



**Copenhagen
Business School**
HANDELSHØJSKOLEN

CAND.MERC.(MAT.)-STUDIET
KANDIDATAFHANDLING, FORÅR 2013

EN PRODUKTIVITETSANALYSE AF LANDBRUGET I EU

AN ANALYSIS OF AGRICULTURAL PRODUCTIVITY IN THE EUROPEAN UNION

Forfatter: Michael Vierø - _____

Afleveringsdato: 31. maj 2013

Vejleder: Mette Asmild, Institut for fødevare- og ressource økonomi (IFRO) - Københavns
Universitet

Antal anslag: 154.210 anslag inkl. mellemrum, 17 tabeller og 15 figurer

Afhandlingens omfang: 77,63 normalsider

Abstract

The main objective of this study is to measure and compare levels and trends in the agricultural efficiency and productivity among the countries in the European Union. To complete the analysis, I use data drawn from the Farm Accountancy Data Network (FADN), and my study covers the period 1995-2009.

In the thesis I present the theories behind *Data Envelopment Analysis* (DEA) and *the Malmquist index* as a way of measuring respectively efficiency and productivity. DEA evaluates each DMU (Decision Making Unit) by choosing a benchmark, and seeks the best proportional way of improvement. The outcome of a DEA is an efficiency score between 0 and 1, where 1 indicates that the DMU is efficient, and a value between 0 and 1 that a DMU is inefficient. Malmquists total factor productivity index use the non-parametric technique of data envelopment to fit distance functions. This approach is non-stochastic and assumes a constant returns to scale on the frontier technology. Malmquists productivity index is commonly used, because it can decompose the productivity growth into a Catching-up and a Technical change part.

In the assignment I use DEA for two purposes. First I apply a DEA to each year separately, in order to clarify how a country performs relative to the other countries of the same year. My analysis shows, that the Netherlands, Denmark and Malta are located on the efficient frontier every year, which means, that according to DEA there is no potential improvement for those countries. Generally the results of the changes in efficiency from year to year, gives an idea of how the Catch-up effect contributes to Malmquists TFP (total factor productivity) growth.

Second, I carry out a DEA for the entire dataset at once to compare the efficiency scores across the years. The results show that Denmark, the Netherlands and Malta are the countries with the highest average annual efficiency scores, which means that farms in Denmark, the Netherlands and Malta are those which on average perform best.

Furthermore I apply the Malmquist index approach, to measure and decompose productivity growth in the agricultural production for the "old nations" of the European Union¹. The reason for restricting the calculations to only comprising the old countries, is to examine whether there is a coincident between the expansions of the European Union² and the productivity of the old nations. According to the Malmquist analysis, Denmark is the country with the highest average annual productivity growth (2,22 %). Besides that the results of the Malmquist approach show, that the old EU countries over the time period 1995-2009 experienced an overall increase in productivity. The average annual productivity developments for the old countries are very stable in the interval [-3%, 3%] right up to, and including 2006. In 2007, the Malmquist index increases by 7.47%, which is followed by significant productivity declines in 2008 and 2009. Overall, the analysis in this thesis states that the presence of the new countries of the EU from a productivity/relative efficiency point of view, has no impact on that of the old countries.

In general, this paper proposes an updated view of agricultural productivity in the EU countries, where each country can compare their regions productivity developments with the productivity developments in other EU countries. The results of both the DEA and Malmquist approaches show that the Netherlands and Denmark are among the best performing countries in terms of efficiency and productivity.

¹ Countries that are members of the EU on January 1, 1995.

² The European Union contained 15 countries in 1995 (the old countries) and was in 2004 and 2007 expanded by respectively 10 and 2 countries.

Indholdsfortegnelse

1. INTRODUKTION	6
1.1. INDLEDNING	6
1.2. MOTIVATION	7
1.3. PROBLEMFORMULERING	7
1.4. AFGRÆNSNING	8
1.5. AFHANDLINGENS OPBYGNING	9
1.6. VIDENSKABSTEORETISKE OVERVEJELSER	10
2. TEORI	11
2.1. DEA MODELLEN	11
2.1.1. MATEMATISK UDLEDNING AF DEA	12
2.1.1.1. Mulighedsområde og den efficiente rand	13
2.1.1.2. Farrells efficiensindeks	17
2.1.2. LØSNINGSMETODE	19
2.1.2.1. CCR modellen	19
2.1.3. EKSEMPEL PÅ DEN OUTPUTORIENTEREDE DEA MODEL	22
2.2. MALMQUIST INDEKSET	27
2.2.1. LØSNINGSMETODE	28
2.2.2. EKSEMPEL PÅ DET OUTPUTORIENTEREDE MALMQUIST INDEKS	33
3. DATABESKRIVELSE	40
3.1. DATASÆTTET	40
3.1.1. DATABASEN	40
3.1.2. TIDSPERIODE	41
3.1.3. LANDENE	41
3.1.4. INDDDELING I REGIONER	43
3.1.5. VARIABLE	44
3.1.6. FRA FADN DATABASEN TIL BRUGBART DATASÆT	47
3.1.7. ANTAL OBSERVATIONER OG INDDDELING I GRUPPER	48
3.2. STATISTISK BESKRIVELSE AF DATA	51
4. RESULTATER	56
4.1. DEA-ANALYSE PÅ HVERT ÅR FOR SIG	56
4.2. DEA-ANALYSE PÅ DET POOLEDE DATASÆT	62

4.3. MALMQUIST INDEKSET	70
<u>5. DISKUSSION</u>	<u>84</u>
5.1. SAMMENHOLDELSE AF RESULTATERNE I MIN ANALYSE MED ANDRES RESULTATER	84
5.2. FORDELE OG ULEMPER VED DE ANVENDTE MODELLER	85
5.3. FREMTIDIGE UNDERSØGELSER	89
<u>6. KONKLUSION</u>	<u>93</u>
<u>7. LITTERATURLISTE</u>	<u>97</u>
7.1. BØGER OG ARTIKLER	97
7.2. INTERVIEWS	100
7.3. WEBSITES	100
<u>8. FIGURLISTE</u>	<u>102</u>
<u>9. TABELLISTE</u>	<u>102</u>
<u>10. BILAG</u>	<u>103</u>
10.1. BILAG 1 - DE LINEÆRE PROGRAMMER FOR DEA-ANALYSEN PÅ HVERT ÅR FOR SIG	103
10.2. BILAG 2 - DET LINEÆRE PROGRAM FOR DEA-ANALYSEN PÅ DET POOLEDE DATASÆT	104
10.3. BILAG 3 - DE LINEÆRE PROGRAMMER FOR MALMQUIST-ANALYSEN PÅ EU15-LANDENE	104
10.4. BILAG 4 - MANN -WHITNEY-TESTEN	106

1. Introduktion

I dette kapitel vil der blive redegjort for kandidatafhandlingens overordnede område og for motivationen bag at studere og analysere problemstillingen. Herunder formuleres det problem, som gøres til genstand for nærmere undersøgelse i den resterende del af afhandlingen. I forlængelse af problemformuleringen vil der blive foretaget en afgrænsning af denne, og dernæst redegøres for afhandlingens struktur og opbygning. Den sidste del af kapitlet indeholder en kort gennemgang af videnskabsteoretiske overvejelser i forbindelse med udarbejdelse af afhandlingen.

1.1. Indledning

I de sidste 6 årtier har produktivitetsvæksten i landbruget været genstand for intensiv forskning. Forskere, politikere, udviklings- og landbrugsøkonomer har i denne periode undersøgt årsagerne til produktivitetsvækst og produktivitetsforskelle mellem lande og regioner. Man er ikke blot interesseret i at måle niveauer og tendenser i landbrugets produktivitet, men også i hvilke kilder der kan tilskrives denne produktivitetsændring. Generelt regnes en produktivitetsvækst i landbrugssektoren, som en væsentlig faktor til at opfylde kravene til mad og råvarer, der udspringer af en stabil befolkningstilvækst. I løbet af bl.a. 1970'erne og 80'erne er en række større analyser af forskelle i landbrugsproduktiviteten på tværs af landegrænser blevet gennemført, herunder Hayami og Ruttan (1970) og Lau og Yotopoulos (1989).

I de seneste årtier er antallet af artikler og studier, der undersøger forskelle i landbrugsproduktiviteten og vækstrater på tværs af landegrænser udvidet betydeligt. F.eks. gennemfører Fulginiti og Perrin (1997) en landbrugsundersøgelse for at se på ændringer i landbrugsproduktiviteten for 18 af de mindst udviklede lande (LDC lande) i perioden 1961-1985. De fandt ud af, at produktiviteten i disse lande i gennemsnit er faldet. Stigningen i antallet af analyser indenfor landbrugsproduktivitet er hovedsageligt drevet af tre faktorer. For det første skyldes det en større tilgængelighed af paneldata, som det der produceres af FAO³ og FADN⁴. Dels skyldes det udvikling af nye empiriske teknikker og analytisk software til at analysere denne type data, såsom Data Envelopment Analysis (DEA) og Stochastic Frontier Analysis (SFA), der f.eks. er

³ FAO = The Food and Agriculture Organisation of the United Nations

⁴ FADN = The Farm Accountancy Data Network

beskrevet af Coelli et. al. (1998). For det tredje er der et ønske om at vurdere, i hvilken grad den grønne revolution⁵, og andre programmer, har forbedret produktiviteten i landbruget.

1.2. Motivation

Motivationen bag emnevalget i denne afhandling er, at den økonomiske globalisering i de seneste årtier har ført til en øget konkurrence i landbrugssektoren. Globaliseringen har forøget bedriftenes konkurrence ikke bare indenfor egne grænser, men også på tværs af landegrænser. Det gælder om at have den bedste vare til den rette pris, samt at vide mest og være bedst til at udnytte den viden man har. Indenfor landbrug kunne det f.eks. være viden om et nyt såningsprodukt eller udvikling af en ny maskine, som giver en konkurrencefordel ift. andre bedrifter. Derudover er produktivitetsstigninger en forudsætning for øget vækst og velfærd, og produktivitet kan anskues som forholdet mellem produktion og ressourceforbrug, hvilket i de seneste år har været et ledelsesmæssigt fokusområde. Derfor er dette indsatsområde en af de vigtigste kilder til en bæredygtig organisatorisk effektivitet, og en systematisk forståelse af de faktorer, der påvirker produktiviteten, er særdeles vigtig.

1.3. Problemformulering

I denne afhandling er hovedformålet at måle og sammenligne niveauer og udviklingstendenser for produktionen og produktiviteten indenfor landbruget i den Europæiske Union (EU) i perioden 1995-2009. Til at undersøge dette vil jeg anvende produktionsanalysen Data Envelopment Analysis (DEA), samt Malmquists produktivitetsudviklingsindeks.

Jeg vil i denne afhandling beskrive og forklare de to analysemetoder, og Malmquist indekset vil blive dekomponeret multiplikativt i to delkomponenter, der beskriver ændringer i *Technical change* hhv. *Catch-up* effekter.

Med udgangspunkt i ovenstående teoretisk begreber ønsker jeg at analysere et konkret datasæt bestående af paneldata for generel landbrugsregnskabsstatistik hentet i Farm Accountancy Data Network (FADN) databasen fra Europa Kommissionens hjemmeside⁶. I det anvendte datasæt består hvert land af en eller flere regioner.

⁵ Den grønne revolution henviser til en række forsknings, udviklings og teknologioverførsels initiativer, der forekom mellem 1940'erne og slutningen af 1970'erne, hvilket øgede landbrugsproduktionen i verden, begyndende mest markant i slutningen af 1960'erne.

⁶ Data kan hentes på hjemmesiden: http://ec.europa.eu/agriculture/rica/database/database_en.cfm

De spørgsmål jeg ønsker at besvare i forbindelse med analysen kan opstilles i nedenstående punkter:

- Hvilke lande i EU er ud fra valgte inputs og outputs mest efficiente?
 - Hvordan har efficiensen udviklet sig?
- Hvilke landes regioner bruges hyppigst som benchmark?
- Er der sammenfald mellem de 2 udvidelser af EU (i 2004 og 2007) og produktiviteten af de "gamle" EU-lande⁷?
 - Hvordan har produktiviteten udviklet sig?
 - Hvordan har *Catching-up* effekten og randeffekten udviklet sig?

1.4. Afgrænsning⁸

Der foreligger et omfattende modelapparat, der kan benyttes i forbindelse med produktivitetssanalyser. Afhandlingen er søgt afgrænset, så der kun forekommer beskrivelse og analyse af forhold og modeller, som jeg anser for værende relevante ift. afhandlingens problemstilling. Jeg har i denne afhandling valgt at anvende efficiensanalysen (benchmarkinganalysen) DEA, da der ved denne metode ikke er behov for eksplicit at angive en matematisk form for produktionsfunktionen, som det f.eks. er tilfældet for SFA (Stochastic Frontier Analysis) og COLS (Corrected Ordinary Least Squares).

Desuden er DEA særdeles brugbar, da de fundne benchmarks beskriver konkrete forbilleder i modsætning til i SFA (Stochastic Frontier Analysis), hvor benchmarks er et slags "gennemsnit af de bedste" og i COLS (Corrected Ordinary Least Squares), hvor benchmarks fremkommer via en regression. DEA gør det herved nemmere at forholde sig til et benchmark, hvilket har stor betydning i en læringsproces.

Jeg vil i denne afhandling bruge Malmquist-analysen til at finde ændringer i produktiviteten, da produktivitetsændringen kan dekomponeres i to led, samt at Malmquist indekset bruger den underliggende teori fra DEA. Derudover har Malmquist indekset ift. andre indeks som f.eks. Fisher og Paasche den fordel, at der ikke kræves informationer om priser på input og output eller antagelser om teknologi og opførsel for denne.

⁷ Lande som er medlem af EU per 1. januar 1995.

⁸ Kilde til dette afsnit: Konkurrence- og Forbrugerstyrelsen (2010)

I afhandlingen har jeg valgt at bruge data fra FADN databasen, da denne database i generel opfattelse er præget af god kvalitet, samt at det er muligt at få landbrugsresultater for alle EU-lande. Desuden har jeg i nærværende afhandling (delvist pga. begrænset tilgængelighed af data) afgrænset analysen til at omhandle landbrugssektoren i EU i tidsperioden fra 1995 til 2009. Endvidere er mit hovedfokus på de gamle lande i EU. Yderligere kommentarer om valg og overvejelser angående det anvendte datasæt vil blive uddybet i kapitel 3.

1.5. Afhandlingens opbygning

Afhandlingen er inddelt i 4 hovedkapitler, der beskriver de anvendte teorier, det anvendte datasæt, de empiriske resultater på baggrund af det givne datasæt, samt et kapitel hvor der bl.a. diskuteres fordele og ulemper af de anvendte modeller, og hvor de fremkomne resultater sammenlignes med, hvad andre indenfor litteraturen er kommet frem til.

Første hovedkapitel, kapitel 2, er inddelt i to afsnit, hvor de antagelser der ligger til grund for de anvendte produktivitetsanalyser bliver introduceret. I afsnittene om DEA og Malmquist indekset forklarer og illustrer jeg teorierne via et tal eksempel.

Kapitel 3 er ligeledes inddelt i to sektioner. I første del beskrives valg og overvejelser ift. udvælgelsen af data, og anden del indeholder en statistisk beskrivelse af det anvendte datasæt.

