826	0,981	0,999	1,211	1,018	1,007	1,000	1,059	1,004
DMU13	1,182	0,949	1,040	1,000	1,012	0,968	0,921	0,979	0,999
DMU14	1,131	1,003	0,913	1,043	1,032	1,010	1,301	1,018	0,950
DMU15	1,007	0,965	0,895	0,893	1,004	0,970	0,876	1,023	0,968
DMU16	0,925	0,935	1,072	0,938	1,032	1,031	0,879	0,992	0,958
DMU17	0,745	0,974	1,198	0,962	0,976	1,053	0,878	0,973	0,982
DMU18	1,101	0,931	1,026	1,302	1,059	1,038	0,913	1,014	1,037
DMU19	0,896	1,005	1,024	0,805	1,004	0,988	1,070	0,984	1,013
Gjennomsnitt	0,982	0,994	0,995	1,022	1,014	1,005	0,931	0,990	1,001
Standardavvik	0,122	0,035	0,066	0,191	0,026	0,034	0,148	0,030	0,043
Maks	1,202	1,052	1,198	1,545	1,060	1,082	1,301	1,059	1,124
Min	0,745	0,931	0,891	0,644	0,972	0,947	0,616	0,940	0,926

Indeksene ovenfor viser effektivitetsforbedringen relativt til hver modell sin front. Hver indeks er derfor påvirket av både det målte helseforetakets endring i ytelse, men også de beste innen hver modell sin endring i ytelse. Helseforetak med indeksen nøyaktig lik 1 ligger på modellens front i begge periodene.

Kun DMU 6 øker sin effektivitet i alle modellene til den ligger på fronten i 2010. Ellers er det en rekke helseforetak som øker scoren i alle modellene samtidig innenfor en periode, like mange synker i ytelse. Mesteparten av helseforetakene har sporadisk endring i ytelse.

**Tabell 15: Malmquist teknologisk endring (CRS)**

DMU	2008-2009(09)			2009-2010(10)			2010-2011(11)		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
DMU1	0,963	0,966	1,045	1,122	0,957	0,935	1,533	0,990	0,997
DMU2	1,216	0,966	1,047	1,131	0,957	0,978	1,279	0,990	1,000
DMU3	1,216	0,966	1,000	1,122	0,957	0,862	1,176	0,990	1,036
DMU4	0,943	0,966	1,025	1,105	0,957	1,005	1,242	0,990	1,002
DMU5	1,177	0,966	1,012	1,215	0,957	1,006	1,220	0,990	0,988
DMU6	1,004	0,966	1,014	1,074	0,957	0,993	1,501	0,990	0,982
DMU7	0,991	0,966	0,959	1,219	0,957	1,083	1,057	0,990	1,018
DMU8	1,176	0,966	1,005	1,360	0,957	1,015	0,948	0,990	0,991
DMU9	1,215	0,966	0,979	1,070	0,957	1,002	0,947	0,990	1,005
DMU10	1,104	0,966	0,991	1,053	0,957	0,996	1,303	0,990	1,002
DMU11	0,897	0,966	1,034	1,175	0,957	0,933	1,208	0,990	1,008
DMU12	1,093	0,966	1,023	1,182	0,957	0,982	1,334	0,990	1,003
DMU13	1,252	0,966	1,000	1,093	0,957	0,992	0,761	0,990	1,000
DMU14	1,050	0,966	1,048	1,130	0,957	0,950	1,231	0,990	1,000
DMU15	1,046	0,966	1,009	1,201	0,957	1,040	1,550	0,990	1,016
DMU16	1,062	0,966	1,006	1,167	0,957	0,993	1,228	0,990	0,995
DMU17	1,130	0,966	0,932	1,222	0,957	1,018	1,098	0,990	1,007
DMU18	1,064	0,966	0,987	1,114	0,957	0,924	1,231	0,990	1,011
DMU19	0,878	0,966	1,031	1,243	0,957	1,026	1,063	0,990	1,009
Gjennomsnitt	1,078	0,966	1,008	1,158	0,957	0,986	1,206	0,990	1,004
Standardavvik	0,110	0,000	0,029	0,073	0,000	0,048	0,197	0,000	0,012
Maks	1,252	0,966	1,048	1,360	0,957	1,083	1,550	0,990	1,036
Min	0,878	0,966	0,932	1,053	0,957	0,862	0,761	0,990	0,982

Tabell 15 viser teknologisk endring ved CRS, altså hvordan frontene til hver modell endrer seg relativt til hvert helseforetak. Som fellesnevner ser vi at indeksen for endring i kostnadsfronten er lik og under 1 for alle helseforetak. Dette kommer av at kun et helseforetak, en output og en input utgjør den kostnadseffektive fronten. Dette viser at det

mest kostnadseffektive helseforetaket også reduserer sin kostnadseffektivitet i løpet av fireårsperioden. Fronten for kvalitet er forbedrer seg relativt til de aller fleste helseforetak. Dette kan tolkes som at de helseforetak med allerede god kvalitet også forbedrer sin kvalitet. For fronten til teknisk effektivitet er endringene mer sporadiske, hvor teknisk effektivitet helseforetak både forbedrer og forverrer sin tekniske effektivitet.

**Tabell 16: Malmquist-effektivitetsendring (VRS)**

DMU	2008-2009(09)			2009-2010(10)			2010-2011(11)		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
DMU1	0,970	1,034	0,992	1,545	1,036	1,000	0,616	0,969	1,024
DMU2	0,933	0,942	0,892	1,127	1,005	1,101	0,743	1,076	0,970
DMU3	0,872	1,000	1,000	1,059	1,000	1,000	1,105	1,000	1,000
DMU4	1,000	1,004	1,000	1,000	1,000	1,000	1,000	1,000	1,000
DMU5	0,903	0,985	1,000	0,771	1,015	0,995	0,919	1,006	1,000
DMU6	1,156	1,010	1,000	1,111	1,026	1,000	1,000	1,000	1,000
DMU7	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
DMU8	1,000	0,997	1,000	1,000	0,975	1,000	1,000	1,000	1,000
DMU9	0,908	1,014	0,999	0,993	1,052	0,939	0,820	0,990	1,044
DMU10	0,900	1,038	0,991	1,023	0,997	1,000	0,709	0,944	1,113
DMU11	1,202	0,985	1,050	0,644	0,994	1,050	0,933	1,011	0,905
DMU12	0,826	1,000	1,041	1,211	1,025	1,038	1,000	1,093	0,987
DMU13	1,182	0,935	1,106	1,000	1,004	0,968	0,921	0,997	0,997
DMU14	1,131	0,979	0,943	1,043	1,022	0,995	1,301	1,014	0,988
DMU15	1,007	0,983	0,910	0,893	1,007	1,000	0,876	1,033	1,000
DMU16	0,925	1,000	1,000	0,938	1,000	1,000	0,879	1,000	1,000
DMU17	0,745	0,978	1,138	0,962	0,980	1,077	0,878	0,976	0,982
DMU18	1,101	0,907	1,065	1,302	1,057	1,036	0,913	1,047	1,057
DMU19	0,896	1,302	1,000	0,805	0,956	1,000	1,070	1,224	1,000
Gjennomsnitt	0,982	1,005	1,007	1,022	1,008	1,010	0,931	1,020	1,003
Standardavvik	0,122	0,077	0,057	0,191	0,024	0,036	0,148	0,059	0,039
Maks	1,202	1,302	1,138	1,545	1,057	1,101	1,301	1,224	1,113
Min	0,745	0,907	0,892	0,644	0,956	0,939	0,616	0,944	0,905

Tabell 16 viser hvert helseforetaks effektivitetsendring gitt variabelt skalautbytte. Som i dataomhyllingsanalysene er ikke kvalitetsvariablene under forutsetning av variabelt skalautbytte ettersom kvalitetsscorene er beregnet etter relative verdier. I denne tabellen får flere helseforetak en indeks på 1 enn de gjorde ved effektivitetsendring gitt konstant

skalautbytte (tabell 14). Dette forklares av et større antall effektive helseforetak i effektivitetsmodellene.

Tabellen viser ingen synlige relasjoner mellom kvalitet og effektivitet. Selv om flere helseforetak har ingen eller positiv endring i både kvalitets- og effektivitetsscorene, har en betydelig mengde helseforetak både positiv og negativ endring.

**Tabell 17: Malmquist teknologisk endring (VRS)**

DMU	2008-2009(09)			2009-2010(10)			2010-2011(11)		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
DMU1	0,963	0,979	1,029	1,122	0,973	0,918	1,533	0,989	1,033
DMU2	1,216	0,983	1,035	1,131	0,967	0,971	1,279	0,966	1,004
DMU3	1,216	0,973	1,053	1,122	0,983	0,897	1,176	0,94	1,039
DMU4	0,943	1,013	0,998	1,105	0,968	1,012	1,242	0,958	1,033
DMU5	1,177	0,999	0,995	1,215	0,961	0,991	1,22	0,957	0,976
DMU6	1,004	0,972	1,03	1,074	0,977	0,988	1,501	0,991	0,982
DMU7	0,991	0,959	0,969	1,219	0,929	1,093	1,057	0,98	1,021
DMU8	1,176	0,982	0,996	1,36	0,967	1,024	0,948	0,975	0,989
DMU9	1,215	0,986	0,955	1,07	0,966	1,015	0,947	0,973	1,002
DMU10	1,104	0,941	0,995	1,053	0,93	0,958	1,303	0,977	1,032
DMU11	0,897	1,000	1,006	1,175	0,971	0,951	1,208	1,015	1,016
DMU12	1,093	0,935	0,981	1,182	0,934	0,954	1,334	0,975	0,996
DMU13	1,252	0,983	0,97	1,093	0,968	0,993	0,761	0,976	1,006
DMU14	1,05	1,002	1,026	1,13	0,972	0,973	1,231	1,002	0,978
DMU15	1,046	0,935	0,997	1,201	0,934	1,046	1,55	0,975	1,004
DMU16	1,062	*1,000	*1,000	1,167	*1,000	*1,000	1,228	*1,000	*1,000
DMU17	1,13	0,955	0,994	1,222	0,951	0,995	1,098	0,981	1,005
DMU18	1,064	0,993	0,979	1,114	0,961	0,919	1,231	0,962	1,014
DMU19	0,878	*1,000	*1,000	1,243	0,967	*1,000	1,063	0,805	*1,000
Gjennomsnitt	1,078	0,978	1,000	1,158	0,962	0,984	1,206	0,968	1,007
Standardavvik	0,110	0,023	0,025	0,073	0,018	0,045	0,197	0,042	0,018
Maks	1,252	1,013	1,053	1,360	1,000	1,093	1,550	1,015	1,039
Min	0,878	0,935	0,955	1,053	0,929	0,897	0,761	0,805	0,976

Tabell 17 viser teknologisk endring relativt til hvert helseforetak med forutsetning for variabelt skalautbytte. Også her er kvalitetsmålene forutsatt konstant skalautbytte. Den kostnadseffektive fronten er ikke lik for alle helseforetakene fordi flere helseforetak påvirker

den effektive fronten i motsetning til ett helseforetak ved konstant skalautbytte.

Forutsetningen om variabelt skalautbytte resulterer også i at den teknologiske endringen ikke kan beregnes relativt til DMU 16 og 19 ettersom disse ligger i enden av utvalget.

Fronten for kostnadseffektivitet har en jevn nedgang over fireårsperioden, noe som stemmer godt med den totale Malmquist-produktivitetsindeksen. Front for teknisk effektivitet har en ujevn endring mens kvalitetsfronten forbedrer seg jevnt over.

**Tabell 18: Malmquist-skalaendring**

DMU	2008-2009(09)		2009-2010(10)		2010-2011(11)	
	Kostnad	Teknisk	Kostnad	Teknisk	Kostnad	Teknisk
DMU1	0,998	0,961	0,998	1,006	0,995	0,998
DMU2	0,992	1,010	0,992	0,990	0,971	0,993
DMU3	0,970	0,923	0,970	0,960	0,997	0,997
DMU4	0,988	1,002	0,988	0,993	0,992	0,994
DMU5	0,998	1,033	0,998	1,003	0,994	0,979
DMU6	1,000	0,984	1,000	1,005	0,999	0,999
DMU7	1,002	0,990	1,002	0,991	1,001	0,997
DMU8	0,994	1,009	0,994	0,991	0,994	1,002
DMU9	0,998	1,003	0,998	0,996	0,994	0,997
DMU10	1,029	0,984	1,029	1,004	1,009	0,980
DMU11	0,985	0,973	0,985	0,993	0,993	1,016
DMU12	1,017	1,001	1,017	0,998	0,984	1,025
DMU13	0,996	0,970	0,996	1,000	0,996	0,996
DMU14	0,994	0,989	0,994	0,990	0,992	0,982
DMU15	1,021	0,996	1,021	0,964	1,006	0,980
DMU16	*1,000	*1,000	*1,000	*1,000	*1,000	*1,000
DMU17	1,002	0,987	1,002	1,001	1,006	1,002
DMU18	0,998	0,971	0,998	1,007	0,998	0,978
DMU19	1,038	*1,000	1,038	*1,000	0,990	*1,000
Gjennomsnitt	1,001	0,989	1,001	0,994	0,995	0,996
Standardavvik	0,015	0,023	0,015	0,012	0,008	0,012
Maks	1,038	1,033	1,038	1,007	1,009	1,025
Min	0,970	0,923	0,970	0,960	0,971	0,978

Tabell 17 illustrerer hvert helseforetaks endring i skalaeffektivitet  $\left( = \frac{CRS\text{-effektivitet}}{VRS\text{-effektivitet}} \right)$ , og viser kun til kostnadsmodell og teknisk modell ettersom kun disse har reelle verdier gitt

variabelt skalautbytte. DMU 16 og 19 har ikke fått beregnet en skalaeffektivitetsscore fordi helseforetakene er ytterligheter i både kostnads- og teknisk effektivitetsmodell.

### 6.1.3. Sensitivitetsanalyse

Sensitiviteten til effektivitetsestimaterne analyseres ved hjelp av Edwardsens (2004) effektivitetsstige-metode, som forklart i kapittel 4.6.3. Analysen er utført på hver DMU i datasettet for både teknisk, kostnad og kvalitet, for alle år i perioden. Framfor å tolke resultatene via den grafiske effektivitetsstigen, har vi notert ned de beregnede effektivitetstallene for de tre første stegene. Deretter har vi beregnet den totale gjennomsnittlige endringen i løpet av ett, to og tre steg. Høye effektivitetsendringer antyder høy sensitivitet for feilmålinger og bortfall av referanseenheter. Estimaterne er beregnet ved hjelp av en CRS-modell.

**Tabell 19: Sensitivitetsanalyse for teknisk effektivitet**

Steg/år		2008(09)	2009(09)	2009(10)	2010(10)	2010(11)	2011(11)
Ett steg	Snittendring	0,037	0,024	0,036	0,042	0,051	0,038
	Standardavvik	0,026	0,028	0,038	0,048	0,050	0,048
To steg	Snittendring	0,036	0,023	0,028	0,031	0,033	0,034
	Standardavvik	0,025	0,024	0,036	0,039	0,042	0,035
Tre steg	Snittendring	0,032	0,022	0,026	0,024	0,026	0,028
	Standardavvik	0,022	0,022	0,031	0,034	0,037	0,030

**Tabell 20: Sensitivitetsanalyse for kostnadseffektivitet**

Steg/år		2008(09)	2009(09)	2009(10)	2010(10)	2010(11)	2011(11)
Ett steg	Snittendring	0,012	0,004	0,038	0,023	0,001	0,010
	Standardavvik	0,001	0,001	0,006	0,003	0,000	0,001
To steg	Snittendring	0,009	0,005	0,025	0,033	0,009	0,018
	Standardavvik	0,003	0,001	0,013	0,011	0,008	0,008
Tre steg	Snittendring	0,012	0,004	0,021	0,023	0,011	0,023
	Standardavvik	0,005	0,001	0,012	0,016	0,007	0,010

Resultatene fra sensitivitetsanalysen viser lave endringer for både kostnadseffektivitet og teknisk effektivitet. Høyeste endring er på 5,1 prosentpoeng i løpet av første steg, 2010(11).

De fleste endringen er langt under dette nivået, slik at den totale gjennomsnittsendringen er rundt 2 prosentpoeng. Estimaten virker å være svært robuste.

**Tabell 21: Sensitivitetsanalyse for kvalitet**

Steg/år		2008	2009	2010	2011
Ett steg	Snittendring	0,140	0,142	0,203	0,154
	Standardavvik	0,173	0,199	0,383	0,173
To steg	Snittendring	0,127	0,132	0,153	0,124
	Standardavvik	0,173	0,171	0,281	0,140
Tre steg	Snittendring	0,129	0,127	0,127	0,135
	Standardavvik	0,171	0,165	0,236	0,157

Kvalitetsestimaten har en gjennomsnittlige effektivitetsendringer mellom 10 og 20 prosentpoeng i løpet av første steg. Selv om endringene synker gradvis etter flere ledd, fortsetter endringene å ligge rundt 10 prosentpoeng. Ingen enheter eller årstall skiller seg spesielt ut. De fleste enhetene blir effektive i løpet av de tre første stegene. Siden modellen har et relativt høyt antall variabler, og siden ingen av variablene korrelerer, kan vi forvente mange effektive enheter. Frafall av referanseenheter, vil da i løpet av få steg føre til nye effektive enheter og nokså høye endringer. Resultatene viser dermed at estimaten er temmelig sensitive til feilmålingen og for fjerning av referanseenheter.

## 6.2. Trinn 2: Statistisk analyse

Det skilles fortsatt mellom resultatene fra dataomhyllingsanalysene og Malmquist-indeksene. Dataomhyllingsanalysene er videre inndelt i resultater gitt konstant skalautbytte og resultater gitt variabelt skalautbytte. For Malmquist-indeksene inndeler vi regresjonsanalysene i den totale ytelsesendringen, effektivitetsendring gitt konstant skalautbytte og effektivitetsendring gitt variabelt skalautbytte.

For å ha kunne utføre sikrere statistiske tester, er det ønskelig å ha flere observasjoner enn de 19 helseforetakene vi måler. I stedet for å teste hvert år for seg, blir perioden slått sammen til en virtuell periode. Det medfører da dataomhyllingsanalyseresultatene får 76 observasjoner, mens resultatene fra Malmquist har 57 observasjoner. Dette lar seg gjøre siden estimaten for kvalitet og effektivitet alle er relative verdier som etter å ha blitt beregnet i sitt tilhørende år,



er sammenlignbare med tidligere år. Under følger en tabell med deskriptive statistikk for hvert av de nye datasettene.

**Tabell 22: Deskriptiv statistikk for DEA-resultater**

	CRS			VRS	
	Kvalitet	Kostnad	Teknisk	Kostnad	Teknisk
Gjennomsnitt	0,94	0,83	0,93	0,97	1,05
Standardavvik	0,37	0,11	0,12	0,29	0,27
Minimum	0,50	0,55	0,68	0,67	0,77
Maximum	2,77	1,03	1,18	2,22	2,22
Antall DMU	76	76	76	76	76

**Tabell 23: Deskriptiv statistikk for MPI-resultater**

	Produktivitetsendring			Effektivitetsendring (CRS)			Effektivitetsendring (VRS)		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
Gjennomsnitt	1,12	0,97	1,00	0,98	1,00	1,00	0,98	1,01	1,01
Standardavvik	0,21	0,03	0,05	0,16	0,03	0,05	0,16	0,06	0,05
Minimum	0,70	0,90	0,86	0,62	0,93	0,89	0,62	0,91	0,89
Maximum	1,73	1,05	1,13	1,54	1,06	1,20	1,54	1,30	1,14
Antall DMU	57	57	57	57	57	57	57	57	57

### 6.2.1. Gruppendifferanseanalyser

Gruppendifferanseanalysene utført under viser statistiske differanser mellom de helseforetak som ligger på kvalitetsfronten, kostandseffektivitetsfronten og teknisk effektivitetsfront. Formålet med disse analysene er å se hvorvidt helseforetak med best kvalitet har gjennomsnittlig bedre eller verre effektivitet, og om helseforetak med best effektivitet har gjennomsnittlig bedre eller dårligere kvalitet. I tabellene presenteres også standardavviket for de forskjellige gruppene gjennomsnitt. Til slutt presenterer tabellene en p-verdi beregnet igjennom en «Kolmogorov-Smirnov»-test. P-verdier under 0,05 viser til signifikante forskjeller mellom de to gruppene som sammenliknes.

## Gruppedifferanseanalyser på dataomhyllingsanalyser

**Tabell 24: Gruppedifferanseanalyseresultater - kvalitet (CRS)**

Kvalitet (CRS)	CRS			VRS	
	Kvalitet	Kostnad	Teknisk	Kostnad	Teknisk
Score > 1	1,372	<b>0,905</b>	<b>0,998</b>	<b>0,952</b>	<b>1,057</b>
Standardavvik	0,367	<b>0,108</b>	<b>0,117</b>	<b>0,092</b>	<b>0,086</b>
Score < 1	0,750	<b>0,803</b>	<b>0,900</b>	<b>0,984</b>	<b>1,049</b>
Standardavvik	0,116	<b>0,097</b>	<b>0,104</b>	<b>0,335</b>	<b>0,311</b>
Differanse	0,622	<b>0,102</b>	<b>0,098</b>	<b>-0,033</b>	<b>0,008</b>
P-verdi	0,000	<b>0,000</b>	<b>0,000</b>	<b>0,018</b>	<b>0,002</b>

Gitt konstants skalautbytte har DMU-ene som scorer høyt på kvalitet gjennomsnittlig 10 prosentpoeng høyere effektivitetsscore enn de som scorer under 1 i kvalitet. Ved variabelt skalautbytte er det ingen markant forskjell mellom effektivitetstallene til de som scorer høyt og lavt på kvalitet; -3,3 prosentpoeng for kostnadseffektivitet og nesten 0,8 prosent for teknisk effektivitet. «Kormogorov-Smirnov»-testen viser at helseforetakene med best og dårligst kvalitet er signifikant forskjellig for begge effektivitetsmodellene ved CRS og VRS.

**Tabell 25: Gruppedifferanseanalyseresultater - effektivitet (CRS)**

	CRS			VRS	
	Kvalitet	Kostnad	Teknisk	Kostnad	Teknisk
<b>Kostnad (CRS)</b>					
Score > 1	<b>1,140</b>	1,015	1,111	1,024	1,137
Standardavvik	<b>0,169</b>	0,009	0,041	0,012	0,046
Score < 1	<b>0,927</b>	0,823	0,919	0,972	1,047
Standardavvik	<b>0,368</b>	0,105	0,112	0,292	0,270
Differanse	<b>0,213</b>	0,192	0,191	0,053	0,090
P-verdi	<b>0,104</b>	0,001	0,009	0,123	0,384
<b>Teknisk effektivitet (CRS)</b>					
Score > 1	<b>1,182</b>	0,927	1,078	1,152	1,247
Standardavvik	<b>0,454</b>	0,104	0,052	0,382	0,340
Score < 1	<b>0,851</b>	0,800	0,876	0,911	0,982
Standardavvik	<b>0,277</b>	0,092	0,084	0,207	0,187
Differanse	<b>0,331</b>	0,127	0,201	0,241	0,265
P-verdi	<b>0,000</b>	0,000	0,000	0,000	0,000

Helseforetak som scorer over 1 på kostnadseffektivitet gitt konstant skalautbytte, har gjennomsnittlig 21,3 prosentpoeng høyere kvalitet. Differansen er ikke signifikant, trolig på grunn av lavt antall kostnadseffektive observasjoner. De høystscorende på teknisk effektivitet gitt konstant skalautbytte, scorer 33,1 prosentpoeng høyere på kvalitet.