I det næste hovedkapitel, kapitel 4, gives en detaljeret gennemgang af resultaterne af DEA-analysen på hhv. hvert år for sig og hele datasættet på en gang. Kapitlet afsluttes med en gennemgang af resultaterne for Malmquist indekset og dets komponenter.

I kapitel 5 diskuterer jeg bl.a. fordele og ulemper ved de anvendte modeller i kapitel 2, og de fundne resultater i kapitel 4 sammenlignes med hvad andre er kommet frem til. Derudover ser jeg på, hvilke fremtidige undersøgelser der vil være oplagte at foretage i forlængelse af analysen i denne afhandling.

Ud fra teorien i kapitel 2, databeskrivelsen i kapitel 3, resultaterne fremkommet i kapitel 4 og diskussionen i kapitel 5 vil jeg, som sidste del af afhandlingen drage mine konklusioner sammenholdt med problemformuleringen.

1.6. Videnskabsteoretiske overvejelser

Problemstillingen i denne afhandling løses ved anvendelse af teori fra DEA- og Malmquist-analysen, og en primært positivistisk tilgang vil blive brugt. Jeg har valgt at benytte mig af en deduktiv fremgangsmåde til indsamling og udvikling af viden. Dvs., at jeg med udgangspunkt i teorien vil sige noget om empirien. Mit belæg i denne afhandling er kvantitative data omhandlende landbrugsresultater fra landene i EU (fra FADN-databasen). Det er altså ud fra den grundlæggende teori i DEA og Malmquist, at jeg ift. disse data ønsker at komme med konkluderende bemærkninger ift. problemstillingen.

2. Teori

Dette kapitel er opdelt i to sektioner. I første del introducerer jeg en teoretisk beskrivelse af produktivitetsanalysemodellen DEA, hvor en konstant skalaafkast (CRS) produktionsteknologi forudsættes (forklares yderligere i afsnit 2.1.1.1). For at give en større intuition af teorien, vil modellen blive gennemgået i et eksempel. Efter gennemgangen af DEA modellen vil jeg beskrive Malmquist indekset (der anvender DEA teorien), som angiver en produktivitetsændring mellem to tidsperioder. Denne sektion vil også blive afsluttet med et eksempel. Jeg har i både DEA- og Malmquist-analysen valgt at ligge mit fokus på de outputorienterede modeller, da det er en fair antagelse, at man indenfor landbrug (som er den industri, hvori min analyse vil foregå) normalt prøver at maksimere output fra et givent sæt af input snarere end det modsatte⁹.

2.1. DEA modellen

I dette afsnit beskrives det teoretiske fundament for den anvendte analyseteknik DEA. Der vil ydermere fremkomme eksempler, som understøtter og forklarer DEA. De primære kilder til dette afsnit er Coelli et. al. (1998), Thanassoulis (2001), Bramsen og Nielsen (2004) og Portela et. al (2008).

Den grundlæggende idé i moderne benchmarking er sammenligning med *best practice*. Best practice kan beskrives, som det bedste man i praksis kan opnå. I stedet for at spekulere i hvad der teoretisk kan opnås, estimeres dette ved at se på, hvad konkrete forsyninger reelt har formået. I en benchmarkinganalyse gennemføres en systematisk undersøgelse på et passende aggregeret niveau af et stort antal virksomheder.

Til en sådan analyse er DEA en af de mest udbredte metoder. DEA er en matematisk metode, der anvender lineær programmering til at sammenligne en række forsyninger med hinanden og bestemme effektiviteten af de forskellige forsyninger. DEA blev oprindeligt udviklet som en metode til at vurdere efficiensen for sammenlignelige organisatoriske enheder så som banker, skoler, hospitaler og restauranter.

⁹ Der vil selvfølgelig altid findes undtagelser. F.eks. hvis en landmand er forpligtet til at opfylde bestemte outputkvoter, og dermed vil forsøge at minimere sine input.

Den teoretiske baggrund for udviklingen af DEA startede i 1957, da Farrell foreslog den stykvis lineær konvekse indhylningsstilgang til at estimere *den efficiente rand* (bliver uddybet i afsnit 2.1.1.1). I årtierne efter Farrell foreslog akademikere forskellige matematiske programmeringsmetoder, men ingen af fremgangsmåderne fik stor opmærksomhed indtil artiklen af operationsanalytikerne Charnes et. al. (1978), hvor udtrykket *Data Envelopment Analysis* (DEA) for første gang blev brugt.

Charnes et. al. (1978) opstiller problemet som en række lineære programmeringsproblemer og gør dermed Farrells teori operationel. Mere specifikt foreslår Charnes et. al. (1978) en DEA model, som antager konstant skalaafkast (CRS). Siden da er der blevet skrevet en lang række papirer og bøger, der har anvendt og udvidet DEA modellen. Bl.a. har Färe et. al. (1983) og Banker et. al. (1984) beskæftiget sig med modeller med variabelt skalaafkast (VRS).

2.1.1. Matematisk udledning af DEA

I dette afsnit vil den matematiske udledning af DEA modellen blive gennemgået, mens længere forklaringer på løsningsmetoden følger i det næste afsnit.

Som skrevet kort i forrige afsnit, er DEA modellen udviklet med det formål at kunne måle den relative produktivitet for en række homogene produktionsenheder/beslutningsenheder (kaldet DMU - Decision Making Units). Kravet om homogenitet (homogenitetskravet) skal forstås i den forstand, at hver DMU bruger de samme ressourcer til at skaffe de samme resultater omend i varierende mængder. Denne produktionsteknologi kan man vha. DEA analysere og finde den relative produktivitet for. Hver DMU, k , kan beskrives som et produktionsplan $(x_1^k, \dots, x_m^k, y_1^k, \dots, y_r^k)$, hvor m input $x_i^k \in R_+^m$ transformeres til r output $y_j^k \in R_+^r$. I DEA er produktionsressourcerne lig input, og produktionens outcome er lig output. Transformationen af disse foretaget på alle DMU'er er illustreret nedenfor.

Figur 2.1: Transformation af input til output



Kilde: Egen tilblivelse

Når man analyserer en industri, er selve identifikationen af input og output særdeles vigtig. Det gælder om, at input indeholder alle ressourcer, som bruges til at producere, og output skal afspejle, det en DMU producerer.

DEA er særdeles anvendelig, da metoden gør det muligt at måle produktiviteten ved multiple input og output, selvom der ikke er nogen fast definerede priser eller vægte på disse. DEA finder disse vægte, som en del af løsningen sammen med efficiensgraden. DEA metoden er så udbredt, da metoden indebærer, at de forskellige input og output kan opgøres i forskellige måleenheder og benævnelser. Metoden kan selv korrigere herfor. Det afgørende er, at de forskellige input og output er målbare. Ydermere analyserer DEA vha. termer som "bedst" eller "efficient" snarere end en gennemsnitlig præstation, hvilket gør DEA yderst anvendelig, når det drejer sig om at måle performance internt i en industri.

2.1.1.1. Mulighedsområde og den efficiente rand

Tidligere i dette kapitel nævnte jeg produktionsteknologi, der er et udtryk for de samlede forudsætninger og muligheder, som produktionsvirksomheder i en bestemt industri er underlagt. Ift. benchmarking er opgaven at finde den DMU, som er bedst til at udnytte teknologien.

Den matematiske definition af mulighedsområdet eller den empiriske referenceteknologi T er:

$$T = \{(x, y) \in R_+^{m+r} | x \text{ kan producere } y\}$$

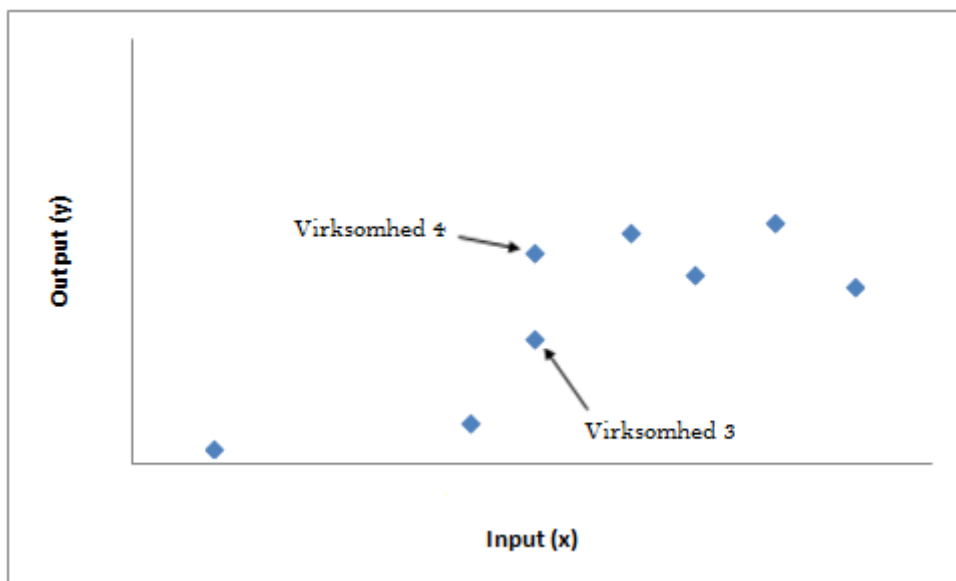
$$x \text{ inputvektor, } m \text{ input, } x = \{x_1, \dots, x_m\}$$

$$y \text{ outputvektor, } r \text{ output, } y = \{y_1, \dots, y_r\}$$

For en given DMU, k , antages det, at $x_i^k \geq 0 \forall i = 1, \dots, m$ og $y_j^k \geq 0 \forall j = 1, \dots, r$, samt at hver DMU har mindst én positiv input- og én positiv outputværdi. Mulighedsområdet T indeholder da mulige korrespondancer af input $x \in R_+^m$, der kan producere output $y \in R_+^r$, dvs. T indeholder altså ikke negative værdier.

Jeg har valgt at illustrere teknologien på nedenstående figur, hvor der for simplicitetens skyld kun optræder ét input og ét output.

Figur 2.2: Produktionsteknologi for ét input og ét output



Kilde: Egen tilblivelse

På figuren er illustreret en række virksomheder med samme type produktion men med forskellige input- og outputkombinationer. Virksomhederne har ikke klaret sig lige godt alle sammen til trods for, at jeg antager, at de har haft mulighed for at anvende den samme teknologi.

For at fastslå hvilke virksomheder der er efficiente, bruges kriteriet at ”producere mere med mindre”. Mere nøjagtigt gælder der, at en beslutningstager er efficient, hvis den ikke er domineret af en anden beslutningstager. Dette er gældende, hvis der ikke eksisterer en anden beslutningstager, der enten har:

1. mere output med samme eller mindre input eller
2. samme output med mindre input

I forbindelse med ovenstående figur er virksomhed 4 således efficient, hvorimod virksomhed 3 er inefficient, idet virksomhed 4 har opnået et større output med akkurat det samme input (virksomhed 4 ligger lodret over virksomhed 3).

Den matematiske definition på efficiens er, at en given virksomhed (DMU) $k \in T$ med inputvektoren $x^k = (x_1^k, \dots, x_m^k)$ og outputvektoren $y^k = (y_1^k, \dots, y_r^k)$ sammensat i vektoren $z^k = (z_1^k, \dots, z_i^k, \dots, z_{m+r}^k)$ er efficient, hvis der ikke findes en anden DMU l med $z^l \in T$, hvor

$z_i^l > z_i^k \exists i \wedge z_i^l \geq z_i^k \forall i$. De matematiske formler betyder, at de DMU'er, der anvender den mindste kombination af m input til at producere en given mængde af de r output, er efficiente. Sagt på en anden måde, en DMU er efficient, hvis og kun hvis ingen af dets input eller output kan forbedres uden at forværre nogle af sine andre input eller output.

Anvendes DEA modellen på et datasæt med n DMU'er, kan det klarlægges, hvilke DMU'er som er hhv. efficiente og inefficiente. Dette gøres ifølge ovenstående matematiske definition på efficiens. De efficiente DMU'er indhyller mulighedsområdet, og dette er defineret som:

$$T(x, y) = \left\{ (x, y) \left| \sum_{k=1}^n \lambda^k y^k \geq y; \sum_{k=1}^n \lambda^k x^k \leq x; \lambda^k \geq 0 \forall k = 1, \dots, n \right. \right\}$$

Området er underlagt følgende antagelser:

1. Det er muligt at gøre ingenting:

Origo (nulpunktet) indgår i mulighedsområdet.

2. Fri bortskaffelse af input:

Hvis $(x', y') \in T$ og $x \geq x'$, så vil $(x, y') \in T$. $x \geq x'$ betyder, at mindst ét element i x er større end det tilsvarende element i x' . Fri bortskaffelse af input siger altså, at hvis det er muligt at producere y' enheder med x' input, er det også muligt at producere y' enheder med x input, hvor $x \geq x'$. Flere input kan altid producere minimum samme output.

3. Fri bortskaffelse af output:

Hvis $(x', y') \in T$ og $y \leq y'$, så $(x', y) \in T$. Forklaringen af denne antagelse følger helt analogt med antagelse 2. Det er altid muligt at producere mindre output med samme inputniveau.

4. "No free lunch":

$(x', 0) \in T$, men hvis $y' \geq 0$ så $(0, y') \notin T$. Der kan ikke produceres noget output, uden at der bruges noget input.

5. Konstant skalaafkast (CRS):

Hvis $(x', y') \in T$, så gælder der for $\lambda > 0$, at $(\lambda x', \lambda y') \in T$. Med konstant skalaafkast er der tale om en teknologi, hvor der frit kan skaleres en given produktionsplan op og ned, men stadig opnå mulige produktionsplaner.

6. Konveksitet:

Hvis $(x, y) \in T$ og $(x', y') \in T$ så vil $(\lambda(x, y) + (1 - \lambda)(x', y')) \in T \forall \lambda \in [0; 1]$. Hvilket betyder, at hvis der findes 2 forskellige efficiente produktionsplaner (x, y) og (x', y') , er det også muligt at lave alle mulige konvekse kombinationer ("vejede gennemsnit") af de 2 produktioner. Dvs., at antallet af brugbare produktionsmuligheder er konveks.

7. Det minimale ekstrapolationsprincip:

Alle DMU'er $\{(x^k, y^k), k = 1, \dots, n\} \in T$, og T er den mindste delmængde af R_+^{m+r} , som opfylder ovenstående antagelser.