**Tabell 26: Gruppendifferanseanalyseresultater - effektivitet (VRS)**

	CRS		VRS		
	Kvalitet	Kostnad	Teknisk	Kostnad	Teknisk
<b>Kostnads-effektivitet (VRS)</b>					
Score > 1	<b>0,987</b>	0,870	0,991	1,210	1,243
Standardavvik	<b>0,269</b>	0,133	0,153	0,399	0,386
Score < 1	<b>0,920</b>	0,812	0,896	0,866	0,962
Standardavvik	<b>0,395</b>	0,096	0,087	0,087	0,091
Differanse	<b>0,066</b>	0,058	0,096	0,344	0,281
P-verdi	<b>0,000</b>	0,048	0,000	0,000	0,000
<b>Teknisk effektivitet (VRS)</b>					
Score > 1	<b>1,021</b>	0,859	0,971	1,069	1,158
Standardavvik	<b>0,408</b>	0,124	0,133	0,338	0,302
Score < 1	<b>0,824</b>	0,799	0,872	0,844	0,905
Standardavvik	<b>0,248</b>	0,077	0,052	0,083	0,058
Differanse	<b>0,196</b>	0,060	0,100	0,224	0,253
P-verdi	<b>0,061</b>	0,001	0,000	0,000	0,000

Ved variabel skalautbytte har de kostnadseffektive enhetene en kvalitetsscore som er 6,6 prosentpoeng høyere enn de lavscorende på kostnadseffektivitet. Forskjellene mellom gruppene er signifikante. DMU-ene som scorer over 1 på teknisk effektivitet, gitt variabelt skalautbytte, har 19,6 prosentpoeng høyere kvalitet enn de resterende. Disse resultatene er kun signifikante ved konfidensnivå på 90 %.

## Gruppedifferanseanalyser på Malmquist-produktivtetsindeks

I de videre tabellene presenteres gjennomsnittsanalysene av Malmquist produktivtetsindeks og Malmquist effektivtetsendring for CRS og VRS i hver sin tabell. I motsetning til dataomhyllingsanalysene skilles nå mellom helseforetak som har forbedret og forverret sin kvalitet og effektivitet. Også her er gjennomsnittene og standardavvikene beregnet etter det totale antall endringer i fireårsperioden ( $n = 57$ ). Matrisen under er også inndelt i de som har forbedret og forverret seg i hver av de tre modellene.

**Tabell 27: Gruppedifferanseanalyseresultater - Malmquist-produktivtetsendring**

Referanse:	KVALITET			KOSTNAD			TEKNISK		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
Score > 1	1,209	<b>0,972</b>	<b>0,990</b>	<b>1,234</b>	1,014	<b>0,975</b>	<b>1,097</b>	<b>0,960</b>	1,038
Standardavvik	0,165	<b>0,032</b>	<b>0,046</b>	<b>0,253</b>	0,011	<b>0,033</b>	<b>0,170</b>	<b>0,032</b>	0,032
Score < 1	0,899	<b>0,966</b>	<b>1,020</b>	<b>1,085</b>	0,958	<b>1,005</b>	<b>1,135</b>	<b>0,980</b>	0,961
Standardavvik	0,088	<b>0,028</b>	<b>0,050</b>	<b>0,176</b>	0,023	<b>0,051</b>	<b>0,230</b>	<b>0,026</b>	0,029
Differanse	0,310	<b>0,006</b>	<b>-0,030</b>	<b>0,149</b>	0,056	<b>-0,031</b>	<b>-0,038</b>	<b>-0,021</b>	0,077
P-verdi	0,000	<b>0,987</b>	<b>0,088</b>	<b>0,279</b>	0,000	<b>0,236</b>	<b>0,084</b>	<b>0,213</b>	0,000

Tabellen over viser gjennomsnittresultatene for den totale Malmquist-produktivtetsindeksen. I tilfellene med forbedret kvalitet har helseforetakene gjort det gjennomsnittlig litt bedre på kostnadsproduktivitet og verre på teknisk produktivitet enn helseforetak som forverret sin kvalitet. Samme mønster finner vi også med kostnadsproduktivitet og teknisk produktivitet. Vi ser også at differansen er større for kvalitetsgjennomsnittene i forhold til både kostnads- og teknisk modell. Samtidig er det ingen spesielt signifikante gruppeforskjeller når vi skiller mellom helseforetak som har forbedret og forverret seg. Setter vi konfidensnivået til 90 % finnes to unntak. Helseforetak som forbedrer sin kvalitet gjør det gjennomsnittlig litt verre på teknisk produktivitet enn helseforetak som forverrer sin kvalitet. Helseforetak som forbedrer sin tekniske produktivitet vil gjennomsnittlig forverre sin kvalitet opp mot helseforetak som forverrer sin tekniske produktivitet.

**Tabell 28: Gruppendifferanseanalyseresultater - Malmquist-effektivitetsendring (CRS)**

Referanse:	KVALITET			KOSTNAD			TEKNISK		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
Score > 1	1,092	<b>1,001</b>	<b>0,996</b>	<b>1,019</b>	1,024	<b>0,989</b>	<b>0,986</b>	<b>0,992</b>	1,030
Standardavvik	0,125	<b>0,033</b>	<b>0,035</b>	<b>0,170</b>	0,019	<b>0,035</b>	<b>0,152</b>	<b>0,033</b>	0,042
Score < 1	0,860	<b>0,997</b>	<b>1,005</b>	<b>0,933</b>	0,972	<b>1,013</b>	<b>0,969</b>	<b>1,009</b>	0,963
Standardavvik	0,094	<b>0,031</b>	<b>0,062</b>	<b>0,136</b>	0,019	<b>0,060</b>	<b>0,170</b>	<b>0,028</b>	0,029
Differanse	0,232	<b>0,004</b>	<b>-0,008</b>	<b>0,086</b>	0,052	<b>-0,023</b>	<b>0,016</b>	<b>-0,017</b>	0,067
P-verdi	0,000	<b>0,987</b>	<b>0,087</b>	<b>0,279</b>	0,000	<b>0,236</b>	<b>0,084</b>	<b>0,213</b>	0,000

De gjennomsnittlige Malmquist effektivitetsendringene for konstant skalautbytte vises i tabellen ovenfor. Her ser vi samme tendensen som i den totale produktivitetsindeksen. De som forbedrer sin kvalitet i forhold til de beste forbedrer også sin kostnadseffektivitet og forverrer sin tekniske effektivitet i forhold til de beste i disse modellene. Også her er standardavviket betydelig større enn de gjennomsnittlige differansene. Setter vi referansen til teknisk modell, ser vi at de som forbedrer sin tekniske effektivitet også har gjennomsnittlig bedre endring i kvalitet relativt til de beste. Dette henter også til uklare forhold mellom teknisk effektivitetsendring og kvalitetsendring relativt til de beste innen kvalitet.

**Tabell 29: Gruppendifferanseanalyseresultater - Malmquist-effektivitetsendring (VRS)**

Referanse:	KVALITET			KOSTNAD			TEKNISK		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
Score > 1	1,092	<b>1,010</b>	<b>1,006</b>	<b>1,005</b>	1,030	<b>1,000</b>	<b>0,984</b>	<b>1,009</b>	1,023
Standardavvik	0,125	<b>0,052</b>	<b>0,038</b>	<b>0,151</b>	0,060	<b>0,030</b>	<b>0,171</b>	<b>0,064</b>	0,037
Score < 1	0,860	<b>1,012</b>	<b>1,007</b>	<b>0,925</b>	0,972	<b>1,020</b>	<b>0,965</b>	<b>1,016</b>	0,966
Standardavvik	0,094	<b>0,064</b>	<b>0,051</b>	<b>0,166</b>	0,024	<b>0,063</b>	<b>0,127</b>	<b>0,037</b>	0,035
Differanse	0,232	<b>-0,003</b>	<b>-0,001</b>	<b>0,079</b>	0,058	<b>-0,020</b>	<b>0,019</b>	<b>-0,007</b>	0,057
P-verdi	0,000	<b>0,905</b>	<b>0,905</b>	<b>0,344</b>	0,000	<b>0,064</b>	<b>0,821</b>	<b>0,028</b>	0,000

Gjennomsnittlige Malmquist effektivitetsendringer ved variabelt skalautbytte viser også relativt vage forskjeller mellom de beste og dårligste endringene med hensyn til hver modell. De med best endring i kvalitet relativt til kvalitetsfronten har dårligere effektivitetsendring i både kostnads- og teknisk modell. Likevel ser vi at de med positiv endring i kostnadseffektivitet og teknisk effektivitet har gjennomsnittlig bedre kvalitetsendring relativt til fronten i forhold til de med nedgang i kostnadseffektivitet og teknisk effektivitet. Igjen er

standardavvikene betydelig større enn differansene av gjennomsnittene, og forskjellene er derfor svært vage.

### 6.2.2. Korrelasjon

I tabellene under presenteres korrelasjonskoeffisientene mellom modellene for kvalitet, kostnadseffektivitet og teknisk effektivitet. Tabellene beskriver samvariasjonene mellom kvalitet og effektivitet ut over skillene mellom de gode og dårlige innenfor hver modell.

**Tabell 30: Korrelasjonsanalyse DEA**

	CRS			VRS		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
Kvalitet	1,000			1,000		
Kostnad	0,360	1,000		-0,067	1,000	
Teknisk	0,306	0,677	1,000	-0,029	0,947	1,000

Ved konstant skalautbytte har kvalitet svak positiv korrelasjon over 0,3 for både teknisk effektivitet og kostnadseffektivitet. Tar man hensyn til variabelt skalautbytte er korrelasjon nesten fraværende. Ser at samvariasjonen mellom kostnadseffektivitet og teknisk effektivitet er positiv i begge tilfellene; ved konstant skalautbytte 0,667, med variabelt skalautbytte 0,947.

**Tabell 31: Korrelasjonsanalyse MPI**

	Produktivitetsendring			Effektivitetsendring CRS			Effektivitetsendring VRS		
	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk	Kvalitet	Kostnad	Teknisk
Kvalitet	1,000			1,000			1,000		
Kostnad	0,192	1,000		0,281	1,000		0,073	1,000	
Teknisk	-0,249	-0,357	1,000	-0,221	-0,285	1,000	-0,066	-0,160	1,000

Korrelasjonsanalysen for total Malmquist-produktivitetsindeks viser en svak positiv korrelasjon på 0,192 mellom kvalitet og kostnadseffektivitet, og en svak negativ korrelasjon mellom kvalitet og teknisk effektivitet på -0,249. Det er også en svak, negativ korrelasjon mellom kostnadseffektivitet og teknisk effektivitet på -0,357. Korrelasjonsanalysen for effektivitetsendringer gitt konstant skalautbytte viser også en svak positiv korrelasjon mellom kvalitet og kostnadseffektivitet på 0,281. Teknisk effektivitetsendring korrelerer svakt og negativt med både kvalitet (-0,221) og kostnadseffektivitetsendring (-0,285). Gitt variabelt skalautbytte ser man en svært svak positiv korrelasjon mellom kvalitet og kostnadseffektivitet

på 0,073. Teknisk effektivitetsendring korrelerer svært svakt og negativt med kvalitet (-0,066) og svakt negativt med kostnadseffektivitetsendringen (-0,160).

### 6.2.3. Regresjon

I regresjonsanalysene søker å finne en signifikant årsakssammenheng mellom prosesskvalitet, kostnadseffektivitet og teknisk effektivitet. Ettersom årsakssetning ikke er kjent, kjøres regresjonen med både kvalitet og effektivitet som forklarende variabler.

Følgende modeller testes med kvalitet som avhengig variabel:

$$1. \text{ Kvalitet} = \beta_0 + \beta_1 * \text{kostnad} + \beta_2 * \text{teknisk} + \varepsilon$$

Følgende modeller testes med effektivitetsbegrepene som avhengig variabel:

$$1. \text{ Teknisk effektivitet} = \beta_0 + \beta_1 * \text{kvalitet} + \varepsilon$$

$$2. \text{ Kostnadseffektivitet} = \beta_0 + \beta_1 * \text{kvalitet} + \varepsilon$$

Videre utføres begge settene med tester på både resultatene fra dataomhyllingsanalysen og Malmquist-produktivitetsindeks. Likedan som for korrelasjonsanalysen analyseres et sett for hele analyseperioden for å få et mer solid grunnlag for de statistiske beregningene.

### Regresjon på dataomhyllingsanalyser

**Modell 1:**  $\text{Kvalitet} = \beta_0 + \beta_1 * \text{kostnad} + \beta_2 * \text{teknisk} + \varepsilon$

**Tabell 32: Regresjon DEA modell-1 CRS**

Regresjon		Koeffisient	$\beta_n$	P-verdi
R <sup>2</sup>	0,137	Skjæringspunkt	-0,164	0,627
R <sub>A</sub> <sup>2</sup>	0,113	Kostnad	0,929	0,060
Standardfeil <sup>10</sup>	0,344	Teknisk	0,353	0,443
P-verdi	0,005			

<sup>10</sup> Standardfeilen beregnes som  $s_\varepsilon = \sqrt{\frac{SSE}{n-k-1}}$ , og måler de observerte verdienes spredningen rundt de estimerte verdiene til regresjonsmodellen. Hvor intervallet  $\mp 2s_\varepsilon$  bør vurderes med et praktisk skjønn. Estimater med for høy standardfeil, bør forkastes selv om de har signifikante p-verdier. (Groebner et al., 2008)

Den justerte forklaringsgrader er på 11,3% og modellen er signifikant med en p-verdi på 0,005. Dette forteller oss at 11,3% av variasjonen i kvalitet kan totalt forklares av variasjonen til kostnadseffektivitet og teknisk effektivitet. Av koeffisientene ser man at det er kostnadseffektiviteten som har en signifikant positiv innvirkning på kvaliteten. Teknisk effektivitet har noe lavere positiv innvirkning, men dette forholdet er ikke signifikant. Til tross for middels sterk korrelasjon mellom de forklarende variablene, ansees ikke multikollinearitet å være et problem siden beregnet VIF<sup>11</sup> er på 1,85. Standardfeilen ansees derimot som svært høy.

Ved variabelt skalautbytte er korrelasjon mellom de forklarende variablene for høy til at denne modellen gir gyldige resultater. Tilnærmet perfekt korrelasjon og en beregnet VIF lik 9,62, antyder store problemer med multikollinearitet. I stedet for testes to modeller med enkel regresjon, hvor kvalitet er avhengigvariabel og kostnadseffektivitet og teknisk effektivitet utgjør forklarende variabel hver sin gang.

**Tabell 33: Regresjon DEA modell-1 VRS**

Regresjon	Kvalitet-Kost	Kvalitet-Teknisk
R <sup>2</sup>	0,067	0,029
R <sub>A</sub> <sup>2</sup>	0,004	0,001
Standardfeil	0,367	0,368
P-verdi	0,565	0,803

Resultantene viser at det er tilnærmet lik null forklaringsgrad mellom kvalitet og effektivitetsbegrepene ved variabelt skalautbytte. Ingen signifikante årsakssammenhenger.

**Modell 2:**  $Kostnadseffektivitet = \beta_0 + \beta_1 * kvalitet + \varepsilon$

**Tabell 34: Regresjon DEA modell-2**

Regresjon	CRS	VRS
R <sup>2</sup>	0,004	0,001
Standardfeil	0,367	0,368
P-verdi	0,565	0,803
Koeffisient	-0,085	-0,040

<sup>11</sup> VIF («variance inflation factor») =  $\frac{1}{1-R_j^2}$ . Hvor R<sub>j</sub><sup>2</sup> er forklaringsgraden til en regresjonsanalyse hvor uavhengig variabel j er analysert mot de gjenværende k – 1 uavhengige variablene. Generelt indikerer VIF ≥ 5 for høy multikollinearitet, og variabelen bør ekskluderes fra modellen. (Groebner et al., 2008)



Variasjonen i kvalitetsscoren forklarer tilnærmet null av variasjonen til kostnadseffektiviteten.

**Modell 3:**  $Teknisk\ effektivitet = \beta_0 + \beta_1 * kvalitet + \varepsilon$

**Tabell 35: Regresjon DEA-4**

Regresjon	CRS	VRS
R <sup>2</sup>	0,093	0,001
Standardfeil	0,113	0,368
P-verdi	0,007	0,803
Koeffisienter	0,099	-0,040

Kvalitet forklarer 9,3% av teknisk effektivitet, gitt konstant skalautbytte, og dette er et signifikant forhold. Ved variabelt skalautbytte derimot, er forklaringsgraden tilnærmet lik null.

### Regresjon på Malmquist-produktivitetsindeks

**Modell 1:**  $Kvalitet = \beta_0 + \beta_1 * kostnad + \beta_2 * teknisk + \varepsilon$

**Tabell 36: Regresjon MPI-1**

Regresjon	Produktivitet	Effektivitet (CRS)	Effektivitet (VRS)
R <sup>2</sup>	0,074	0,101	0,008
R <sub>A</sub> <sup>2</sup>	0,040	0,067	-0,028
Standardfeil	0,201	0,156	0,164
P-verdi	0,126	0,057	0,800

	Produktivitet		Effektivitet (CRS)		Effektivitet (VRS)	
	$\beta_n$	p-verdi	$\beta_n$	p-verdi	$\beta_n$	p-verdi
Skjæringspunkt	1,216	0,328	0,295	0,742	1,000	0,144
Kostnad	0,778	0,404	1,179	0,083	0,177	0,645
Teknisk	-0,855	0,146	-0,494	0,258	-0,199	0,688

Produktivitetsendringen og effektivitetsendringen ved variabelt skalautbytte, har ingen signifikant forklaringsgrad på kvalitetsendringen. Heller ingen koeffisienter har et signifikant forhold. Derimot viser effektivitetsendringen ved konstant skalautbytte å ha en 6,7 %

forklaringsgrad, som er signifikant ved et 90 % konfidensnivå. Dette skyldes en positiv innvirkningen fra kostnadskoeffisienten som er signifikant ved et 90 % konfidensnivå.

**Modell 2:**  $Kostnadseffektivitet = \beta_0 + \beta_1 * kvalitet + \varepsilon$

**Tabell 37: Regresjon MPI-2**

Regresjon	Produktivitet	Effektivitet (CRS)	Effektivitet (VRS)
R <sup>2</sup>	0,037	0,079	0,005
Standardfeil	0,019	0,062	-0,013
P-verdi	0,031	0,032	0,059
Koeffisient	0,029	0,057	0,026

Også her har kvalitet en signifikant forklaringsgrad til kostnadseffektivitet, men den er lav. Høyeste verdi ved effektivitetsendringen ved konstant skalautbytte er 7,9 %.

**Modell 3:**  $Teknisk effektivitet = \beta_0 + \beta_1 * kvalitet + \varepsilon$

**Tabell 38: Regresjon MPI-3**

Regresjon	Produktivitet	Effektivitet (CRS)	Effektivitet (VRS)
R <sup>2</sup>	0,062	0,049	0,004
Standardfeil	0,049	0,050	0,046
P-verdi	0,062	0,098	0,628
Koeffisient	-0,060	-0,069	-0,018

Kvalitetsendringen har en svak forklaringsgrad på endringen i teknisk effektivitet, men den er signifikant for både produktivitetsendring og effektivitetsendringen gitt konstant skalautbytte.

## 7. Diskusjon og konklusjon

Dette kapittelet er dedikert til diskusjon av resultatene. Vi skal først se på viktige momenter ved metoden som kan ha påvirket resultatene, slik at svakheter og styrker ved den valgte metoden blir belyst. Deretter vil vi svare på problemstillingen ved hjelp av studiets resultater, og diskutere våre resultater opp mot andre studiers resultater. Kapittelet ender i en kort konklusjon av oppgaven.

### 7.1. Oppsummering og viktige momenter

Resultatene av dataomhyllingsanalysene viser en svak relasjon mellom kvalitet og effektivitet; enkelte helseforetak ligger på begge frontene samtidig, men det er ingen klar trend. Gruppedifferanseanalysene viser at det er en signifikant forskjell på differansene mellom helseforetak som gjør det godt og dårlig i de forskjellige modellene; differansene er imidlertid lave. Korrelasjonsanalysene bekrefter den svake positive samvariasjonen mellom kvalitet og effektivitet.

I gjennomgangen av Malmquist-produktivitetsindeksene har vi vanskelig for å finne noen som helst relasjon mellom kvalitet og effektivitet. Gruppedifferanseanalysene gav få signifikante relasjoner. Differansen var negativ i alle tilfeller, men nærmest ubetydelig. Den svake relasjonen bekreftes av korrelasjonskoeffisientene. Kvaliteten korrelerte positivt med kostnadseffektiviteten og negativt med teknisk effektivitet.

I søken etter direkte årsakssammenhenger mellom kvalitet og effektivitet ble regresjonsanalyser med forskjellige hypotetiske forutsetninger utført. Analysene viser til noen signifikante verdier, men både med kvalitet og effektivitet som forklaringsvariabler er forklaringsgraden lav.

Studiens resultater påvirkes av metoden. Valg av variabler, modeller og analyseverktøy kan endre resultatene og i verste fall føre til feiltolkning. Kvalitetsresultatene påvirkes sterkt av variabelvalget, og har betydelig høyere standardavvik enn effektivitetsmodellene. Dette skyldes hovedsakelig at alle kvalitetsvariablene er relative verdier med lav korrelasjon til hverandre. Samtidig er outputen satt lik en, noe som gjør at variabelt skalautbytte ikke kan beregnes.

Få helseforetak beregnes til kostnadseffektive. Dette fordi modellen kun inneholder en input og en output. Dette er også grunnen til at kostnadseffektivitetsfronten endrer seg likt for alle helseforetak gitt konstant skalautbytte.

Modell for teknisk effektivitet korrelerer kun moderat med kostnadseffektivitetsmodellen. En del av dette kan skyldes at inputene til teknisk effektivitetsmodell i praksis brukes til mer enn kun produksjon av DRG-poeng, for eksempel forskning og utdanning. Samtidig mangler teknisk effektivitetsmodell en variabel for medisinsk utstyr og vareforbruk, noe som inkluderes i kostnadsmodellen.