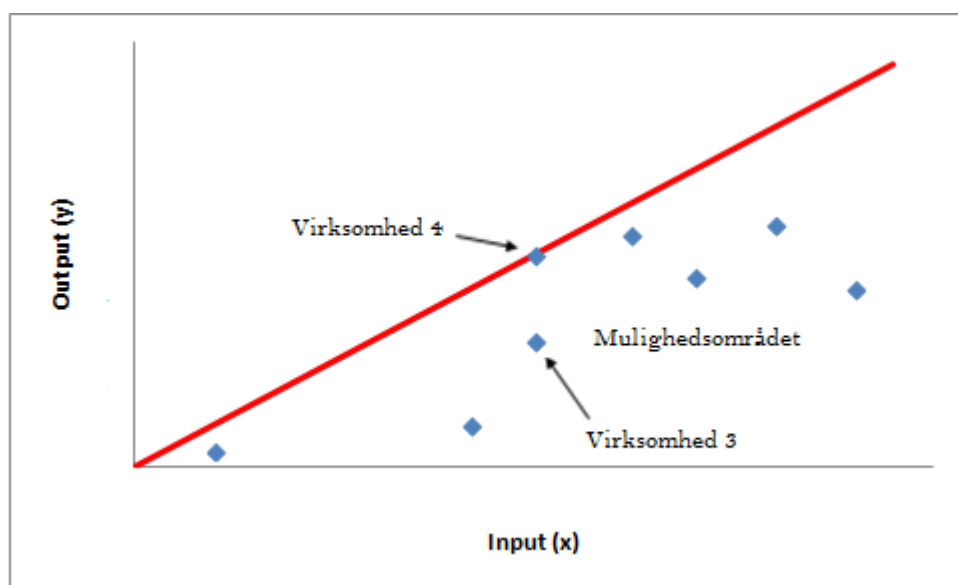
Grunden til at jeg i denne afhandling antager en CRS teknologi er, at CRS er særdeles anvendeligt inden for produktionsområder, hvor input indgår i et fast miks for at producere output, som f.eks. i landbrug. Her kan der skaleres op og ned i "opskrifterne" efter egen kapacitet, og det er rimeligt at antage, at output fordobles, hvis input fordobles, hvilket er selve definitionen på konstant skalaafkast. Ydermere er CRS en berettiget antagelse, da landbrug i generel opfattelse er præget af fuld konkurrence. Under fuld konkurrence giver det ikke mening at "belønne" enheder for ikke at være i den optimale størrelse, som en VRS model ville gøre. Derimod vil en CRS model under fuld konkurrence tvinge alle DMU'er til at have det rette bytteforhold mellem input og output, og efficiensscoren vil afsløre, hvis de ikke har.

Med disse antagelser er udledt et lukket produktionsmulighedsområde, som er indhyllet/udspændt af en front bestående af efficiente virksomheder og konvekse kombinationer af disse. Mængden af alle de konvekse kombinationer af dominerende/efficiente produktionsmuligheder kaldes *den efficiente rand (isokvanten)*, og det er disse, der benyttes som benchmarks. DMU'er, som ikke ligger på den efficiente rand, kaldes inefficente, og disse måles i DEA op imod den efficiente rand.

Ift. eksemplet gennemgået tidligere i afsnittet vil virksomhed 4 fremstå, som den virksomhed der har den højeste produktivitet (er efficient) - producerer mest med mindst input (har den største

værdi af output/input ratio). Mulighedsområdet er derfor udspændt af virksomhed 4, og da der antages konstant skalaafkast, kan der, som definitionen siger skaleres op og ned efter eget ønske. Dermed består hele den røde linje, *den efficiente rand*, af mulige produktionsplaner. Mulighedsområdet er udspændt af en enkelt virksomhed, da der i eksemplet kun indgår ét input og ét output. I øvrigt bliver virksomhed 3 under konstant skalaafkast erklæret inefficiant, da den domineres af virksomhed 4. Mulighedsområdet i eksemplet er illustreret på nedenstående figur.

Figur 2.3: Udledt mulighedsområde for ét input og ét output



Kilde: Egen tilblivelse

2.1.1.2. Farrells efficiensindeks

I benchmarkinganalyser vurderes efficiensscoren for hver enkelt DMU ved afstanden til den efficiente rand af produktionsmulighedsområdet, som angiver best practice. Denne afstand kan måles på mange måder, men den helt dominerede metode i litteraturen er brugen af Farrells efficiensindeks. Ideen i Farrells *inputorienterede efficiensindeks* er at undersøge, om det er muligt at reducere input uden at ændre output. Når man har multiple input og output, ser man på en forholdsmæssig nedsættelse af alle inputs.

Det inputorienterede Farrell efficiensindeks af et plan (x, y) ift. T er defineret som:

$$E = \min\{\theta | (\theta x, y) \in T\}$$

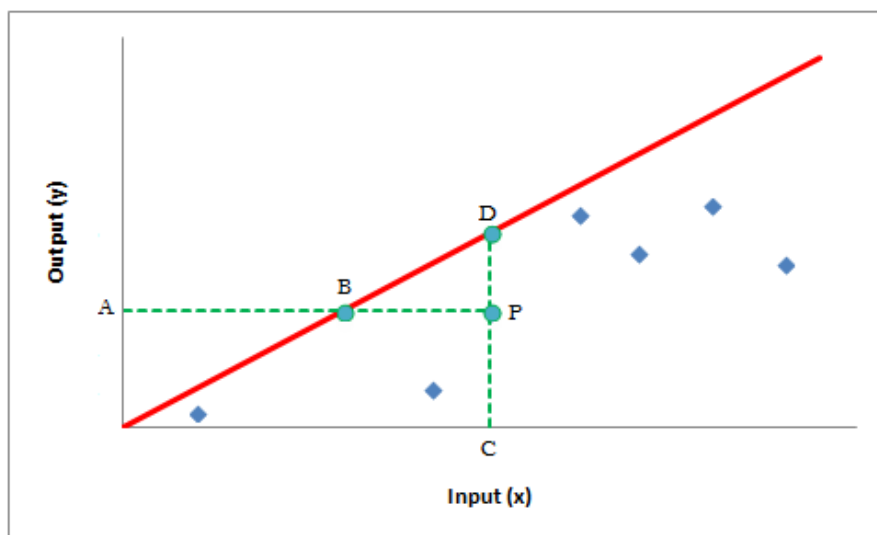
E måler den mulige %-vise besparelse for en DMU ud fra dens anvendte inputniveau ift. de efficiente DMU'er. Hvis $E = 0,9$, indikeres det, at man kunne have sparet 10% af det anvendte inputniveau og stadig have produceret samme output. For det *outputorienterede Farrell efficiensindeks* undersøges det, om der kan øges output uden at ændre på input, og indekset er defineret som:

$$F = \max\{\phi | (x, \phi y) \in T\}$$

Efficiensgraden angiver, hvor meget større output burde have været, givet det nuværende forbrug af input. Hvis $F = 1,3$, betyder det, at man kunne have udvidet output med 30 % uden at bruge yderligere ressourcer. I de to formler er θ ($\theta \in]0; 1]$) og ϕ ($\phi \in [1; \infty[$) den hhv. mindst og størst mulige skalar, der enten skal ganges på inputvektoren eller outputvektoren, således at man enten kommer ind eller ud på randen og stadig producerer y eller forbruger x .

Nedenfor har jeg illustreret, hvordan Farrells to efficiensindeks (E og F) ser ud for virksomhed 3 (punkt P) i det tidligere gennemgået eksempel.

Figur 2.4: Farrells input- og outputorienterede efficiensindeks



Kilde: Egen tilblivelse

På figuren er Farrells input- og outputorienterede efficiensindeks hhv. $E = \frac{\overline{AB}}{\overline{AP}}$ og $F = \frac{\overline{CD}}{\overline{CP}}$. Det noteres yderligere, at $E = \frac{1}{F} \left(\frac{\overline{AB}}{\overline{AP}} = \frac{\overline{CP}}{\overline{CD}} \right)$, hvilket i øvrigt altid er tilfældet, når der antages konstant skalaafkast. For Farrells inputorienterede efficiensindeks gælder, at jo tættere E er på 0, jo mindre

efficient er en virksomhed. For Farrells outputorienterede efficiensindeks gælder, at jo tættere F er på uendelig, jo mindre efficient er en virksomhed. Der gælder for begge efficiensindeks, at de efficiente DMU'er antager værdien 1.

2.1.2. Løsningsmetode

Generelt skelnes der mellem input- og outputorienterede DEA modeller. Som netop gennemgået giver den inputorienterede DEA model en efficiensscore, der angiver, hvor meget den aktuelle DMU kan reducere input med uden at afvige fra det producerede outputniveau. Den outputorienterede DEA model kaldes det duale problem til det inputorienterede, hvor det er inputniveauet, som holdes fast på et bestemt niveau, mens det klarlægges, hvor meget outputniveauet kan forøges. Ændringerne i hhv. inputniveau (reduktion) og outputniveau (forøgelse) vil ske proportionalt i henhold til DEA teorien.

2.1.2.1. CCR modellen

Modellen udviklet af Charnes et. al. (1978), den såkaldte CCR model, bliver anvendt som den grundlæggende DEA model. Charnes et. al. omskrev efficiensproblemet til en række lineære programmeringsproblemer, og denne omskrivning vil i dette afsnit blive gennemgået.

Som skrevet i starten af dette kapitel kommer DEA-analysens popularitet af, at det er muligt at måle produktiviteten ved multiple input og output, selvom der ikke er nogen fast definerede priser eller vægte på disse. Dermed opstår et behov for at sammenveje de enkelte input og output, så man får et enkelt mål for disse, og derfor defineres produktiviteten/efficiensmålet, som angiver forholdet mellem vægtede input og output.

$$P = \frac{\text{vægtet output}}{\text{vægtet input}} = \frac{u \cdot y}{v \cdot x}$$

, hvor $x, v \in R^m$ og $y, u \in R^r$.

I ligningen bruges u og v som vægte for hhv. output og input, hvorimod y og x anvendes som vektorer af hhv. output og input. Ud fra disse vægte og vektorer kan Produktivitsbrøken, P , skrives som:

$$P = \frac{\sum_{j=1}^r u_j y_j}{\sum_{i=1}^m v_i x_i}$$

I formelen bliver x og y opsplittet, hvilket gør det muligt at fastslå, hvor produktiviteten er for lav. Målet er nu at identificere de vægte, som maksimerer P for en given produktionsplan (DMU (k_0)) under bibetingelse af (ubb.), at alle brøker bliver normaliseret. Dette gøres ved løsning af det ikke lineære programmeringsproblem, som er opskrevet nedenfor.

$$\max_{u,v} \frac{\sum_{j=1}^r u_j y_j^{k_0}}{\sum_{i=1}^m v_i x_i^{k_0}}$$

ubb.

$$\frac{\sum_{j=1}^r u_j y_j^k}{\sum_{i=1}^m v_i x_i^k} \leq 1 \forall k = 1, \dots, n \wedge u_j, v_i \geq 0$$

I ovenstående ligningssystem er det k_0 , der, for alle i input og j output med de tilhørende vægte, evalueres. Enheden evalueres op imod alle n DMU'er, hvis forhold mellem output og input ikke må overstige 1 for nogen produktionsenhed ved de valgte vægte. For maksimeringsproblemet er løsningen (v^*, u^*) , hvilket medfører, at også $(\lambda v^*, \lambda u^*)$ er løsning for alle $\lambda > 0$. Der er altså uendeligt mange løsninger, og derfor omskrives ovenstående programmeringsproblem til et lineært problem ved enten at fastholde input eller output.

$$\min_{u,v} \sum_{i=1}^m v_i x_i^{k_0}$$

ubb.

$$\sum_{j=1}^r u_j y_j^{k_0} = 1$$

$$\sum_{i=1}^m v_i x_i^k - \sum_{j=1}^r u_j y_j^k \geq 0 \forall k$$

$$v_i, u_j \geq 0$$

I ligningssystemet findes der ikke negative vægte v_i og u_j , hvor det vægtede input minimeres ubb., at output fastholdes på et efficient niveau. Helt konkret gøres dette ved at summere til én for den evaluerede DMU, hvorefter alle brøkerne normaliseres (næstsidste bibetingelse). Der lægges selvfølgelig mærke til, at maksimeringsproblemet nu er blevet et minimeringsproblem af input. Til multiplikatorproblemet ovenfor findes der altid et dualt problem, som kan omskrives vha. Farrells efficiensscore/indeksmål. Hermed opstår den outputorienterede DEA model under antagelse af CRS (konstant skalaafkast).

$$\max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ubb.

$$\sum_{k=1}^n \lambda^k x_i^k \leq x_i^{k_0}, i = 1, \dots, m$$

$$\sum_{k=1}^n \lambda^k y_j^k \geq \phi y_j^{k_0}, j = 1, \dots, r$$

$$\lambda^k \geq 0, k = 1, \dots, n$$

Det lineære programmeringsproblem indeholder $m + r$ bibetingelser med et positivitetskrav på alle vægte. I ligningssystemet er det igen k_0 , som evalueres. Programmet løses ved at finde den maksimale skalar ϕ , hvilken proportionalt øger outputvektoren $y_j^{k_0}$, så denne stadig anvender den givne inputmængde. Modellen skal løses for alle n forsyninger, således at der er bestemt et ϕ og en λ -vektor for hver af de n forsyninger. Scorerne for ϕ gør det muligt, at rangordne de inefficiente DMU'er, så de enheder med højest ϕ er mindst efficiente.

Løsningen angiver derudover implicit de tidligere nævnte benchmark, som oftest udgøres af en kombination af flere af de efficiente DMU'er og ikke nødvendigvis af en allerede eksisterende DMU. Benchmark findes ud fra den optimale skaleringsfaktor λ^* givet af løsningen. De enheder, som indgår i benchmark, kaldes for referencemængden (*efficient peers*) Q^{k_0} , og de er defineret nedenfor.

$Q^{k_0} = \{k_0 | \lambda^{*k_0} > 0\}$, for $k_0 \in \{1, \dots, n\}$. Inefficiente DMU'er/observationer (har $\phi > 1$, og dermed $\frac{1}{\phi} < 1$) kan aldrig optræde som benchmark for andre DMU'er og har derfor lambda'er (vægte) lig 0, når alle observationer evalueres.

2.1.3. Eksempel på den outputorienterede DEA model

Jeg har i dette afsnit valgt at forklare DEA teorien vha. et to dimensionalt eksempel. Jeg kan hermed vise grafiske afbildninger, hvilket giver en bedre intuitiv forståelse af teorien.

Jeg ser på et eksempel med 7 DMU'er, der hver med ét input x producerer to output y_1 og y_2 . De 7 beslutningsenheder antages at producere samme typer output ved anvendelse af samme type input, jf. homogenitetskravet. Input og output for de 7 DMU'er er angivet i følgende tabel.