Både dataomhyllingsanalysene og Malmquist-produktivitetsindeksene er ikke-parametriske modeller. Dette gjør dem mer sensitive mot små endringer i datasettet. Dette er spesielt viktig for prosesskvalitetsmodellen ettersom kvaliteten har betydelig mer spredte resultater enn effektivitetsmodellene. Dataomhyllingsanalysene forutsetter substituerbare inputer, men den sterke korrelasjonen mellom inputer i teknisk modell indikerer ikke-substituerbare inputs. Vi velger ikke Barnum et al. (2009) sin løsning ettersom dette vil resultere i mange effektivitetsscorer per helseforetak, noe som er ugunstig for oppgaven. Samtidig er dataomhyllingsanalysen bredt brukt i effektivitetsanalyser av helsesektoren. Bruken av dataomhyllingsanalyser medfører at helseforetak som er generelt gode på alle inputer ansees som effektive mens en modell for ikke-substituerbare inputer fortsatt anser dem ineffektive. Dette ender likevel ikke på rangeringen av de effektive helseforetakene, som er det viktigste.

For Malmquist-produktivitetsindekser kan skalaendringer være et problem ettersom indeksen kan feilberegnes ved store skalaendringer. For studien er ikke dette et problem ettersom helseforetakene ikke har hatt store skalaendringer i løpet av perioden.

Både dataomhyllingsanalyser og Malmquist-produktivitetsindekser er gode verktøy som gir brukbare resultater til tross for svakhetene. Priser og vekter må ikke fastsettes på forhånd ved beregning av scorene og indeksene. Verktøyene krever heller ingen funksjon for optimale forhold mellom input og output som observasjonene relateres til.

Resultatene fra regresjonsanalysene må tolkes med forsiktighet. Dette fordi regresjonsmodellen har en rekke teoretiske forutsetninger som er vanskelig å innfri i virkeligheten. Dette blir enda vanskeligere forutsatt et relativt lite datasett. Vi risikerer derfor å beregne urealistiske, signifikante verdier, eller ikke å finne signifikante verdier der de eksisterer.

## 7.2. Tolkning av resultater

Med metodens styrker og svakheter tatt i betraktning mener vi at det finnes en svak positiv relasjon mellom kvaliteten og effektiviteten. Undersøkelsene henter til at helseforetak med høy prosesskvalitet også har høy effektivitet, og at helseforetak med høy effektivitet også har høy kvalitet.

Eksempelet til statssekretær Kåss om Sykehuset i Vestfold samsvarer godt med våre resultater [6]. Fra og med 2009 er helseforetaket beregnet til å være kostnadseffektiv, både av vår analyse og SAMDATA-analysen Kåss henviser til. Videre påstås det at kvaliteten økte etter første budsjettoverskudd i 2010, noe vår kvalitetsberegninger av helseforetaket (DMU 6) også viser. En lignende trend finner vi hos Sykehuset i Telemark HF (DMU 7) og Sørlandets Sykehus HF (DMU 8), som gjennom hele analyseperioden har veldig høy kvalitet og ligger nesten hele tiden på kvalitets- og effektivitetsfronten. Trenden støttes flere helseforetak med både lav kvalitet og effektivitet.

Dette gjelder ikke i alle tilfeller, og helseforetak kan ha høy kvalitet uten god effektivitet eller omvendt. En forbedring i ressursutnyttelsen – i form av døgnplasser og årsverk – til helseforetak ser også ut til å forverre endringene i prosesskvaliteten. Dette kan hinte til at for høy effektivitet slår negativt ut på kvaliteten, slik at det kan lønne seg å stoppe effektiviseringen på et visst nivå for å fokusere mer på kvalitet. En oppskrift Sykehuset i Vestfold virker å ha hatt god suksess med. Funnet til endring over tid støtter delvis helsepersonellens påstander om at økt effektivitet gir dårlig kvalitet.

Ifølge våre statistiske tester kan helseforetakenes endringer i kostnadseffektivitet verken knyttes opp mot ressursutnyttelsen eller prosesskvaliteten. Dette er interessant ettersom man kan forvente at kostnadseffektiviteten i behandling kan relateres sterkt til ressursutnyttelsen i behandling. I perioden har helseforetak generelt redusert sin kostnadseffektivitet til tross for både forbedring og forverring i ressursutnyttelsen. Den manglende relasjonen kan potensielt forklares ved at kostnader relatert til medisinsk utstyr utgjør en betydelig del av kostnadsbildet. Hvis ressursutnyttelsen forbedres samtidig som kostnadseffektiviteten går ned, kan dette være et tegn på betydelig økning i behandlingsrelaterte varekostnader og medisinsk utstyr. Dette kan også tyde på at effektiv bruk av varekostnader og medisinsk utstyr ikke påvirker prosesskvaliteten direkte. Dette fordi endring i teknisk effektivitet har sterkere relasjon til endring i prosesskvaliteten enn endring i kostnadseffektivitet.

De funn som er gjort er både vage og til dels motstridende. Effektive helseforetak har god sjans for å ha god kvalitet, men endringer i teknisk effektivitet påvirker prosesskvaliteten negativt. Dette tyder på at kvaliteten og effektiviteten i helseforetak ikke påvirker hverandre direkte. En mulighet er flere forskjellige interne og eksterne forhold som påvirker den endelige kvaliteten og effektiviteten. Styrken og forholdet til en potensiell samvariasjon mellom kvaliteten og effektiviteten vil derfor være påvirket av miksen av interne og eksterne forhold som påvirker dem. Et forhold kan, for eksempel; kun påvirke prosesskvaliteten, kun påvirke effektiviteten, påvirke begge positivt eller negativt, eller påvirke kvaliteten og effektiviteten omvendt av hverandre. Den svært vage negative relasjonen mellom endring i prosesskvalitet og teknisk effektivitet kan skyldes en trend hvor helseforetak fokuserer på enten ressursutnyttelse eller kvaliteten. Et slikt fokus kan være latent i støyen av andre forhold som påvirker kvaliteten og effektiviteten innen spesialisthelsetjenesten.

Resultatkvalitetstiltakene ved Akershus universitetssykehus (DMU 2) viser at i perioden 08-09 klarte de å redusere antallet sykehusinfeksjoner betydelig, noe som skulle frigjøre ressursene i form av liggedøgn og personell [7]. I våre analyser har helseforetaket økt sin prosesskvalitet – en økning som kommer i form av færre korridorpasienter. Likevel er helseforetaket blitt mindre effektiv i perioden 2008-2009. Dette kan enten være fordi tiltakene ikke gav de effektivitetsgevinstene helseforetaket så for seg, eller fordi andre forhold påvirket effektiviteten betydelig mer i negativ retning. Hvis dette er typisk for resultatene av tiltak, kan det være årsak til det svake forholdet mellom prosesskvaliteten og effektiviteten. Disse sammenhengene blir likevel bare spekulasjoner.

### 7.3. Sammenlikning mot tidligere litteratur

De svake samvariasjonene fra våre undersøkelser har både likheter og ulikheter med tidligere studier. Ulikheter i resultater betyr ikke nødvendigvis at et av studiene er gale ettersom forutsetninger, begreper og metoder er forskjellige mellom forskjellige studier.

Forskjellene er store mellom våre resultater og resultatene fra Morey et al. (1992). I deres studie var det en klar relasjon mellom penger brukt og kvaliteten, hvor en økning i kostnader førte til høyere kvalitet. Denne relasjonen er ikke like klar i vår studie. Det er to grunner til disse forskjellene. For det første er kvalitetsbegrepet forskjellig. Mens vår studie fokuserer på prosesskvaliteten, har Morey et al. (1992) fokusert på en justert overlevelsesrate som kvalitetsfaktor. Som tidligere redegjort er det et klart skille mellom resultatorientert kvalitet

og prosesskvalitet. For det andre har Morey et al. (1992) strenge grenser for sitt studie, og så mange eksterne faktorer som mulig er kartlagt for å kun måle forholdet mellom kostnader og kvalitet. I vår studie er både eksterne og interne forhold i stand til å påvirke kostnadene og prosesskvaliteten.

Studien av Nayar & Ozcan (2007) fant en positiv korrelasjon mellom kvaliteten og den tekniske effektiviteten til sykehus. Dette er heller ikke i samsvar med våre resultater ettersom vi ikke fant en klar korrelasjon mellom kvaliteten og teknisk- eller kostnadseffektivitet. Denne forskjellen kan skyldes en differanse i metoden til våre studier. I vår studie har vi brukt tre modeller; modell for kostnadseffektivitet, modell for teknisk effektivitet og modell for prosesskvaliteten. I studiet til Nayar & Ozcan (2007) ble det brukt en modell for effektivitet og en modell for effektivitet og kvalitet. Dette medfører at effektivitetsmålene ikke settes opp mot kvalitetsmålene alene og relasjonen vil påvirkes ved at effektivitetsmålene delvis settes opp mot seg selv, noe som resulterer i mye høyere korrelasjon mellom modellene. Samtidig viser Nayar & Ozcan (2007) til at sykehus med lav teknisk effektivitet også kan ha høy kvalitet, noe som avviker fra deres resultater.

Studien av Valdmanis et al. (2008) fant at sykehus med høy kvalitet hadde gjennomsnittlig høyere effektivitet enn sykehus med lav kvalitet. Dette er konsistent med våre funn hvor helseforetak med høy kvalitet har gjennomsnittlig høyere effektivitet. Samtidig kunne høy kvalitet relateres til høye totalkostnader. Dette resultatet er skjevt i forhold til våre resultater. En mulig grunn kan være at Valdmanis et al. (2008) ser på totalkostnadene i stedet for kun kostnader beregnet opp mot pasientbehandling. Et annet interessant moment er at Valdmanis et al. (2008) mener kvaliteten påvirker effektiviteten. I vår studie strekker vi ingen klare årsakssammenhenger mellom effektiviteten og kvaliteten ettersom det er vanskelig å se hva som skjer først. Konklusjonen sier også at ressursreallokering kan øke kvaliteten uten å medføre økte kostnader. Dette samsvarer med vår diskusjon om endringer i interne forhold som kan forbedre, forverre eller vedlikeholde kvaliteten og effektiviteten i helseforetakene.

#### 7.4. Konklusjon

Med hensyn til oppgavens problemstilling kan vi ikke konkludere med at det eksisterer en direkte relasjon mellom prosesskvaliteten og effektiviteten i spesialisthelsetjenesten. Å avvise at en direkte relasjon eksisterer er også urimelig ettersom de vage resultatene kan være påvirket av metodens svakheter og mangfoldet av faktorer som påvirker kvalitet og

effektivitet. Studiens indikasjon likevel viser til at helseforetak med høy effektivitet kan ha høy prosesskvalitet, men at forbedring i teknisk effektivitet kan ha negativ effekt på endring i prosesskvalitet. I alle tilfeller er trendene svake. Effektivitet har liten påvirkningskraft på kvalitet, og kvalitet liten påvirkningskraft på effektivitet. Siden en direkte relasjon ikke kan påvises, konkluderer vi med at samvariasjonen sannsynligvis kommer fra andre forhold som påvirker kvalitet og effektivitet i samme retning.

Det negative synet på «New Public Management» og effektivisering generelt kan ikke generaliseres ettersom kvaliteten og effektiviteten ikke påvirker hverandre negativt i betydelig grad. Eventuelle protester mot effektiviseringen kan derimot rettes mot enkelttiltak som skader kvaliteten ettersom det ikke er noen garanti for at enkelttiltak ikke påvirker kvaliteten negativt. Tvert imot ser det ut til at flere helseforetak har hatt høy forbedring innen kvalitet til tross for effektiviseringen. Det er derfor feil å si at helseforetak får dårligere kvalitet på grunn av effektiviseringen i seg selv. Eksempelvis kan denne studiens resultater brukes til å sammenlikne helseforetaks prosesskvalitet og effektivitet opp mot deres interne og eksterne forhold, og slik se hvilke tiltak som fungerer best. Eventuelt kan hvert enkelt helseforetak studere egen utvikling opp mot andres utvikling og de tiltakene som er utført.

#### 7.4.1. Veien videre

Studien har resultert i at helseforetak har en svak korrelasjon mellom sin kvalitet og effektivitet, men uten at relasjoner er blitt funnet mellom endringer i kvalitet og effektivitet. En mulighet til de vage resultatene er at studien har en stor bredde. Studien kan derfor tas videre i flere retninger. For det første kan studien tas ned fra helseforetaksnivå til sykehusnivå eller avdelingsnivå for å studere forholdet mellom kvalitet og effektivitet. Dette ville redusere antallet forhold som påvirker hver observasjon betydelig, og samtidig skape flere observasjoner per år.

For det andre er kvalitetsbegrepet til denne oppgaven begrenset til kun prosesskvalitet. Dette gir ikke et fullstendig bilde av kvaliteten. Når nasjonale kvalitetsindikatorer er mer fullstendige kan struktur- og resultatindikatorer tas med. For det tredje kan effektive og ineffektive helseforetak settes opp mot hverandre. På denne måten kan interne og eksterne forhold avdekkes og kartlegges for å finne hvilke effekter som påvirker kvalitet og effektiviteten positivt og negativt.



## 8. Referanser

### Artikler og bøker

- American Medical Association, Council of Medical Service. (1986). Quality of care. *The Journal of the American Medical Association* 256, 1032-1034.
- Andersen, P., & Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management science*, 39(10), 1261-1264.
- Andrews, & Pregibon. (1978). Finding the Outliers that Matter. *Journal of the Royal Statistical Society, Series B* 40(1), 85-93.
- Banker, R. D. (1993). Maximym Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation. *Management Science*, 39(10), 1265-1273.
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management science*, 30(9), 1078-1092.
- Barnum, Walton, Shields, & Schumock. (2009). Measuring Hospital Efficiency with Data Envelopment Analysis: Nonsubstitutable vs. Substitutable Inputs and Outputs. *Journal of Medical Systems*, 35(6), 1393-1401.
- Biørn, E., Hagen, T. P., Iversen, T., & Magnussen, J. (2009). How different are hospital's responses to a financial reform? The impact on efficiency of activity-based financing. *Health Care Manag Sci.*
- Blumberg, B., Cooper, D., & Schindler, S. (2011). *Business Research Methods (3rd european version)*. McGraw-Hill.
- Blumenthal, D. (1996). Quality of Care - what is it? *The New England Journal of Medicine*, 335, 891-894.
- Bogetoft, P., & Otto, L. (2011). *Benchmarking with DEA, SFA and R*. Springer.
- Brook, R. H., & Cleary, P. D. (1996, September 26). Quality of Care Part 1. *The New England Journal of Medicine*, 12, 891-970.
- Campbell, S., Roland, M., & Buetow, S. (2000). Defining quality of care. *Social Science & Medicine*, 51, 1611-1625.
- Caves, D., Christensen, L., W., & Diewert, E. (1982). The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity. *Econometric*, 50(6), 1393-1414.
- Charnes, Cooper, & Rhodes. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429-444.
- Chassin, M., Galvin, R., & Quality, t. N. (1998). The Urgent Need to Improve Health Care Quality. *The Journal of the American Medical Association*, 280(11), 1000-1006.

- Coelli, Rao, O'Donnell, & Battese. (2005). *An Introduction To Efficiency And Productivity Analysis (6th edition)*. Springer.
- Cooper, Seiford, & Zhu. (2011). *Handbook on Data Envelopment Analysis (2nd edition)*. Springer.
- Donabedian, A. (1988). The Quality of Care. *The Journal of the American Medical Association*, 260(12), 1743-1748.
- Dyson, Allen, Camanho, Podinovski, Sarrico, & Shale. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, 245-259.
- Edvardsen, D. (2004). *Four Essays on the Measurement of Productive Efficiency*. University of Gothenburg.
- Farrell, M. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3), 253-290.
- Färe, Grosskopf, Norris, & Zhang. (1994). Productivity Growth, Technical Progress and Efficiency Change in Industrialized Countries. *The American Economic Review*, 84, 66-83.
- Grifell-Tatjé, & Lovell. (1995). A note on the Malmquist productivity index. *Economics Letters*, 47, 169-175.
- Groebner, Shannon, Fry, & Smith. (2008). *Business Statistics. A Decision-Making Approach (7th edition)*. Pearson.
- Hagen, & Kjekshus. (2003). Ga sammenslåinger av sykehus bedre effektivitet? *Health Organization Research Norway 2003:2*.
- Hagen, T. P., Iversen, T., & Magnussen, J. (2001). *Sykehusenes effektivitetsutvikling 1992-1999: Hvilke effekter ga innsatsstyrt finansiering?* Oslo: Health Economics Research Programme.
- Hill, Griffiths, & Lim. (2012). *Principles of Econometrics (4th edition)*. Wiley.
- Hood, C. (1995). THE «NEW PUBLIC MANAGEMENT IN THE 1980's: VARIATIONS ON A THEME. *Accounting, Organizations and Society*, 20, 93-109.
- Kittelsen, S. A., & Førsum, F. R. (2001). Empirisk forskningsresultater om effektivitet i offentlig tjenesteproduksjon\*. *Økonomisk forum Nr.6*, 22-29.
- Kittelsen, S. A., Magnussen, J., & Anthun, K. S. (2007). *Sykehusproduktiviteten etter statlig overtakelse: En nordisk komparativ analyse*. Oslo: Health Economics Research Programme.
- Kjekshus, L. (2000). Om å sammenlikne sykehusprestasjoner. *Tidsskrift for Den norske legeforening*.
- Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318), 399-402.

- Lind, Marchal, & Wathon. (2010). *Statistical Techniques in Business & Economics (5th edition)*. McGraw-Hill/Irwin.
- Lohr, K. N. (1990). *Medicare: a strategy for quality assurance*. National Academies Press.
- Magnussen, J. (1996). Efficiency Measurement and the Operationalization of Hospital Production.
- McDonald, J. (2009). Using Least Squares and Tobit in Second Stage DEA Efficiency Analyses. *European Journal of Operational Research*, 729-798.
- Morey, R. C., Fine, D. J., Loree, S. W., Retzlaff-Roberts, D. L., & Tsubakitani, S. (1992). The trade-off between hospital cost and quality of care: An exploratory empirical analysis. *Medical Care*, 677-698.
- Nayar, P., & Ozcan, Y. A. (2008). Data Envelopment Analysis Comparison of Hospital Efficiency and Quality. *Journal of Medical Systems*, 32, 193-199.
- Pedersen, M. (2007). *SAMDATA Sektorrapport for somatisk spesialisthelsetjeneste 2006 3/07*. SINTEF.
- Ray, S. C., & Desli, E. (1997). Productivity growth, technical progress, and efficiency change in industrialized countries: comment. *The American Economic Review*, 87(5), 1033-1039.
- Valdmanis, V., Rosko, M. D., & Mutter, R. L. (2008). Hospital Quality, Efficiency, and Input Slack Differential. *Health Services Research*, 43(5), 1830-1848.
- Wilson, & Carey. (2004). Nonparametric Analysis of Returns to Scale in the US Hospital Industry. *Journal of Applied Econometrics*, 19, 505-524.
- Wilson, P. (1993). Detecting outliers in deterministic nonparametric frontier models with multiple outputs. *Journal of Business & Economic Statistics*, 11(3), 319-323.
- Zhu, J. (1996). Robustness of the efficient DMUs in data envelopment analysis. *European Journal of Operational Research*, 90(3), 451-460.

### **Offentlige dokumenter**

- NOU 2012:6 *Arbeidsrettede tiltak*
- NOU 2004:2 *Effekter og effektivitet. Effekter av statlig innsats for regional utvikling og distriktpolitiske mål.*
- Ot.prp. nr. 66(2000-2001) *Om lov om helseforetak m.m. (helseforetaksloven)*
- St.meld. nr. 44 (1995-96) *Ventetidsgarantien - kriterier og finansiering*
- St.prp. nr. 1 (2006-2007). *Nasjonal helseplan 2007-2010.*
- Sosial- og helsedirektoratet. (2005). *Nasjonal strategi for kvalitetsforbedring i Sosial- og helsetjenesten*. IS-1162.

## Internettkilder

- [1] Helseregnskap 2012. (13.03.13). Hentet fra SSB: <http://www.ssb.no/nasjonalregnskap-og-konjunkturer/statistikker/helsesat> (27.05.13)
- [2] Sykehusmatematikk. (10.02.13). Hentet fra Nordlys: <http://www.nordlys.no/debatt/kronikk/article6492885.ece> (27.05.13)
- [3] Styrk fag og kvalitet i sykehusene! (28.05.2009) Hentet fra Tidsskrift for Den norske legeforening: <http://tidsskriftet.no/article/1841440> (27.05.13)
- [4] Kritisk til helsereformen. (10.08.2011). Hentet fra: <http://www.sykepleien.no/nyhet/692483/kritiske-til-helsereformen> (27.05.13)
- [5] Effektivitet gir kvalitet. Hentet fra Dagens Medisin: <http://www.dagensmedisin.no/nyheter/effektivitet-gir-kvalitet> (27.05.13)
- [6] Effektivitet og kvalitet i helsesektoren. (18.10.11) Hentet fra SAS Institute: [http://www.sas.com/offices/europe/norway/pdf/Robin\\_Martin\\_Kaass\\_Effektivitet\\_kvalitet\\_helsesektoren\\_Helseforum.pdf](http://www.sas.com/offices/europe/norway/pdf/Robin_Martin_Kaass_Effektivitet_kvalitet_helsesektoren_Helseforum.pdf) (27.05.13)
- [7] Forbedringstiltak som har ført til bedre kvalitet og effektivitet ved Akershus Universitetssykehus (18.10.11) Hentet fra SAS Institute: [http://www.sas.com/offices/europe/norway/pdf/Helge\\_Svaar\\_Forbedringstiltak\\_Helseforum.pdf](http://www.sas.com/offices/europe/norway/pdf/Helge_Svaar_Forbedringstiltak_Helseforum.pdf) (27.05.13)
- [8] Helseforetak. Hentet fra Store norske leksikon: <http://snl.no/helseforetak> (27.05.13)
- [9] Samdata. (06.12.2011) Hentet fra Helsedirektoratet: <http://www.helsedirektoratet.no/tall-analyse/samdata> (27.05.13)
- [10] DRG-systemet. Hentet fra Helsedirektoratet: <http://www.helsedirektoratet.no/finansiering/drg/Sider/default.aspx> (27.05.13)
- [11] Om Samdata. (12.12.2011). Hentet fra Helsedirektoratet: <http://www.helsedirektoratet.no/tall-analyse/samdata/om-samdata/Sider/default.aspx> (27.05.13)
- [12] Nasjonale kvalitetsindikatorer. Hentet fra Helsedirektoratet: <http://nesstar2.shdir.no/kvalind/> (27.05.13)
- [13] Datagrunnlag. (12.12.11). Hentet fra Helsedirektoratet: <http://www.helsedirektoratet.no/tall-analyse/samdata/datagrunnlag/Sider/default.aspx> (28.05.13)
- [14] Spesialisthelsetjenesten. Hentet fra Statistisk sentralbyrå: <https://www.ssb.no/statistikkbanken/selecttable/hovedtabellHjem.asp?KortNavnWeb=speshelse&CMSSubjectArea=helse&checked=true> (28.05.13)
- [15] Om kvalitet i helsetjenesten. Hentet fra: <http://helsenorge.no/Helsetjenester/Sider/Om-kvalitet-i-helsetjenesten.aspx> (28.05.13)
- [16] Sykehuskøer – en gordisk knute. (24.04.13) Hentet fra Dagens Medisin: <http://www.dagensmedisin.no/debatt/sykehuskoer---en-gordisk-knute-siden-1985/> (28.05.13)