Tabel 2.1: Data for DEA eksempel

DMU	Input (x)	Output 1 (y_1)	Output 2 (y_2)	$\frac{\text{Output 1}}{\text{Input}} = \frac{y_1}{x}$	$\frac{\text{Output 2}}{\text{Input}} = \frac{y_2}{x}$
A	2	4	3	2	1,5
B	1	5	3	5	3
C	2	14	4	7	2
D	1	8	1	8	1
E	1	5	1	5	1
F	3	12	6	4	2
G	3	9	10	3	3,33

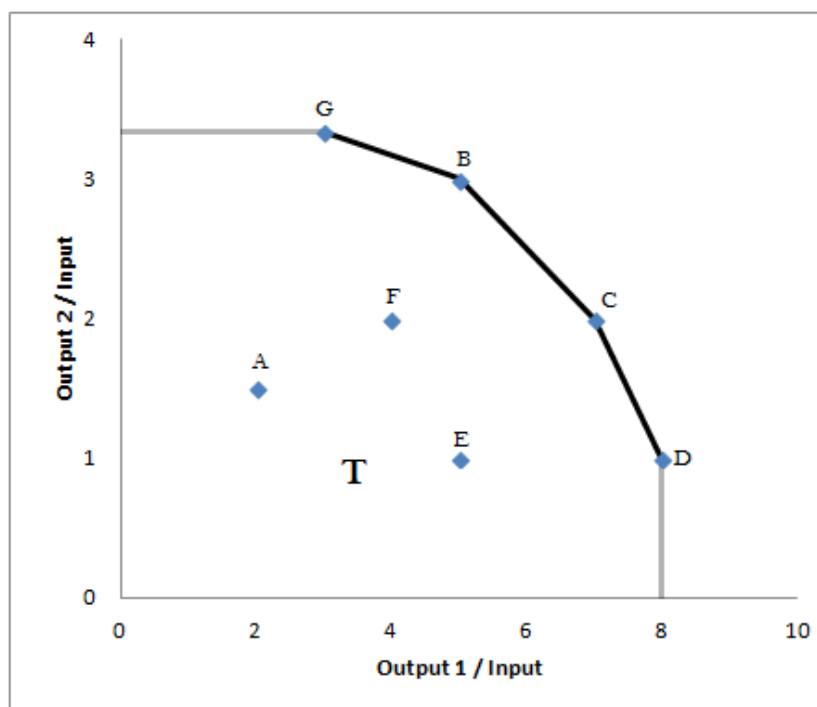
Kilde: Egen tilblivelse

Da DMU'erne har forskellige inputmængder, har jeg under antagelse af konstant skalaafkast "normeret" disse ved at dividere de to output med inputmængden¹⁰ (se de 2 yderste kolonner i tabellen). Derfor kan der laves en grafisk afbildning af T . De 7 DMU'er indtegnes i et koordinatsystem, og de efficiente DMU'er er dem, som har den højeste kombination af output ved brug af 1 stk. input. Disse forbindes for derved at danne den efficiente rand, dvs. den efficiente rand er en stykvis lineær overflade af de efficiente DMU'er. Da antagelsen om fri bortskaffelse af output

¹⁰ Dette er kun tilladt/passende pga. antagelsen om CRS (tilladt at skalere op og ned).

også skal være opfyldt, består den efficiente rand ydermere af det vandrette linjestykke ud fra *G* og det lodrette linjestykke under *D*. Dette er illustreret på nedenstående figur.

Figur 2.5: Grafisk eksempel på den outputorienterede DEA model

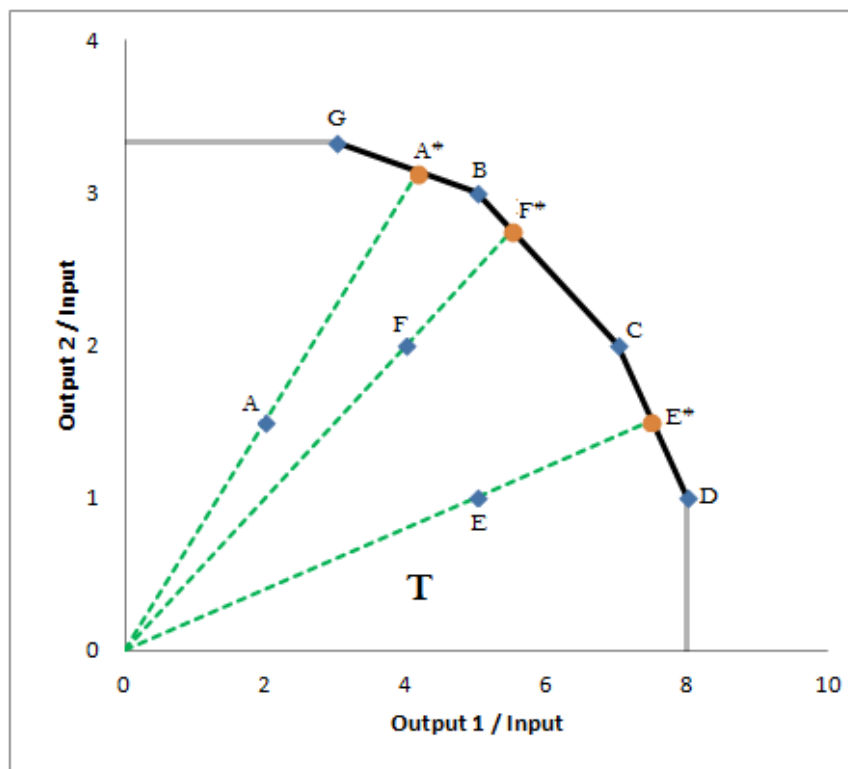


Kilde: Egen tilblivelse

Den efficiente rand dannes af de efficiente DMU'er *B*, *C*, *D* og *G*, som har den største kombination af output ved brug af 1 stk. input. Disse DMU'er har naturligvis et Farrells indekssmål på 1.

DMU'erne *A*, *E* og *F* er, som det fremgår af grafen tydeligvis ikke efficiente, idet de ligger inde i mulighedsområdet *T*. På nedenstående figur har jeg illustreret disse DMU'ers efficiensmål, deres mulige proportionale forøgelse i outputmængden, samt hvilke DMU'er de bliver målt op imod.

Figur 2.6: Referencepunkter for de inefficente DMU'er



Kilde: Egen tilblivelse

Af figuren ses det bl.a., at A^* er referencepunkt for DMU'en A (angiver projektionen af A på den efficiente rand). A^* kaldes benchmark for A, og dette punkt ligger selvfølgelig på den efficiente rand. Referencepunktet A^* angiver i dette tilfælde en hypotetisk observation, som er en kombination af produktionsenhederne G og B. Farrells efficiensindeks giver den skalar, hvorved produktionsenheden A proportionalt skal øge sine 2 output for at blive efficient - for at blive A^* . På grafen er Farrells outputorienterede efficiensindeks for A: $\phi^A = \frac{\overline{OA^*}}{\overline{OA}}$.

I eksemplet gennemgået ovenfor så jeg, at randen, udover at være en konveks kombination af de efficiente enheder, består af et vandret og lodret linjestykke. Hvis der på disse linjestykker findes produktionsenheder vil de få en Farrell efficiens på 1, hvilket medfører en klassifikation som efficiente. Det ses dog tydeligt, at enheder liggende på disse linjestykker ikke er fuldt efficiente, da

der i "knækket" findes en anden enhed, der med samme input producerer mere af det ene output. En produktionsenhed liggende på et af disse linjestykker kaldes en (svagt) domineret enhed¹¹.

CCR programmet for det gennemgåede eksempel kan opskrives, og jeg har valgt at opskrive det for DMU A.

$$\max_{\phi^A, \lambda} \phi^A$$

ubb.

$$\lambda^A + \lambda^B + \lambda^C + \lambda^D + \lambda^E + \lambda^F + \lambda^G \leq 1$$

$$2\lambda^A + 5\lambda^B + 7\lambda^C + 8\lambda^D + 5\lambda^E + 4\lambda^F + 3\lambda^G \geq 2\phi^A$$

$$1,5\lambda^A + 3\lambda^B + 2\lambda^C + \lambda^D + \lambda^E + 2\lambda^F + 3,33\lambda^G \geq 1,5\phi^A$$

$$\lambda^k \geq 0, k = 1, \dots, n$$

Ovenstående ligningssystem indeholder én bibetingelse for input, to bibetingelser for output og ét positivitetskrav for hver af de 7 vægte. Alle input er sat til 1, fordi jeg allerede har taget højde for forskelligt input, da jeg anvender værdierne for $\frac{\text{Output 1}}{\text{Input}} = \frac{y_1}{x}$ og $\frac{\text{Output 2}}{\text{Input}} = \frac{y_2}{x}$ i bibetingelserne for output. Problemet har jeg løst vha. Solveren i Excel, og jeg får:

$$\phi^A = 2,091 \Leftrightarrow \frac{1}{\phi^A} = 0,478, (\lambda^A; \lambda^B; \lambda^C; \lambda^D; \lambda^E; \lambda^F; \lambda^G) = (0; 0,591; 0; 0; 0; 0; 0,409)$$

Af tallene ses, at DMU A er ca. 48 % efficient, hvilket betyder, at A kan øge produktionen af de to output proportionalt med 109,1 % uden at forøge input. Yderligere lægger jeg mærke til, at DMU B og G er de to produktionsenheder, der indgår i A's referencemængde med de angivne skaleringsfaktorer. Det ses, at produktionsenhed B er den mest dominerende i referencemængden, hvilket også svagt kan anes på Figur 2.6. På figuren ligger benchmark for A (A*) nemlig tættere på DMU B end på DMU G. Vægtene for DMU A, som er fundet i DEA modellen, er nedenfor brugt til at beregne koordinaterne for A*.

¹¹ En domineret enhed (kaldes også pareto-inefficient) kan findes ved at tilføje slack variable til det lineære programmeringsproblem, samt kræve at alle vægte er strengt positive. Da vil en efficient DMU få $\phi = 1$, hvis alle slack variable er lig 0. Herved undgås at den efficiente enhed befinder sig på det vandrette eller lodrette stykke af den efficiente rand. Dette er dog ikke noget, som vil blive brugt i analyserne i nærværende opgave.

Output/input ratios	B	G	Benchmark for A (A^*)
---------------------	-----	-----	-----------------------------

$$\begin{matrix} y_1/x \\ y_2/x \end{matrix} \quad 0,591 \times \begin{bmatrix} 5 \\ 3 \end{bmatrix} + \quad 0,409 \times \begin{bmatrix} 3 \\ 3,33 \end{bmatrix} = \quad \begin{bmatrix} 4,182 \\ 3,136 \end{bmatrix}$$

Her er de fundne vægte multipliceret på hhv. B og G 's koordinater, hvormed jeg udregner koordinaterne for A 's benchmark, A^* , til at antage værdierne 4,182 (y_1/x) og 3,136 (y_2/x), hvilket stemmer overens med Figur 2.6. Forskellen på benchmark A^* og det "rigtige" A er som følger:

Output/input ratios	A^*	A	$A^* - A$
y_1/x	4,182	2	2,182
y_2/x	3,136	1,5	1,636

Dette indikerer, at en blanding af de driftsteknikker, som anvendes af B og G , vil resultere i en sammensat hypotetisk virksomhed (A^*), som har større værdier for både y_1/x og y_2/x end A . Tallene viser, at ved vedtagelse af en blanding af de faktiske teknikker, der anvendes ved B og G , skal A kunne øge output y_1 og y_2 med hhv. 2,182 og 1,636 enheder per input. De to vægte summer i øvrigt til 1 pga. den første bibetingelse, og jeg lægger ydermere mærke til, at løsningen er optimal, da jeg har den størst mulige sum af λ 'er uden at bryde bibetingelsen. Nedenfor har jeg listet resultaterne for alle de 7 DMU'er, og disse er også fundet vha. Solveren.

Tabel 2.2: Resultater for DEA eksempel

DMU	ϕ	$\frac{1}{\phi}$	λ^A	λ^B	λ^C	λ^D	λ^E	λ^F	λ^G
<i>A</i>	2,091	0,478	0	0,591	0	0	0	0	0,409
<i>B</i>	1	1	0	1	0	0	0	0	0
<i>C</i>	1	1	0	0	1	0	0	0	0
<i>D</i>	1	1	0	0	0	1	0	0	0
<i>E</i>	1,500	0,667	0	0	0,500	0,500	0	0	0
<i>F</i>	1,375	0,727	0	0,750	0,250	0	0	0	0
<i>G</i>	1	1	0	0	0	0	0	0	1

Kilde: Egne beregninger

Af tabellen ses, at det kun er de efficiente enheder, som indgår som benchmark, hvilket går helt i tråd med teorien om DEA. Det gennemgåede eksempel kunne også have været løst ved at bruge de faktiske værdier for input og output i stedet for de normerede, hvilket ville have givet de præcis samme resultater som ovenfor. Yderligere skal det siges, at vægtene under CRS ikke nødvendigvis skal summere til 1 for hver DMU, som det er tilfældet her.

2.2. Malmquist indekset

I dette afsnit beskrives anvendelsen og teorien bag Malmquists produktivitetsindeks, og jeg vil, som i DEA afsnittet, fremføre et eksempel, som understøtter og forklarer teorien bag Malmquist indekset. De primære kilder til dette afsnit er Coelli et. al. (1998), Thanassoulis (2001) og Rao og Coelli (2003).

Som beskrevet i det forrige afsnit finder DEA efficiensscoren for en vilkårlig DMU i én given tidsperiode. For at tage "skridtet" videre kan man finde produktivitetsændringen for en vilkårlig DMU fra en given periode til en anden. Til at bestemme denne ændring har jeg valgt at anvende Malmquists TFP (Total factor productivity) indeks, som blev opfundet af svenskeren Sten Malmquist.

Oprindeligt foreslog Sten Malmquist i 1953 et mængdeindeks til måling af levestandarden med henblik på en forbrugsanalyse, men senere er Malmquist indekset og dens variationer overvejende blevet anvendt i forbindelse med produktionsanalyser. I starten blev en produktivitetsændring forklaret med tekniske ændringer, men senere er det blevet almindeligt accepteret, at ændringen i efficiens også kan bidrage. En sådan opsplittelse af det oprindelige Malmquist indeks fremmer forståelsen af de afgørende faktorer for at skabe en bedre ydeevne og giver dermed værdifulde oplysninger for ledere og planlæggere i både den private og den offentlige sektor.

Den mest skelsættende artikel om Malmquist indekset er Grosskopf et. al. (1994), som tager Malmquist indeksets TFP vækst defineret i Caves et. al. (1982b) og illustrerer, hvordan komponenterne i Malmquist indekset kan estimeres vha. DEA lignende metoder. Yderligere viser de, hvordan Malmquist indekset kan splittes i forskellige komponenter (efficiensændring og teknisk ændring).

Sammenlignet med andre produktivitetsindeks har Malmquist den styrke, at den ikke kræver adfærdsmæssige forudsætninger såsom omkostningsminimering eller profitmaksimering, hvilket gør den anvendelig i situationer, hvor producenterne mål enten afviger, er ukendt eller ikke er opnået. Yderligere kræver indekset ikke prisoplysninger, hvilket indebærer, at det kan anvendes i situationer, hvor enten priserne ikke eksisterer, forvrænges eller har ringe økonomisk betydning.

2.2.1. Løsningsmetode

Malmquist indekset kan defineres ved brug af en input- eller en outputorienteret strategi. I det følgende har jeg igen valgt at beskrive og anvende den outputorienterede strategi, men den inputorienterede strategi kan defineres og bruges på en tilsvarende måde.

Det outputorienterede Malmquist indeks fokuserer på det maksimale niveau af output, som kan produceres ved brug af givne input og en given produktionsteknologi ift. det observerede niveau af output. Denne produktionsteknologi er underlagt de samme antagelser som i afsnit 2.1.1.

En produktionsteknologi kan defineres vha. output-sæt, $T(x)$, som repræsenterer sættet af alle output vektorer, y , som kan fremstilles vha. input-vektoren, x . Dvs.:

$$T(x) = \{y: x \text{ kan producere } y\}$$

Output distance funktionen er defineret vha. output-sæt, $T(x)$, som:

$$d^0(x, y) = \min \left\{ \phi: \left(\frac{y}{\phi} \right) \in T(x) \right\}$$

Afstandsfunktionen, $d^0(x, y)$, antager en værdi mindre end eller lig med 1, hvis outputvektoren, y , er en del af de mulige produktions sæt, $T(x)$. $d^0(x, y) = 1$, hvis og kun hvis (x, y) ligger på den efficiente rand. I den foreliggende undersøgelse bruger jeg DEA lignende metoder til at beregne disse afstandsmål.