## Vedlegg

### Vedlegg 1: Liste over helseforetak

<b>DMU</b>	<b>Helseforetak</b>
DMU1	Sykehuset i Østfold HF
DMU2	Akershus Universitetssykehus HF
DMU3	Oslo Universitetssykehus HF
DMU4	Sykehuset Innlandet HF
DMU5	Vestre Viken HF
DMU6	Sykehuset i Vestfold HF
DMU7	Sykehuset i Telemark HF
DMU8	Sørlandets Sykehus HF
DMU9	Helse Stavanger HF
DMU10	Helse Fonna HF
DMU11	Helse Bergen HF
DMU12	Helse Førde HF
DMU13	Helse Møre og Romsdal HF
DMU14	St. Olavs Hospital HF
DMU15	Helse Nord-Trøndelag HF
DMU16	Helgelandsykehuset HF
DMU17	Nordlandsykehuset HF
DMU18	Universitetssykehuset i Nord-Norge HF
DMU19	Helse Finnmark HF

Vedlegg 2: Liste over tilgjengelige kvalitetsindikatorer fra Helsedirektoratet

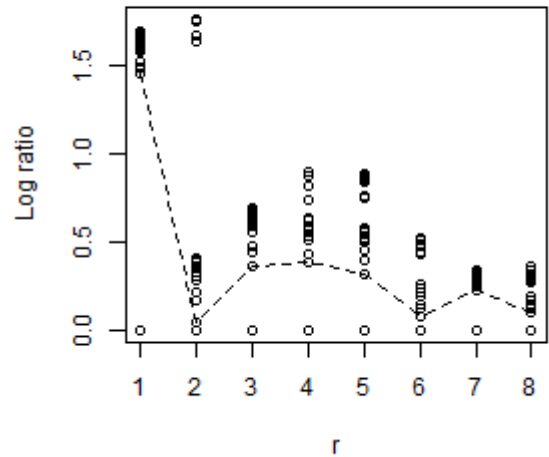
Indikator	Utvalgelse		
	Ufullstendig	Uegnet	OK
Korridorpasienter			X
Epikriser sendt innen 7 dager			X
Sykehusinfeksjoner	X		
Lårhalsbruddoperasjoner innen 48 timer			X
Strykninger fra planlagt operasjonsprogram			X
Individuell plan barnehabilitering	X		
Keisersnittoperasjoner		X	
Trombolysebehandlinger	X		
Perinealruptur (fødselsrifter) 3. og 4. grad	X		
Underekstremitetsamputasjoner blant pasienter med diabetes		X	
Tid fra henvisning til første behandling for tykktarmskreft	X		
Tid fra henvisning til første behandling for lungekreft	X		
Tid fra henvisning til første behandling for brystkreft	X		
Andel oppdaterte ventetider for Fritt Sykehusvalg - somatikk		X	
5 års overlevelsesrate tykktarmskreft pr. helseregion	X		
5 års overlevelsesrate endetarmskreft pr. helseregion	X		
5 års overlevelse lungekreft pr. helseregion	X		
5 års overlevelsesrate brystkreft pr. helseregion	X		
5 års overlevelse prostatakreft pr. helseregion	X		
30-dagers overlevelse etter sykehusinnleggelse ved hoftebrudd	X		
30-dagers overlevelse etter sykehusinnleggelse ved hjerteinfarkt	X		
30-dagers overlevelse etter sykehusinnleggelse ved hjerneslag	X		
30-dagers overlevelse etter sykehusinnleggelse uansett årsak	X		
Ventelistedata		X	

### Vedlegg 3: Outlier-analyse

#### Kostnadseffektivitet

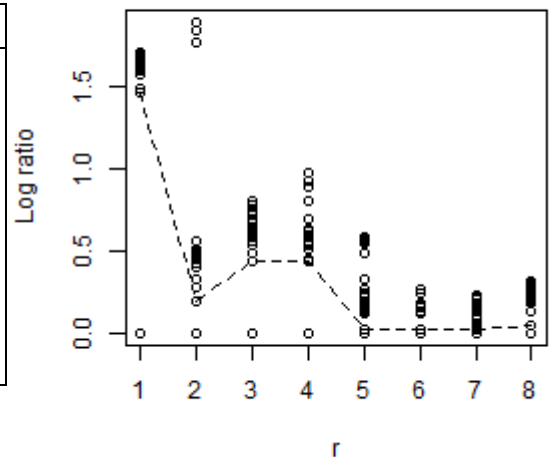
2008(09)

<i>i</i>									$R^i$
1	3								0,16
2	14	3							0,09
3	14	18	3						0,04
4	14	11	18	3					0,02
5	14	11	18	4	3				0,01
6	14	11	18	4	5	3			0,01
7	14	8	11	18	4	5	3		0,01
8	7	14	8	11	18	4	5	3	0,00



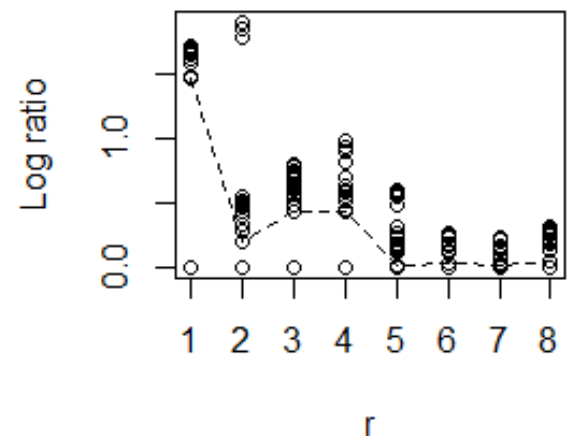
2009(09)

<i>i</i>									$R^i$
1	3								0,17
2	18	3							0,10
3	14	18	3						0,05
4	14	11	18	3					0,02
5	2	14	11	18	3				0,01
6	2	14	11	18	4	3			0,01
7	2	14	11	5	18	4	3		0,00
8	2	14	8	11	5	18	4	3	0,00



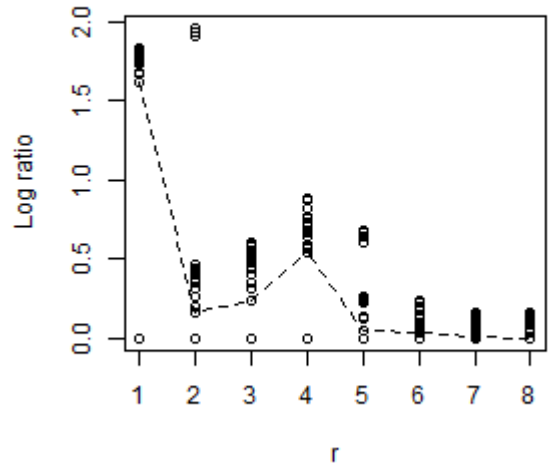
2010(10)

<i>i</i>									$R^i$
1	3								0,17
2	18	3							0,10
3	14	18	3						0,05
4	14	11	18	3					0,02
5	2	14	11	18	3				0,01
6	2	14	11	18	4	3			0,01
7	2	14	11	5	18	4	3		0,00
8	2	14	8	11	5	18	4	3	0,00



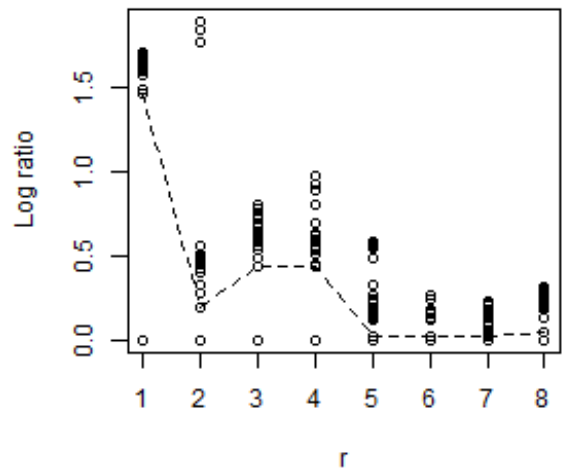
2010(11)

<i>i</i>									$R^i$
1	3								0,15
2	18	3							0,09
3	14	18	3						0,05
4	14	11	18	3					0,02
5	14	11	18	4	3				0,01
6	14	11	5	18	4	3			0,01
7	14	6	11	5	18	4	3		0,01
8	14	7	6	17	11	18	4	3	0,00



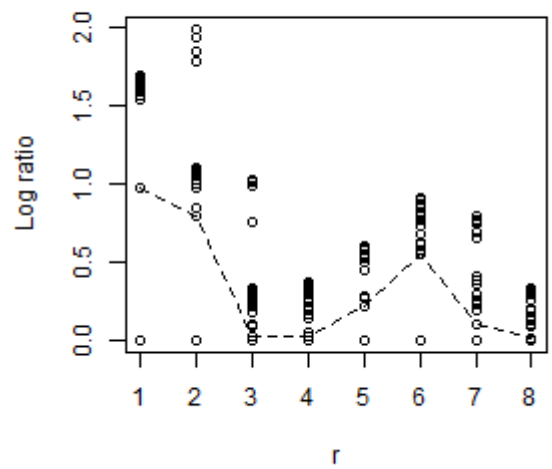
2011(10)

<i>i</i>									$R^i$
1	3								0,15
2	18	3							0,09
3	14	18	3						0,05
4	14	11	18	3					0,02
5	14	11	18	4	3				0,01
6	14	11	5	18	4	3			0,01
7	14	6	11	5	18	4	3		0,01
8	14	7	6	17	11	18	4	3	0,00



2011(11)

<i>i</i>									$R^i$
1	3								0,18
2	5	3							0,05
3	11	5	3						0,03
4	14	11	5	3					0,02
5	2	14	18	5	3				0,01
6	2	14	18	11	5	3			0,00
7	2	14	18	4	11	5	3		0,00
8	2	14	18	17	4	11	5	3	0,00

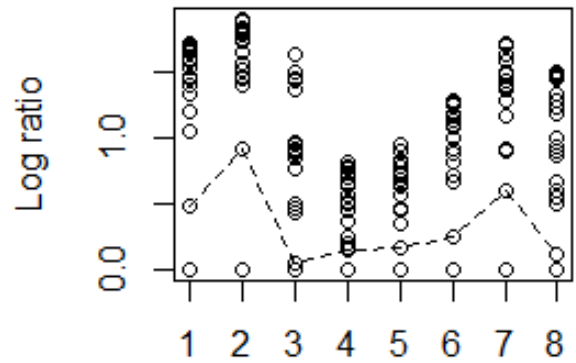




## Teknisk effektivitet

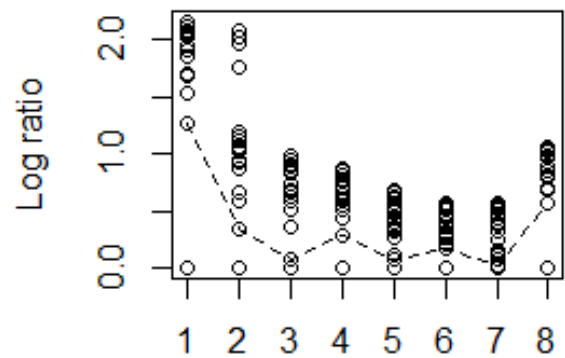
2008(09)

<i>i</i>								$R^i$	
1	3							0,16	
2	11	3						0,02	
3	18	11	3					0,01	
4	2	18	11	3				0,00	
5	2	4	18	11	3			0,00	
6	2	8	4	18	11	3		0,00	
7	2	8	14	4	18	11	3	0,00	
8	2	8	14	5	4	18	11	3	0,00