Malmquist indekset måler produktivitetsændringer mellem 2 datapunkter ved at udregne ratio af afstande af hvert datapunkt ift. en fælles teknologi. Jeg lader DMU'erne ($k = 1, \dots, n$) operere i perioderne s (basis perioden) og t , og DMU k_0 bruger input x_t (x består af en vektor med m

forskellige input) til at sikre output y_t (y består af en vektor med r forskellige output) i periode t . Følger man Grosskopf et. al. (1994), er det outputorienterede Malmquist indeks mellem periode s og periode t for k_0 givet ved:

$$MI^{k_0}(y_s, x_s, y_t, x_t) = \left[\frac{d_s^{k_0}(y_t, x_t)}{d_s^{k_0}(y_s, x_s)} \cdot \frac{d_t^{k_0}(y_t, x_t)}{d_t^{k_0}(y_s, x_s)} \right]^{1/2}$$

Malmquist indekset består af afstande, og med outputafstanden betragter man en maksimal proportional udvidelse af produktionsvektoren givet en input vektor. I formelen er notationen, at $d_s^{k_0}(y_t, x_t)$ repræsenterer afstanden mellem observationen (k_0) i periode t ift. periode s ' teknologi. Malmquist indekset er faktisk et geometrisk gennemsnit af to Malmquist indekser, ét som er målt ift. periode s ' teknologi, og ét som er målt ift. periode t 's teknologi. $MI^{k_0} > 1$ betyder en positiv vækst i produktiviteten fra periode s til periode t , hvorimod $MI^{k_0} < 1$ indikerer et TFP fald.

Malmquist indekset kan omskrives vha. simpel algebra, så produktivitetsændringen er en kombination af en produktivitetsændring i hele industrien over tid (teknisk ændring) og en efficiensændring for en given beslutningsenhed:

$$\begin{aligned} MI^{k_0}(y_s, x_s, y_t, x_t) &= \left[\frac{d_s^{k_0}(y_t, x_t)}{d_s^{k_0}(y_s, x_s)} \times \frac{d_t^{k_0}(y_t, x_t)}{d_t^{k_0}(y_s, x_s)} \right]^{1/2} = \left(\frac{d_t^{k_0}(y_t, x_t)}{d_s^{k_0}(y_s, x_s)} \right)^{1/2} \cdot \left(\frac{d_s^{k_0}(y_t, x_t)}{d_t^{k_0}(y_s, x_s)} \right)^{1/2} \\ &= \left(\frac{d_t^{k_0}(y_t, x_t)}{d_s^{k_0}(y_s, x_s)} \right)^{1/2} \cdot \left(\frac{d_s^{k_0}(y_t, x_t)}{d_t^{k_0}(y_s, x_s)} \right)^{1/2} \cdot \left(\frac{d_t^{k_0}(y_t, x_t)}{d_s^{k_0}(y_s, x_s)} \cdot \frac{d_s^{k_0}(y_s, x_s)}{d_t^{k_0}(y_t, x_t)} \right)^{1/2} \\ &= \frac{d_t^{k_0}(y_t, x_t)}{d_s^{k_0}(y_s, x_s)} \cdot \left(\frac{d_s^{k_0}(y_t, x_t)}{d_t^{k_0}(y_s, x_s)} \cdot \frac{d_s^{k_0}(y_s, x_s)}{d_t^{k_0}(y_t, x_t)} \right)^{1/2} \\ &= \underbrace{\frac{d_t^{k_0}(y_t, x_t)}{d_s^{k_0}(y_s, x_s)}}_{\text{Catch-up}} \cdot \underbrace{\left[\frac{d_s^{k_0}(y_t, x_t)}{d_t^{k_0}(y_t, x_t)} \cdot \frac{d_s^{k_0}(y_s, x_s)}{d_t^{k_0}(y_s, x_s)} \right]^{1/2}}_{\text{Technical change}} \end{aligned}$$

Dermed er indekset opsplittet i følgende to indeks.

Catch-up indekset:

$$\frac{d_t^{k_0}(y_t, x_t)}{d_s^{k_0}(y_s, x_s)}$$

Catch-up indekset angiver ændringen i det outputorienterede mål for Farrell efficiens mellem periode s og t . Dvs., at efficiensændringen er lig forholdet mellem Farrells efficiensindeks i hhv. periode t og s . Hvis Catch-up indekset > 1 , er DMU k_0 blevet mere efficient ift. sine konkurrenter i periode t end i periode s (dvs. rykket nærmere den efficiente rand), og det modsatte gælder selvfølgelig, hvis Catch-up indekset < 1 . Hvis Catch-up indekset er lig 1, så vil DMU k_0 i begge perioder ligge lige langt fra eller nok mere sandsynligt ligge på den efficiente rand i begge perioder. Det bemærkes, at de to led som indgår i brøken, er de samme, som man finder under DEA bare i forskellige perioder.

Technical change indekset:

$$\left[\frac{d_s^{k_0}(y_t, x_t)}{d_t^{k_0}(y_t, x_t)} \cdot \frac{d_s^{k_0}(y_s, x_s)}{d_t^{k_0}(y_s, x_s)} \right]^{1/2}$$

Den resterende del af Malmquist indekset kaldes Technical change (eller randeffekt), og er det geometriske gennemsnit af skiftet i teknologi mellem de to perioder, evalueret i både x_t og x_s . Det første led i formlen er udtryk for et ratio mellem afstanden fra k_0 i periode t ift. periode s ' teknologi og afstanden fra k_0 i periode t ift. periode t 's teknologi. De to led hvor k_0 i en periode sammenlignes med teknologien i samme periode, er de samme led som indgår i Catch-up indekset. Det betyder, at man faktisk "kun" skal beregne 4 afstande for at finde Malmquist indekset. Desuden skal det siges, at hvis randeffekten > 1 , registres der for DMU k_0 en teknisk forbedring gående fra periode s til periode t , og det modsatte gælder selvfølgelig, hvis Technical change indekset < 1 .

Afstandene i Malmquist indekset kan findes vha. matematisk programmering eller økonometriske teknikker. Den metode jeg bruger i denne afhandling, er de DEA lignende lineære programmer. De lineære systemer kræver et balanceret panel, dvs. at inddelingen af observationer ikke ændrer sig over tid. For hver observation (DMU) er opgaven altså at beregne de 4 afstande for at finde TFP ændringen mellem 2 perioder. Dette kræver, at man som tidligere skrevet løser 4 lineære programmer under antagelse af konstant skalaafkast (CRS) (forklaret i afsnit 2.1.1.1). En CRS teknologi er udover at være en passende antagelse for landbrugsproduktionen en nødvendighed for Malmquists TFP indeks. Grifell-Tatjé og Lovell (1995) bruger f.eks. et simpelt ét input og ét output eksempel til at illustrere, at Malmquist ikke altid korrekt måler TFP ændringer, når en VRS teknologi er antaget.

De nødvendige lineære programmer til at finde Malmquist indekset er:

(1)

$$[d_t^{k_0}(y_t, x_t)]^{-1} = \max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ubb.

$$\sum_{k=1}^n \lambda_t^k x_{i,t}^k \leq x_{i,t}^{k_0}, i = 1, \dots, m$$

$$\sum_{k=1}^n \lambda_t^k y_{j,t}^k \geq \phi y_{j,t}^{k_0}, j = 1, \dots, r$$

$$\lambda_t^k \geq 0, k = 1, \dots, n$$

(2)

$$[d_s^{k_0}(y_s, x_s)]^{-1} = \max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ubb.

$$\sum_{k=1}^n \lambda_s^k x_{i,s}^k \leq x_{i,s}^{k_0}, i = 1, \dots, m$$

$$\sum_{k=1}^n \lambda_s^k y_{j,s}^k \geq \phi y_{j,s}^{k_0}, j = 1, \dots, r$$

$$\lambda_s^k \geq 0, k = 1, \dots, n$$

(3)

$$[d_t^{k_0}(y_s, x_s)]^{-1} = \max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ubb.

$$\sum_{k=1}^n \lambda_t^k x_{i,t}^k \leq x_{i,s}^{k_0}, i = 1, \dots, m$$

$$\sum_{k=1}^n \lambda_t^k y_{j,t}^k \geq \phi y_{j,s}^{k_0}, j = 1, \dots, r$$

$$\lambda_t^k \geq 0, k = 1, \dots, n$$

(4)

$$[d_s^{k_0}(y_t, x_t)]^{-1} = \max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ubb.

$$\sum_{k=1}^n \lambda_s^k x_{i,s}^k \leq x_{i,t}^{k_0}, i = 1, \dots, m$$

$$\sum_{k=1}^n \lambda_s^k y_{j,s}^k \geq \phi y_{j,t}^{k_0}, j = 1, \dots, r$$

$$\lambda_s^k \geq 0, k = 1, \dots, n$$

Af ovenstående genkendes (1) og (2), som DEA programmer for hhv. periode t og basisperioden s . I (3) holdes observationerne i periode s op imod periode t 's teknologi og i (4), er det omvendt.

Det noteres, at i (3) og (4), hvor datapunkterne bliver holdt op imod teknologier fra en anden periode, behøver parameteren ϕ ikke være større end eller lig 1, som den altid er i (1) og (2). Datapunkter kan ligge over mulighedsområdet i både (3) og (4). Det er mest sandsynligt, at dette sker i (4), hvor datapunkter fra periode t bliver sammenlignet med teknologien i en tidligere periode (s). Hvis der er sket teknologisk fremskridt, da vil en værdi på $\phi < 1$ være mulig. Det skal dog lige noteres, at $\phi < 1$ også er mulig i (3), hvis der er sket en teknologisk tilbagegang.

Ud fra de lineære programmer kan Malmquist indekset opskrives:

$$MI^{k_0}(y_s, x_s, y_t, x_t) = \underbrace{\frac{(1) \frac{1}{\phi^{k_0}}}{(2) \frac{1}{\phi^{k_0}}}}_{\text{Catch-up}} \cdot \underbrace{\left[\frac{(4) \frac{1}{\phi^{k_0}}}{(1) \frac{1}{\phi^{k_0}}} \cdot \frac{(2) \frac{1}{\phi^{k_0}}}{(3) \frac{1}{\phi^{k_0}}} \right]}_{\text{Technical change}}^{1/2}$$

, hvor $(1) \frac{1}{\phi^{k_0}}$ angiver den inverse til løsningen af det første lineære program.

Man skal dog have sig for øje, er at de ϕ - og λ parametre, som findes i de 4 ovenstående lineære programmer sandsynligvis ikke er ens, da der løses 4 forskellige lineære programmer. De 4 programmer skal løses for hver observation i et datasæt, og hvis et datasæt består af 25 observationer og 2 tidsperioder, skal 100 lineære programmer løses. Hvis der lægges en ekstra tidsperiode til, skal der løses 3 ekstra lineære programmer for hver observation (det ene lineære program går igen). Dette kan skrives helt generelt, at hvis et datasæt består af n observationer og T tidsperioder, skal der løses $n \cdot (3T - 2)$ lineære programmer.

2.2.2. Eksempel på det outputorienterede Malmquist indeks

Jeg har igen valgt at forklare den gennemgåede teori vha. et to dimensionalt eksempel. Formålet er at give en bedre intuition af, hvordan indekset findes i praksis. Jeg genbruger eksemplet fra afsnit 2.1.3 med de 7 DMU'er, der hver med ét input producerer to output. Det antages igen, at de 7 beslutningsenheder producerer samme type output ved anvendelse af samme type input, jf. homogenitetskravet. I tabellen nedenfor er angivet input, output og output/input ratios for de 7 DMU'er i 2 forskellige perioder. Basisperiode 1 består af de samme data, som blev brugt i DEA eksemplet.

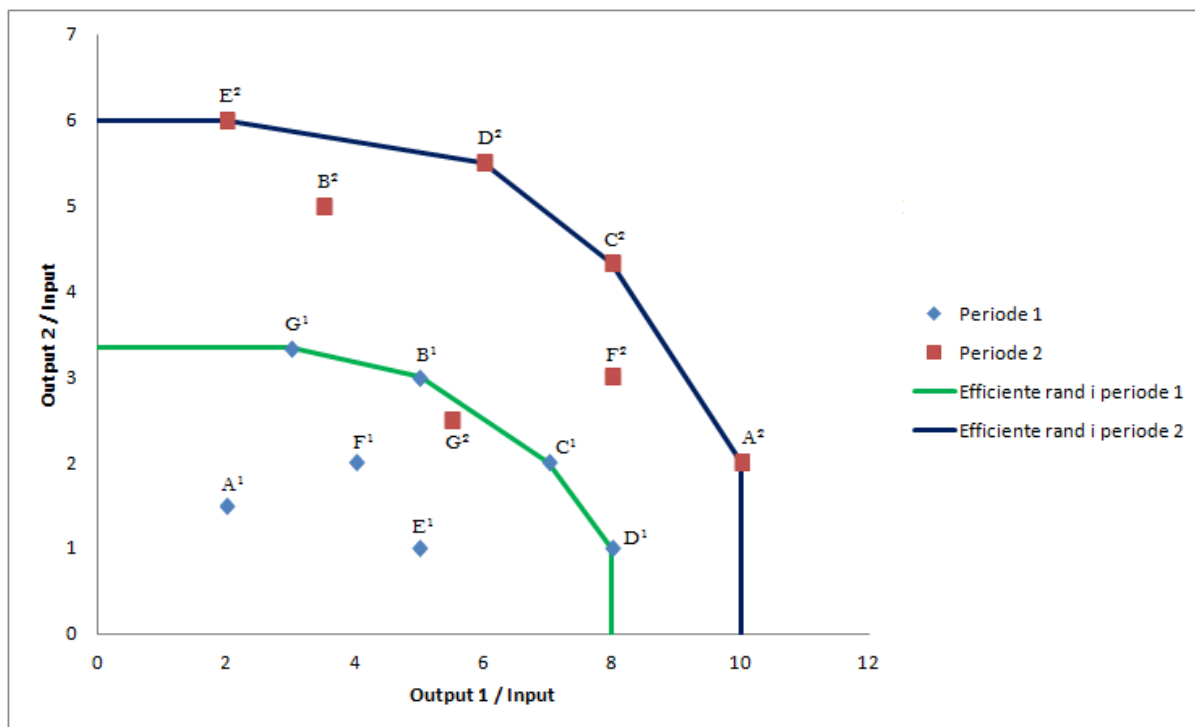
Tabel 2.3: Data for Malmquist eksempel

Periode	DMU	Input (x)	Output 1 (y_1)	Output 2 (y_2)	$\frac{\text{Output 1}}{\text{Input}} = \frac{y_1}{x}$	$\frac{\text{Output 2}}{\text{Input}} = \frac{y_2}{x}$
1	A	2	4	3	2	1,5
1	B	1	5	3	5	3
1	C	2	14	4	7	2
1	D	1	8	1	8	1
1	E	1	5	1	5	1
1	F	3	12	6	4	2
1	G	3	9	10	3	3,33
2	A	2	20	4	10	2
2	B	4	14	20	3,5	5
2	C	6	48	26	8	4,33
2	D	2	12	11	6	5,5
2	E	3	6	18	2	6
2	F	1	8	3	8	3
2	G	4	22	10	5,5	2,5

Kilde: Egen tilblivelse

Ud fra ovenstående datapunkter har jeg valgt at illustrere alle DMU'er fra hver periode i et koordinatsystem med $\frac{\text{Output 1}}{\text{Input}}$ og $\frac{\text{Output 2}}{\text{Input}}$, som hhv. vandret og lodret akse. Formålet med dette er at se hvilke beslutningsenheder, der ligger på den efficiente rand i de to perioder.

Figur 2.7: Den efficiente rand i periode 1 og 2



Kilde: Egen tilblivelse

På figuren er de røde punkter datapunkterne fra periode 2, de blå punkter er datapunkterne fra periode 1, de blå linjestykker udgør den efficiente rand i periode 2 og de grønne linjestykker udgør den efficiente rand i periode 1. Figuren viser, at hele den efficiente rand i periode 2 ligger over den efficiente rand i periode 1, hvilket betyder, at randeffekterne er positive.

I DEA eksemplet fandt jeg, at de efficiente DMU'er var *B*, *C*, *D* og *G*, hvilket ses på figuren. I periode 2 er det derimod DMU'erne *A*, *C*, *D* og *E*, som er efficiente, og dermed ligger på den efficiente rand. Enhederne *C* og *D* ligger på den efficiente rand i begge perioder, hvilket gør, at deres Catch-up indeks er lig 1. Figuren illustrer også, at Catch-up indekset for både *B* og *G* er mindre end 1, da de i periode 1 ligger på den efficiente rand, men ikke ligger der i periode 2. Det modsatte er tilfældet for *A* og *E*.

Nedenfor har jeg valgt at gå i dybden med DMU *B* og har derfor opskrevet de 4 lineære programmer for denne.

(1)

$$[d_2^B(y_2, x_2)]^{-1} = \max_{\phi^B, \lambda_t} \phi^B$$

ubb.

$$\lambda_2^A + \lambda_2^B + \lambda_2^C + \lambda_2^D + \lambda_2^E + \lambda_2^F + \lambda_2^G \leq 1$$

$$10\lambda_2^A + 3,5\lambda_2^B + 8\lambda_2^C + 6\lambda_2^D + 2\lambda_2^E + 8\lambda_2^F + 5,5\lambda_2^G \geq 3,5\phi^B$$

$$2\lambda_2^A + 5\lambda_2^B + 4,33\lambda_2^C + 5,5\lambda_2^D + 6\lambda_2^E + 3\lambda_2^F + 2,5\lambda_2^G \geq 5\phi^B$$

$$\lambda_2^k \geq 0, k = 1, \dots, n$$

(2)

$$[d_1^B(y_1, x_1)]^{-1} = \max_{\phi^B, \lambda} \phi^B$$

ubb.

$$\lambda_1^A + \lambda_1^B + \lambda_1^C + \lambda_1^D + \lambda_1^E + \lambda_1^F + \lambda_1^G \leq 1$$

$$2\lambda_1^A + 5\lambda_1^B + 7\lambda_1^C + 8\lambda_1^D + 5\lambda_1^E + 4\lambda_1^F + 3\lambda_1^G \geq 5\phi^B$$

$$1,5\lambda_1^A + 3\lambda_1^B + 2\lambda_1^C + \lambda_1^D + \lambda_1^E + 2\lambda_1^F + 3,33\lambda_1^G \geq 3\phi^B$$

$$\lambda_1^k \geq 0, k = 1, \dots, n$$

(3)

$$[d_2^B(y_1, x_1)]^{-1} = \max_{\phi^B, \lambda} \phi^B$$

ubb.

$$\lambda_2^A + \lambda_2^B + \lambda_2^C + \lambda_2^D + \lambda_2^E + \lambda_2^F + \lambda_2^G \leq 1$$

$$10\lambda_2^A + 3,5\lambda_2^B + 8\lambda_2^C + 6\lambda_2^D + 2\lambda_2^E + 8\lambda_2^F + 5,5\lambda_2^G \geq 5\phi^B$$

$$2\lambda_2^A + 5\lambda_2^B + 4,33\lambda_2^C + 5,5\lambda_2^D + 6\lambda_2^E + 3\lambda_2^F + 2,5\lambda_2^G \geq 3\phi^B$$

$$\lambda_2^k \geq 0, k = 1, \dots, n$$

(4)

$$[d_1^B(y_2, x_2)]^{-1} = \max_{\phi^B, \lambda} \phi^B$$

ubb.

$$\lambda_1^A + \lambda_1^B + \lambda_1^C + \lambda_1^D + \lambda_1^E + \lambda_1^F + \lambda_1^G \leq 1$$

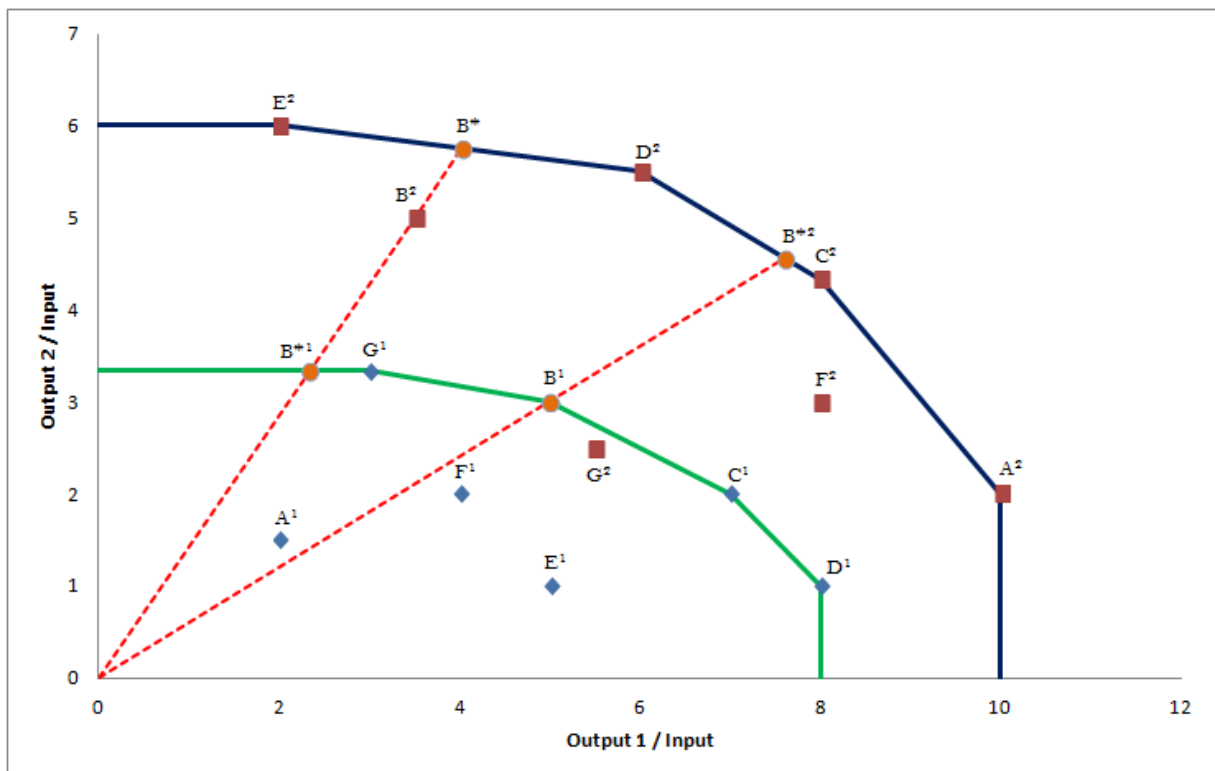
$$2\lambda_1^A + 5\lambda_1^B + 7\lambda_1^C + 8\lambda_1^D + 5\lambda_1^E + 4\lambda_1^F + 3\lambda_1^G \geq 3,5\phi^B$$

$$1,5\lambda_1^A + 3\lambda_1^B + 2\lambda_1^C + \lambda_1^D + \lambda_1^E + 2\lambda_1^F + 3,33\lambda_1^G \geq 5\phi^B$$

$$\lambda_1^k \geq 0, k = 1, \dots, n$$

For at få en ide om hvordan løsningen på de 4 lineære programmer ser ud, har jeg på nedenstående figur indtegnet hjælpelinjer ift. datapunkterne for B i periode 1 og 2.

Figur 2.8: Malmquist's produktivtetsindeks for DMU B



Kilde: Egen tilblivelse

På figuren er der angivet 3 forskellige referencepunkter for DMU B (B^* , B^{*1} og B^{*2}). For datapunktet B^1 (som jo er datapunktet for B i periode 1) er der "kun" angivet et referencepunkt, hvorimod der for B^2 er angivet 2. Grunden til dette er, at B i periode 1 ligger på den efficiente rand, hvilket i øvrigt betyder, at $\frac{1}{\phi^B}$ i (2) er lig 1 (som jeg i øvrigt også fik i DEA eksemplet). $\frac{1}{\phi^B}$ i (1) er ifølge grafen $\frac{\overline{OB^2}}{\overline{OB^*}}$, hvor B i periode 2 bliver holdt op i mod den efficiente rand i samme periode. I (3) bliver B i periode 1 sammenlignet med den efficiente rand i periode 2. Derfor er $\frac{1}{\phi^B}$ i (3) lig $\frac{\overline{OB^1}}{\overline{OB^{*2}}}$. I (4) er det omvendt ift. (3), for her bliver B i periode 2 holdt op imod den efficiente rand i periode 1, hvilket medfører, at $\frac{1}{\phi^B}$ i (4) er lig $\frac{\overline{OB^2}}{\overline{OB^{*1}}}$. Af grafen ses det, at $\frac{1}{\phi^B}$ i (3) er mindre end 1, og at $\frac{1}{\phi^B}$ i (4) er større end 1, dvs. der er sket et teknologisk fremskridt fra periode 1 til periode 2.

De 4 lineære problemer har jeg løst vha. Solver funktionen i Excel, og jeg får:

$$(1) \frac{1}{\phi^B} = 0,87, (2) \frac{1}{\phi^B} = 1, (3) \frac{1}{\phi^B} = 0,6574, (4) \frac{1}{\phi^B} = 1,5$$

Ud fra disse værdier kan Catch-up indekset, Technical change indekset og dermed også Malmquist indekset findes.

Catch-up indekset:

$$\text{Catch-up indekset for } B = \frac{(1) \frac{1}{\phi^B}}{(2) \frac{1}{\phi^B}} = \frac{0,87}{1} = 0,87$$

Technical change indekset:

$$\text{Technical change indekset for } B = \left[\frac{(4) \frac{1}{\phi^B}}{(1) \frac{1}{\phi^B}} \cdot \frac{(2) \frac{1}{\phi^B}}{(3) \frac{1}{\phi^B}} \right]^{1/2} = \left[\frac{1,5}{0,87} \cdot \frac{1}{0,6574} \right]^{1/2} = 1,6195$$

Malmquist indekset:

$$MI^B(y_1, x_1, y_2, x_2) = \underbrace{0,87}_{\text{Catch-up}} \cdot \underbrace{1,6195}_{\text{Technical change}} = 1,4089$$

Malmquist indekset reflekterer produktivitetsændringen for DMU *B* mellem periode 1 og 2.

Indekset angiver, at DMU *B* har haft en produktivitetsforbedring på ca. 41 % i den forstand, at dens driftsindtægter vil være 41 % højere i periode 1 sammenlignet med periode 2, hvis der kontrolleres for dens inputniveau. Det ses ydermere, at produktivitetsforbedringen udelukkende kommer fra randeffekten, da efficiensændringen er mindre end 1.

Nedenfor har jeg listet resultaterne for de 7 DMU'er, hvor jeg har løst $7 \cdot (3 \cdot 2 - 2) = 28$ lineære programmer vha. Solveren.

Tabel 2.4: Resultater for Malmquist eksempel

DMU	$\frac{1}{\phi}$ (1)	$\frac{1}{\phi}$ (2)	$\frac{1}{\phi}$ (3)	$\frac{1}{\phi}$ (4)	<i>Catch – up</i>	<i>Technical change</i>	<i>Malmquist indeks</i>
<i>A</i>	1	0,4783	0,2963	1,3333	2,0909	1,4670	3,0674
<i>B</i>	0,87	1	0,6574	1,5	0,87	1,6195	1,4089
<i>C</i>	1	1	0,7439	1,5152	1	1,4272	1,4272
<i>D</i>	1	1	0,8	1,6957	1	1,4559	1,4559
<i>E</i>	1	0,6667	0,5	1,8	1,5	1,5492	2,3238
<i>F</i>	0,9024	0,7273	0,4878	1,2727	1,2409	1,4501	1,7993
<i>G</i>	0,6524	1	0,5933	0,9545	0,6524	1,5703	1,0245

Kilde: Egne beregninger

Den værdi, som falder mest i øjnene, er helt klart $\frac{1}{\phi^B}$ i (4) for DMU *G*, der som den eneste er mindre end 1. Dette er tilfældet, da *G* i periode 2 ligger under den efficiente rand i periode 1. Yderligere ses det, at både Technical change indekset og Malmquist indekset for alle DMU'er er større end 1. Dette betyder, at der for alle DMU'er er tale om en teknisk- og produktivitetsforbedring, når man bevæger sig fra periode 1 til 2.

3. Databeskrivelse

Dette kapitel er opdelt i to sektioner. Første del omhandler mine valg og overvejelser ift. udvælgelsen af data herunder hvilke variable, jeg har valgt at tage med i analysen, samt hvilke justeringer jeg har foretaget for at få datasættet gjort klart til videre bearbejdning. Dernæst følger en statistisk beskrivelse af datasættet med formålet om at give et overblik over, hvordan værdierne i datasættet ser ud.

3.1. Datasættet

I dette afsnit er mine overvejelser og valg i forbindelse med udvælgelsen af data gennemgået. Afsnittet består af 7 sektioner. I første sektion beskriver jeg den database, hvor jeg har hentet data. Dernæst følger sektioner om hhv. hvilken tidsperiode og hvilke lande, som analysen dækker over. Yderligere kommer der en passage om, hvorfor jeg har valgt at få data på regionsbasis. Endvidere følger en beskrivelse af, hvilke variable jeg har taget med i analysen og et afsnit om hvilke justeringer, der har skullet til for at gøre data helt klar til analysedelen. Afsnittet afsluttes med en sektion om antallet af observationer i datasættet og om, hvordan jeg har inddelt regionerne i forskellige grupper.

3.1.1. Databasen

Den foreliggende undersøgelse er baseret på data udelukkende trukket fra Farm Accountancy Data Network (FADN) databasen på Europa Kommissionens hjemmeside. FADNs regnskabsmæssige netværk og dens underliggende datasystemer sigter mod at konstruere sammenlignelige, repræsentative og pålidelige regnskabsdata fra EU's landbrugssektor. FADNs primære formål er at evaluere indkomsten fra landbrugsbedrifter og vurdere effekterne af den fælles landbrugspolitik i EU¹² (CAP¹³). Data fra FADN er tidligere blevet brugt til bl.a. analyse af justeringer i prisstøtten, men bruges også af akademikere og politiske beslutningstagere til f.eks. at gennemføre forskellige former for landbrugs "performance" undersøgelser (som i denne afhandling). Kvaliteten af data er ikke analyseret i denne afhandling, hvilket vil sige, at jeg antager at FADNs og deres forbindelsers procedurer for indhentning og behandling af data sikrer en god datakvalitet.

¹² Yderligere oplysninger om FADN data kan findes på hjemmesiden: <http://ec.europa.eu/agriculture/rca/index.cfm>

¹³ CAP er en forkortelse for den fælles landbrugspolitik inden for EU-samarbejdet: Common Agricultural Policy.

I databasen findes data omhandlende generel landbrugsregnskabsstatistik for alle medlemslandene i EU, dvs. data er tilgængeligt for 27 lande i Europa.