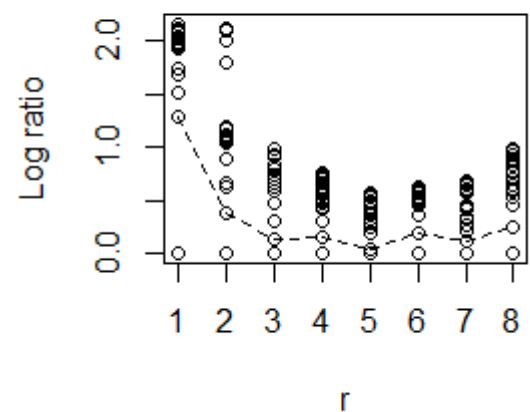
2009(09)

<i>i</i>								$R^i$	
1	3							0,10	
2	11	3						0,03	
3	11	14	3					0,01	
4	18	11	14	3				0,00	
5	4	18	11	14	3			0,00	
6	2	4	18	11	14	3		0,00	
7	7	6	4	18	11	14	3	0,00	
8	7	13	6	4	18	11	14	3	0,00



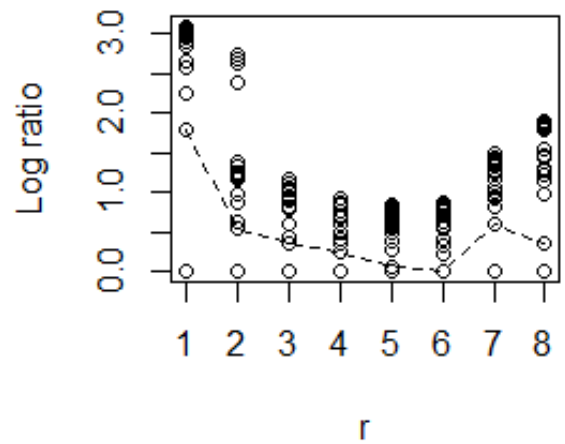
2010(09)

<i>i</i>								$R^i$	
1	3							0,10	
2	11	3						0,03	
3	11	14	3					0,01	
4	18	11	14	3				0,00	
5	18	4	11	14	3			0,00	
6	2	18	4	11	14	3		0,00	
7	7	1	18	4	11	14	3	0,00	
8	7	13	6	18	4	11	14	3	0,00



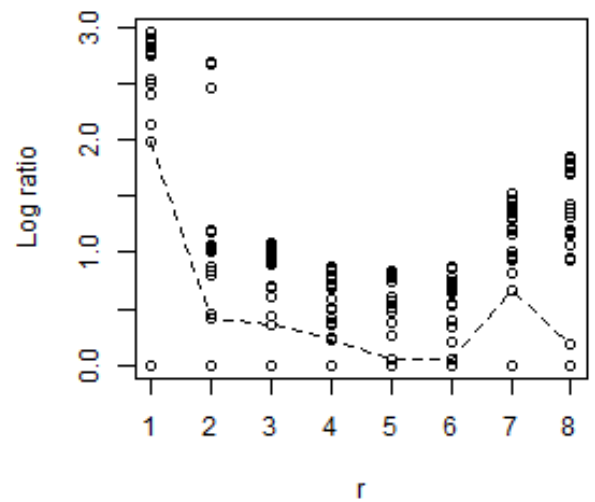
2010(10)

<i>i</i>									$R^i$
1	3								0,04
2	11	3							0,01
3	18	11	3						0,00
4	18	14	11	3					0,00
5	2	18	4	11	3				0,00
6	2	18	4	14	11	3			0,00
7	7	1	18	4	14	11	3		0,00
8	7	2	1	18	4	14	11	3	0,00



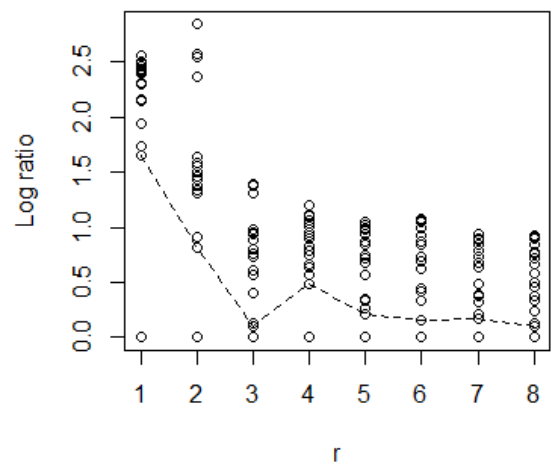
2010(11)

<i>i</i>									$R^i$
1	3								0,05
2	11	3							0,01
3	18	11	3						0,00
4	18	14	11	3					0,00
5	2	4	18	11	3				0,00
6	7	1	18	14	11	3			0,00
7	7	1	4	18	14	11	3		0,00
8	7	2	1	4	18	14	11	3	0,00



2011(11)

<i>i</i>									$R^i$
1	3								0,07
2	11	3							0,01
3	14	11	3						0,00
4	4	14	11	3					0,00
5	7	4	14	11	3				0,00
6	7	1	4	14	11	3			0,00
7	7	1	4	5	14	11	3		0,00
8	7	1	4	5	6	14	11	3	0,00



#### Vedlegg 4: R-script

```
##### DEA #####
library(Benchmarking)

Input <- "prod0809.csv"
Output <- "prod0809_dea.csv"

data <- read.csv2(Input,header=TRUE,sep=";",dec=",")

# definerer variabler
Y <- data[,c(3)]
X <- data[,c(4,5,6,7)]

# DEA
CRS <- 1/eff(sdea(X,Y,RTS="crs",ORIENTATION="out"))
VRS <- 1/eff(sdea(X,Y,RTS="vrs",ORIENTATION="out"))

RES<-cbind(CRS,VRS)
write.csv2(RES,file=Output)
```

```
##### MPI #####
library(FEAR)

Input1 <- "prod0809.csv"
Input2 <- "prod0909.csv"
Output <- "prod08-09_mpi.csv"

data1 <- read.csv2(Input1,header=TRUE,sep=";",dec=",")
data2 <- read.csv2(Input2,header=TRUE,sep=";",dec=",")

x1 <- data1[,c(4,5,6,7)]
x2 <- data2[,c(4,5,6,7)]

y1 <- data1[,c(3)]
y2 <- data2[,c(3)]

id1 <- data1[,c(1)]
id2 <- data2[,c(1)]

m <- malmquist.components(t(x1),t(y1),t(id1),t(x2),t(y2),t(id2),
ORIENTATION=2,NREP=0)

mpi <- malmquist(LIST=m)
mid <- mpi$id
malm <- mpi$malm
eff <- mpi$eff
tech <- mpi$tech
pure.eff <- mpi$pure.eff
pure.tech <- mpi$pure.tech
sch <- mpi$sch

res <- (cbind(mid,malm,eff,tech,pure.eff,pure.tech,sch))
write.csv2(res,file=Output)
```

```

##### Banker-test #####
library(Benchmarking)

Input <- "prod0809.csv"
Output <- "prod0809_res.csv"

data <- read.csv2(Input,header=TRUE,sep=";",dec=",")

# Definerer output
Y <- data[,c(3)]

# Definerer input-varianter
X1 <- data[,c(4,5,6,7,8,9)]
X2 <- data[,c(4,11)]
X3 <- data[,c(4,5,12)]
X4 <- data[,c(4,5,9,13)]
x5 <- data[,c(4,5,14,8,9)]

# Beregner Teknisk Effektivitet
F1 <- eff(dea(X1,Y,RTS="crs",ORIENTATION="in"))
F2 <- eff(dea(X2,Y,RTS="crs",ORIENTATION="in"))
F3 <- eff(dea(X3,Y,RTS="crs",ORIENTATION="in"))
F4 <- eff(dea(X4,Y,RTS="crs",ORIENTATION="in"))
F5 <- eff(dea(X5,Y,RTS="crs",ORIENTATION="in"))

# TEX-test
TEX12 <- (sum(F2-1)/length(F2))/(sum(F1-1)/length(F1))
TEX13 <- (sum(F3-1)/length(F3))/(sum(F1-1)/length(F1))
TEX14 <- (sum(F4-1)/length(F4))/(sum(F1-1)/length(F1))
TEX15 <- (sum(F5-1)/length(F5))/(sum(F1-1)/length(F1))

# THN-test
THN12 <- (sum((F2-1)^2)/length(F2))/(sum((F1-1)^2)/length(F1))
THN13 <- (sum((F3-1)^2)/length(F3))/(sum((F1-1)^2)/length(F1))
THN14 <- (sum((F4-1)^2)/length(F4))/(sum((F1-1)^2)/length(F1))
THN15 <- (sum((F5-1)^2)/length(F5))/(sum((F1-1)^2)/length(F1))

# CBIND & SAVE
TEX <- cbind(TEX12,TEX13,TEX14,TEX15)
THN <- cbind(THN12,THN13,THN14,TEX15)

RES <- cbind(TEX,THN)
write.csv2(RES,file=Output)

```

```

##### Wilson-rutinen #####
library(FEAR)
library(Benchmarking)

Input <- "prod0809.csv"
Output1 <- "prod0809_wil.csv"
Output2 <- "prod0809_wil.png"

data <- read.csv2(Input,header=TRUE,sep=";",dec=",")

# variables
x <- with(data, rbind(dognplasser,leger,annet,behandling))
y <- with(data, rbind(DRG))

# calculation
tap <- ap(x,y,NDEL=8)
res <- print(cbind(tap$imat,tap$r0), na.print="", digit=2)
outlier.ap.plot(tap$ratio)

# output
png(file=Output2, width=300 ,height=300)
write.csv2(res,file=Output1)

```

```

#### Skalatest ####
library(Benchmarking)
data <- read.csv2(Input,header=TRUE,sep=";",dec=",")

Y <- data[,c(3)]
X <- data[,c(4,5,9,13)]

F1 <- eff(sdea(X,Y,RTS="crs",ORIENTATION="out"))
F2 <- eff(sdea(X,Y,RTS="vrs",ORIENTATION="out"))

K<-length(F1)

plot(sort(F1),(1:K)/K,type="s",ylim=c(0,1),ylab="sannsynlighet",
xlab="effektivitetsscore")
lines(sort(F2),(1:K)/K,type="s",lty="dashed")
legend("bottomright",c("F1","F2"),lty=c("solid","dashed"),bty="n")

```

```

##### Kolmogorov-Smirnov test #####
library(Benchmarking)

Input<-"eff_TE.csv"
Output<-"eff_TE_res.csv"

data <- read.csv2(Input,header=TRUE,sep=";",dec=",")

# leser effektivitetstall fra DEA- og MPI-outputfiler
F1 <- data[,c(1)]
F2 <- data[,c(2)]
F3 <- data[,c(3)]
F4 <- data[,c(4)]
F5 <- data[,c(5)]
F6 <- data[,c(6)]

QE<-ks.test(F1, F2, alternative="two.sided")
CE<-ks.test(F3, F4, alternative="two.sided")
TE<-ks.test(F5, F6, alternative="two.sided")

RES<-cbind(QE,CE,TE)
write.csv2(RES,file=Output)

```

```

##### efficiency stepladder #####
library(Benchmarking)

Input<-"prod0809.csv"
Output<-"prod0809_ESL.csv"

data <- read.csv2(Input,header=TRUE,sep=";",dec=",")

x <- with(data, cbind(dognplasser,leger,behandling,annet))
y <- with(data, cbind(DRG))

n <- which(data$DMU=="DMU1")[1]
el <- eladder(n, x, y, RTS="crs")
EFF1 <- cbind(el$eff)

```