3.1.2. Tidsperiode

Oprindeligt havde jeg planlagt at studere perioden fra 1989 til 2009, da det netop er for denne periode, der findes data i databasen. Jeg har dog valgt at basere analysen på landbrugsresultaterne for perioden fra 1995 til 2009, hvilket er gjort ud fra nedenstående argumentation.

- I perioden fra 1989-1994 er der ukomplette serier for regionen Canarias i Spanien.
- Sverige, Finland og Østrig blev først optaget i EU i januar 1995.
- For den østlige del af Tyskland (Brandenburg, Mecklenburg-Vorpommern, Sachsen, Sachsen-Anhalt og Thuringen) er data først tilgængeligt fra 1995¹⁴.

Dvs., at jeg bruger data fra 1995 i stedet for 1989, giver et mere harmonisk datasæt, hvor der ikke er for mange usikkerheder og manglende datapunkter.

3.1.3. Landene

Som tidligere nævnt indeholder FADN databasen landbrugsresultater for alle medlemslandene i EU, og jeg har valgt at inkludere alle disse i analysen. I nedenstående tabel har jeg specificeret hvilke lande, det drejer sig om, samt hvilket år de blev optaget i EU. Ydermere har jeg præciseret, hvilken del af EU's udvikling hvert medlemsland tilhører.

¹⁴ Hvorimod det tidligere Vesttyskland allerede var medlem af EF (nuværende EU), så blev det tidligere Østtyskland først d. 3. oktober 1990 en del af EF. Dette omtales i øvrigt som den "glemte udvidelse", fordi der reelt var tale om en udvidelse af EF med omkring 17 millioner nye EF-borgere. Denne "glemte udvidelse" og den manglende tekniske udvikling i Østtyskland er grunden til, at der først er data fra 1995 i denne del af Tyskland.

Tabel 3.1: Medlemslande i EU inklusiv deres optagelses år

EU's udvikling	Lande
De gamle medlemslande* (15 lande)	Belgien(1952), Danmark(1973), Finland(1995), Frankrig(1952), Grækenland(1981), Irland(1973), Italien(1952), Luxembourg(1952), Holland(1952), Portugal(1986), Spanien(1986), Storbritannien(1973), Sverige(1995), Tyskland(1952) ¹⁵ , Østrig(1995)
Den store udvidelse mod øst** (10 lande)	Cypern, Estland, Letland, Litauen, Malta, Polen, Slovakiet, Slovenien, Tjekkiet, Ungarn
Lille udvidelse*** (2 lande)	Bulgarien, Rumænien

Kilde: <http://www.euo.dk/>

*Lande som var medlem af EU per 1. januar 1995

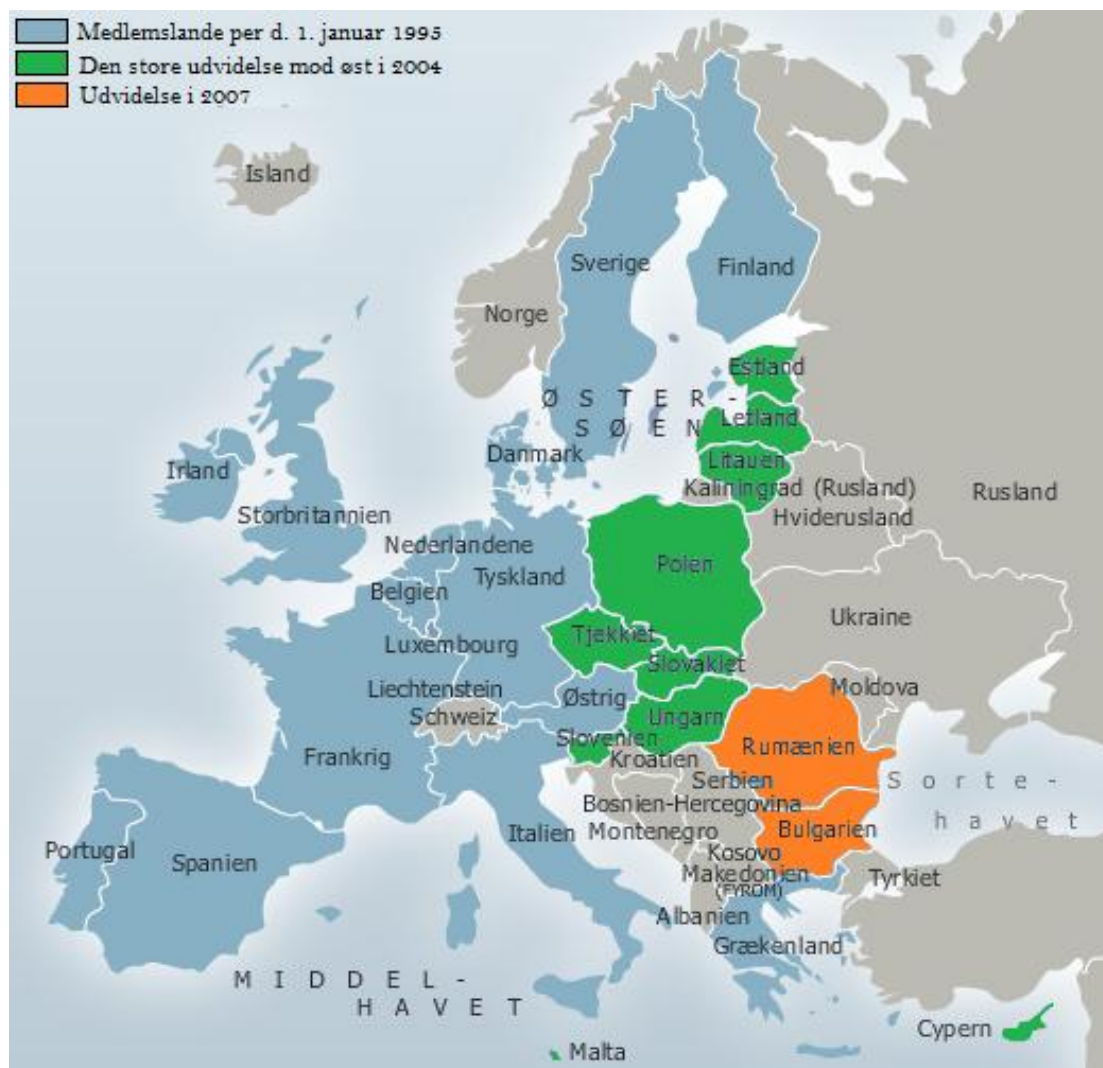
**Lande som blev medlem af EU i maj 2004

***Lande som blev medlem af EU i januar 2007

Denne udvikling af EU har jeg illustreret på nedenstående figur.

¹⁵ Der refereres her til det gamle Vesttyskland.

Figur 3.1: Kort over udviklingen i EU



Kilde: http://www.eu-oplysningen.dk/flash_docs/europakort/ og egen tilblivelse

På figuren angiver de lyseblå lande, landene som var medlem af EU pr. 1. januar 1995. De grønne lande er landene, som blev medlem af EU i forbindelse med den store udvidelse imod øst i 2004, og de orange lande repræsenterer udvidelsen i 2007.

3.1.4. Inddeling i regioner

Jeg har valgt at udtrække data på regionsbasis i stedet for på landebasis, da databasen kun indeholder 27 lande (for få observationer), samt at data ikke er tilgængeligt for alle observationerne over hele den tidsperiode, som jeg ønsker at analysere. Denne inddeling giver følgende fordele:

- Flere observationer giver større fleksibilitet, hvilket skyldes, at jeg undgår en situation, hvor for få observationer og mange input og/eller output kan medvirke til at mange enheder fremstår efficiente (ligger på den efficiente rand (frontier) under DEA).
- Størrelsesforskellene landene imellem kan udjævnes, da de større lande også inddeles i flere regioner end de små.
- Der kan være store forskelle internt i visse lande, som ved inddelingen elimineres.

For at kunne analysere udviklingstendenser i landbrugets produktion og produktivitet i EU over tid, har jeg, som tidligere skrevet, valgt at bruge Malmquists produktivitetsindeks. Denne metode kræver et balanceret panel, dvs. der kræves, at inddelingen af regionerne ikke ændrer sig over tid. Derfor er jeg blevet nødt til at slå et par regioner sammen, hvilket vil blive uddybet i afsnit 3.1.6.

3.1.5. Variable

Der findes i alt 148 forskellige variable i FADN databasen, men med inspiration fra undersøgelsen af landbrug over hele verden af Rao og Coelli (2003) og en analyse af landbruget i Europa af Rungsuriyawiboon og Lissitsa (2010), har jeg udvalgt 8 variable (6 input og 2 output). Variablene er valgt, så input indeholder ressourcerne, som bruges til at producere, og output afspejler, det en region samlet set producerer. De 8 variable er i databasen angivet som et gennemsnit pr. år pr. bedrift (landbrug) inden for regionerne. Nedenfor har jeg listet variablene sammen med en lille forklaring om, hvad de beskriver, og hvilket nummer de har i FADN databasen.

Input:

- *Økonomisk størrelse* (angivet i ESU¹⁶) (SE005)

Den økonomiske størrelse af en regions landbrug angivet i ESU.

- *Arbejdskraft* (angivet i timer) (SE011)

Arbejdsstyrken målt i timer. Variablen indeholder alle erhvervsaktive, som arbejder i landbruget (inkl. ejer(ne) selv).

- *Land* (angivet i ha) (SE025)

¹⁶ ESU = European size unit, er en standard bruttoavance på 1200 €, der bruges til at udtrykke den økonomiske størrelse af et landbrug.

Det udnyttede landbrugsareal. Variablen omfatter ikke arealer, der anvendes til svampe, jord der lejes for mindre end et år på lejlighedsvis basis, skovarealer og andre gård områder (veje, damme, osv.). Variablen indeholder jord i ejer besættelse, lejet jord og jord brug til deling af afgrøder. Den omfatter også landbrugsjord, som midlertidigt ikke er under dyrkning af landbrugsmæssige grunde eller jord, der er blevet fjernet fra produktionen som en del af landbrugspolitiske foranstaltninger. Land udtrykkes i hektar (10.000 m²).

- *Dyr* (angivet i dyreenhed)

Antallet af kvæg, får, geder, svin og fjerkræ til stede på landbruget (års gennemsnit). Variablen indeholder ikke bistader og andre dyr. Dyr der ikke tilhører indehaver af landbruget, men som alligevel er på landbruget tages i betragtning i henhold til deres årlige tilstedeværelse.

Variablen er beregnet på baggrund af følgende variable:

Malkekøer (angivet i stk.) (SE085)

Andet kvæg (angivet i stk.) (SE090)

Får og geder (angivet i stk.) (SE095)

Svin (angivet i stk.) (SE100)

Fjerkræ (angivet i stk.) (SE105)

Da hvert enkelt dyr ikke tæller ens (f.eks. tæller en høne ikke for lige så meget som en ko), har jeg brug for at omregne hver enkelt af ovenstående variable til en fast dyreenhed (animal units), hvilket man i litteraturen ikke er enige om. Jeg har valgt at bruge omregningsfaktoren fra Hayami og Ruttan (1970). Variablen Dyr bliver beregnet på baggrund af følgende formel:

$$\text{Dyr} = 0,8 \cdot \text{Malkekøer} + 0,8 \cdot \text{Andet kvæg} + 0,1 \cdot \text{Får og geder} + 0,2 \cdot \text{Svin} + 0,01 \cdot \text{Fjerkræ}$$

Enheden for variablen Dyr er dyreenheder.

- *Landbrugs specifikke omkostninger* (angivet i €) (SE281)

Variablen indeholder omkostning til frø, planter, gødningstoffer, beskyttelse af afgrøder (plantebeskyttelsesmidler, fælder, lokkemad, fugleskræmsler, frostbeskyttelse) og foder.

- *Andre omkostninger* (angivet i €) (SE336)

Denne variabel dækker omkostning til maskiner og bygninger (vedligeholdelse af udstyr, køb af mindre udstyr, bil udgifter, vedligeholdelse af bygninger, forbedring af jorden, samt forsikring af bygninger), samt omkostning til energi (brændstof, smøremidler og elektricitet).

Output:

- *Udbytte af afgrøder* (angivet i €) (SE135)

Variablen dækker udbytte af afgrøder og planteprodukter, samlet produktion af planter og vegetabiliske produkter = salg + bedriften + Eget forbrug + (ultimo værdiansættelse - primo værdiansættelse).

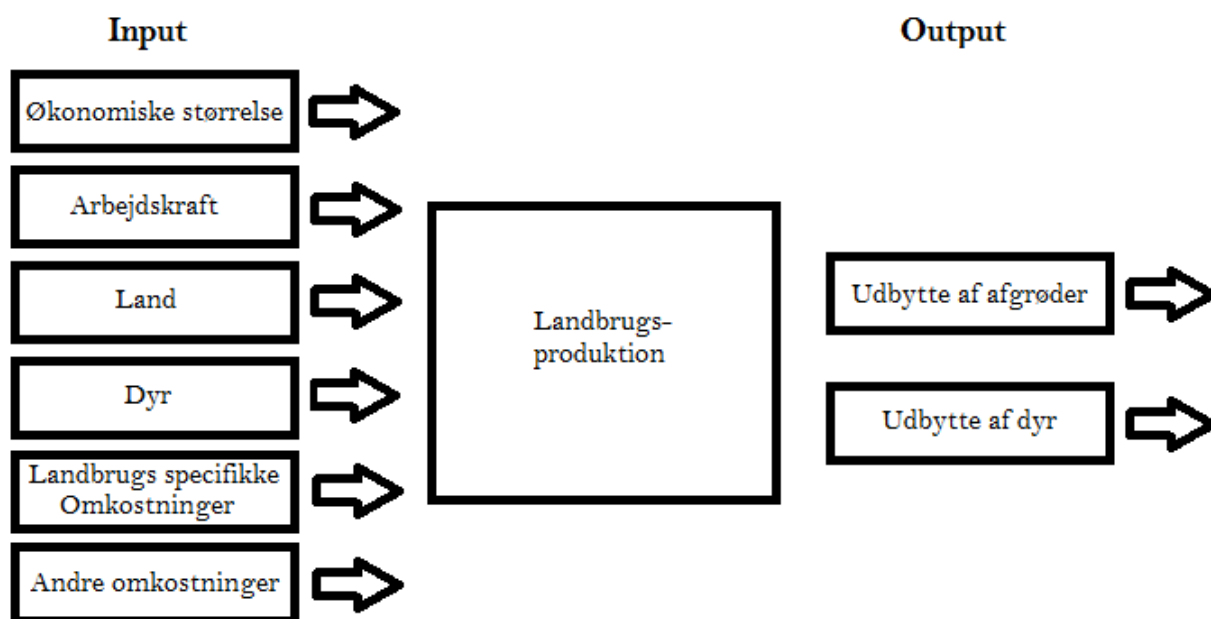
- *Udbytte af dyr* (angivet i €) (SE206)

Variablen indeholder udbytte af mælk og mælkeprodukter, udbytte af kvæg- og kalveproduktion, udbytte af produktionen af svinekød, udbytte af produktionen af får og geder, udbytte af produktionen af fjerkræ, samt udbytte af produktionen af æg.

Jeg er udmærket klar over, at Økonomisk størrelse er en lidt atypisk variabel at bruge i en performance analyse. Variablen har dog ifølge min optik sin berettigelse i analysen, da den indeholder et par ting, som modellen ellers ikke ville have fanget. Investeringer i bygninger og teknologi er eksempler på to elementer, som er indeholdt i Økonomisk størrelse, men som ikke er tilstede i nogle af de andre variable. Specielt investeringer i teknologi er et yderst vigtigt element, som både på den korte og lange bane kan skabe effektivisering og derigennem højne produktiviteten i landbruget.

Modellen som behandles i denne undersøgelse omfatter ovenstående input- og outputvariable, hvilket er illustreret på følgende figur.

Figur 3.2: Variable i modellen



Kilde: Egen tilblivelse

3.1.6. Fra FADN databasen til brugbart datasæt

For at få landbrugsdata fra FADN databasen gjort til et brugbart datasæt, har jeg foretaget et par justeringer. Disse er listet på nedenstående punktform.

1. Regionssammenlægninger:

Jeg har været nødsaget til at sammenlægge nogle regioner pga. ønsket om et balanceret panel, som er en nødvendighed for Malmquist indekset.

- *Vlaanderen og Wallonie* i Belgien er slået sammen til en region kaldet *Belgium*, så Belgien i datasættet består af præcis én region. Dette blev gjort for perioden 2004-2009.
- *Entre Douro e Minho/Beira litoral og Tras-os-Montes/Beira interior* i Portugal er slået sammen til en region kaldet *Norte e Centro*. Dette blev gjort for perioden 1995-2007.

Sammenlægningerne er lavet som et vægtet gennemsnit ift. antallet af landbrug, hvilket er den metode, som generelt anvendes i FADN databasen.

2. Manglende datapunkter:

For regionerne *Baleares* (Spanien) og *Hamburg* (Tyskland) var der ukomplette serier.

Baleares (Spanien):

Alle datapunkterne i 1998 og 1999 fandtes ikke i databasen, hvilket gjorde det nødvendigt at indsætte disse rækker i datasættet. De manglende datapunkter har jeg estimeret ved at lave en lineær fremskrivning mellem 1997 og 2000.

Hamburg (Tyskland):

I denne region manglede datapunkterne i 1995, og på samme måde estimerede jeg dem vha. en lineær tilbageskrivning mellem 1996 og 1997. Ekstrapolationen viste sig derimod at have en stor påvirkning på de øvrige regioner, da Hamburg i DEA-analysen blev benchmark for over 40 % af de andre regioner (i 1995). Derfor besluttede jeg at fjerne regionen Hamburg fra det samlede datasæt.

Mht. de manglende datapunkter så er det et klassisk problem, hvor den "reneste" løsning er at droppe de regioner, hvor der er ukomplette serier – men det er på den anden side også den fremgangsmåde, hvor man mister mest data. Derfor har jeg valgt at beholde regionen Baleares i datasættet. De manglende datapunkter i 1998 og 1999, har jeg estimeret ved at lave en lineær fremskrivning, da ingen af input- og outputvariablene følger en specifik fordeling, og at denne fremskrivning er rimelig simpel. Metoden kræver dog nogle stærke antagelser, om at alle de 8 variable følger en lineær sammenhæng mellem datapunkterne for hhv. 1997 og 2000. Det går dog imod den filosofi, som er grundlæggende i DEA, hvor der ikke kræves antagelse om, hvordan variablene "opfører" sig. Jeg er bevidst om, at dette er en problematisk antagelse, men jeg synes, at det er en mindre "ond" metode end helt at fjerne regionen fra datamængden.

3. Variablen Dyr:

Inputvariablen Dyr blev lavet ud fra omregningsfaktoren, som er angivet i afsnit 3.1.5.

3.1.7. Antal observationer og inddeling i grupper

I databasen er data tilgængeligt fra det år en region/land blev optaget i EU. Det betyder, at der f.eks. for regionen Slovakiet (Slovakiet er i datasættet kun en region) kun er data fra perioden 2004-2009. Denne "problematik" medfører, at datasættet består af et forskelligt antal observationer pr. år, hvilket er noteret i nedenstående tabel.

Tabel 3.2: Antal observationer i datasættet per år

Periode	Antal observationer
1995-2003	100
2004-2006	119
2007-2009	133

Kilde: Egen tilblivelse

Tabellen viser, at antallet af observationer selvfølgelig er stigende i takt med udvidelsen af EU. Med udvidelserne i hhv. 2004 og 2007 bliver datasættet udvidet med hhv. 19 og 14 observationer, så datasættet efter udvidelsen i 2007 består af 133 observationer pr. år. Samlet set består hele datasættet af $100 \cdot 9 + 119 \cdot 3 + 133 \cdot 3 = 1656$ observationer fordelt over de 15 år, som perioden strækker sig over.

Som tidligere skrevet er hvert land inddelt i et forskelligt antal regioner. Denne inddeling er, sammen med hvilken gruppe de tilhører, listet nedenfor.

Tabel 3.3: Antal regioner/observationer i hvert land

Land	Gruppe	Antal observationer
Belgien	EU15	1
Danmark	EU15	1
Finland	EU15	4
Frankrig	EU15	22
Grækenland	EU15	4
Irland	EU15	1
Italien	EU15	21
Luxembourg	EU15	1
Holland	EU15	1
Portugal	EU15	4
Spanien	EU15	17
Storbritannien	EU15	6
Sverige	EU15	3
Tyskland	EU15	13
Østrig	EU15	1
Cypern	EU10	1
Estland	EU10	1
Letland	EU10	1
Litauen	EU10	1
Malta	EU10	1
Polen	EU10	4
Slovakiet	EU10	1
Slovenien	EU10	1
Tjekkiet	EU10	1
Ungarn	EU10	7
Bulgarien	EU2	6
Rumænien	EU2	8

Kilde: Egen tilblivelse

I tabellen ses, at de største europæiske lande også er de lande, som er opdelt i flest regioner. F.eks. er den europæiske stormagt Tyskland opdelt i 13 regioner, hvorimod det lille ørige Malta kun er opdelt i 1. Jeg har valgt at inddele landene og dermed også regionerne i 3 grupper, alt efter hvornår hvert land blev medlem af EU. Den første gruppe navngives EU15 og indeholder de lande som etablerede EU, samt de lande som tilsluttede sig EU senest 1. januar 1995. Navnet på den anden

gruppe er EU10 og indeholder de 10 lande, som blev medlem af EU i maj 2004. Den sidste gruppe bærer navnet EU2 og består af Bulgarien og Rumænien, som begge blev medlem af EU i januar 2007. Ovenstående klassifikation er lavet med henblik på at se, om der er sammenfald mellem de 2 udvidelser i hhv. 2004 og 2007 og landbrugsproduktiviteten i de gamle medlemslande i den Europæiske Union (EU15-landene).

3.2. Statistisk beskrivelse af data

I denne sektion har jeg beskrevet data vha. deskriptiv statistik. I afsnittet vil jeg fremføre tabeller og figurer, som illustrerer, hvordan datasættet hænger sammen.

Nedenfor har jeg præsenteret maksimumsværdien, minimumsværdien, middelværdien og standardafvigelsen (spredningen) for hver enkelt variabel taget over hele datasættet for at give et overblik over, hvordan værdierne for de forskellige variable ser ud.

Tabel 3.4: Beskrivende statistik for input- og outputvariablene for hele datasættet

Variable	Økonomisk størrelse (ESU)	Arbejdskraft (timer)	Land (ha)	Dyr (dyreenhed)	Landbrugs specifikke omkostninger (€)	Andre omkostninger (€)	Udbytte af afgrøder (€)	Udbytte af dyr (€)
Maks	455,00	37.646,09	614,61	216,27	355.253,00	322.581,00	486.517,00	342.819,00
Min	2,20	1.359,21	2,00	0,00	2.321,00	956,00	963,00	0,00
Middel	54,50	4.305,88	68,52	26,18	37.358,89	27.188,89	49.463,37	45.775,15
Std.afv.	62,39	3.595,85	94,26	30,54	47.005,24	39.121,58	59.805,67	53.689,52

Kilde: Egne beregninger fra data

I tabellen er maksimums- og minimumsværdierne angivet for de regioner, som i gennemsnit har den største og mindste værdi per bedrift per år. Middelværdierne og standardafvigelserne er angivet som et gennemsnit pr. år pr. bedrift. Den gennemsnitlige bedrift i datasættet har en økonomisk størrelse

på 54,40 ESU, bruger 4.305,88 timer i arbejdskraft svarende til lidt over 2 fuldtidsansatte per år¹⁷, har et landbrugsareal på 68,52 ha, har ca. 26 dyreenheder på bedriften, bruger knap 64.500 € på landbrugsspecifikke omkostninger og andre omkostninger, har en indtægt på 49.463,37 € på afgrøder og har en indtægt på 45.775,15 € på produktionen af dyr.

Variablene ligger værdimæssigt i forskellige niveauer, dog ligger de 4 variable, som er angivet i € (omkostningsvariablene Landbrugs specifikke omkostninger og Andre omkostninger, samt indtægtsvariablene Udbytte af afgrøder og Udbytte af dyr) alle i nogenlunde samme leje. På indtægtssiden (og dermed også outputsiden) er Udbytte af afgrøder i gennemsnit større end Udbytte af dyr, og på omkostningssiden (inputsiden) er Landbrugs specifikke omkostninger i gennemsnit større end Andre omkostninger (samme mønster ses for standardafvigelsen).

Tabellen giver i det hele taget et godt billede af, hvilket niveau hver variabel ligger i. Bl.a. ser jeg, at den region med det største gennemsnitlige udnyttede landbrugsareal per bedrift ligger på 614,61 ha, hvorimod den mindste har et gennemsnitligt udnyttet landbrugsareal per bedrift på 2,00 ha. Dette er en difference på 612,61 ha, hvilket betyder, at der er stor forskel på størrelsen af det gennemsnitlige landbrugsareal per bedrift alt afhængig af, hvilken region man kigger på. Selve denne forskel i størrelsen af den gennemsnitlige bedrift i en region ses også meget tydeligt i både den Økonomiske størrelse og Arbejdskraft. Variablen Arbejdskraft svinger fra 19,57 fuldtidsansatte til 0,71 fuldtidsansatte, hvilket signalerer en kæmpe forskel på det gennemsnitlige antal ansatte per bedrift i visse regioner kontra andre regioner.

Variablene Dyr og Udbytte af dyr har, som de eneste variable, minimumsværdier på 0, hvilket betyder, at der er regioner, som ikke opdrætter dyr på deres bedrifter, hvormed der selvfølgelig heller ikke optræder en indtægt i forbindelse med opdrætning af dyr (Udbytte af dyr). I datasættet er det kun regionen Canarias i Spanien, som har værdier lig 0 for disse variable, hvilket er gældende i perioden 1995-1999. I de resterende år i datasættet har Canarias positive værdier for begge variable, da de i regionen er begyndt at opdrætte dyr, hvilket medfører, at der heraf naturligvis genereres en indtægt. Skiftet fra værdier lig 0 til positive værdier for Canarias skaber under Malmquist-analysen problemer, hvilket jeg vil komme nærmere ind på i afsnit 4.3.

¹⁷ Jeg har taget udgangspunkt i, at en fuldtidsansat arbejder 37 timer om ugen i 52 uger om året, hvilket svarer til 1924 timer om året.

Nedenfor har jeg opdelt datasættet i 3 tidsperioder for at se, hvordan middelværdierne og standardafvigelserne ændrer sig for EU15-gruppen i forbindelse med de 2 udvidelser af EU. Første tidsperiode dækker over perioden før 1. udvidelse (1995-2003), den næste over perioden fra 1. udvidelse til lige før 2. udvidelse indtræffer (2004-2006) og den sidste over perioden fra 2. udvidelse forekommer til datasættet slutter (2007-2009).

Tabel 3.5: Middelværdi og standardafvigelse for gruppen EU15 angivet pr. år pr. bedrift for 3 tidsperioder

Variable	Økonomisk størrelse (ESU)	Arbejdskraft (timer)	Land (ha)	Dyr (dyreenhed)	Landbrugs specifikke omkostninger (€)	Andre omkostninger (€)	Udbytte af afgrøder (€)	Udbytte af dyr (€)
Middel (1995-2003)	52,47	4189,71	65,05	26,14	32.736,06	23.728,65	43.780,61	42.964,05
Std.afv. (1995-2003)	61,13	3.276,25	90,24	30,42	39.444,98	33.927,18	50.862,89	49.486,47
Middel (2004-2006)	64,92	4036,46	72,74	29,01	40.995,25	32.645,14	54.969,93	52.066,25
Std.afv. (2004-2006)	68,13	3.053,92	94,56	31,35	46.866,88	44.552,20	59.901,11	56.827,99
Middel (2007-2009)	66,77	4.056,42	75,84	30,39	51.718,47	36.681,17	68.964,12	58.913,36
Std.afv. (2007-2009)	66,49	3.166,82	97,19	32,30	63.237,53	48.365,85	81.480,28	65.983,17

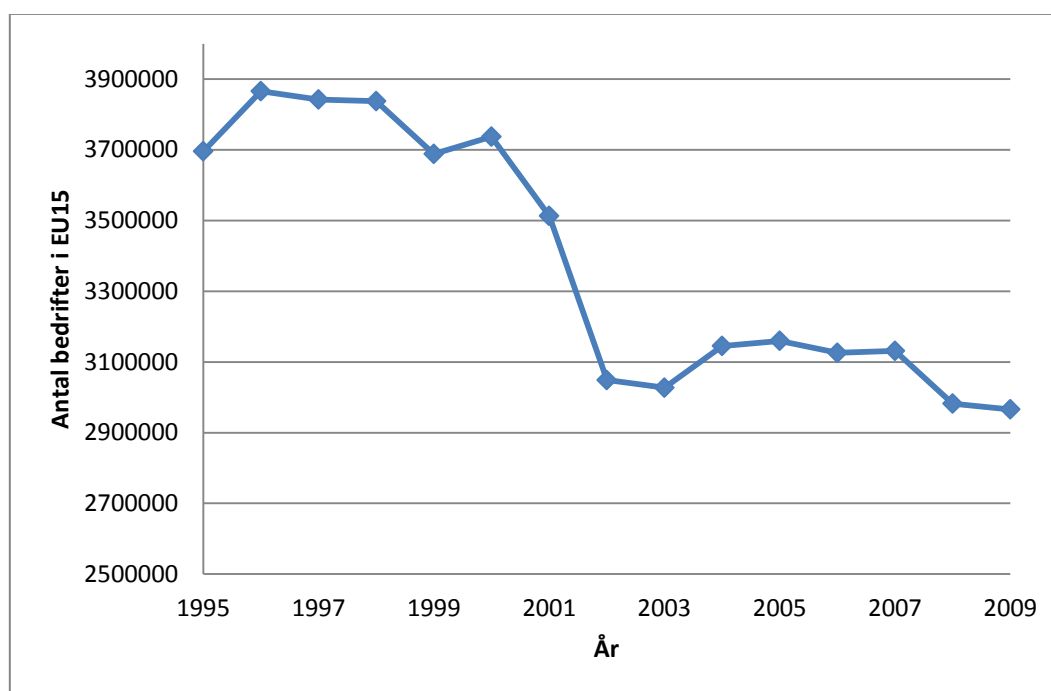
Kilde: Egne beregninger fra data

I tabellen ses en klar tendens. Næsten samtlige middelværdier og standardafvigelser bliver større og større i takt med, at vi nærmer os nutiden. Specielt for omkostnings- og indtægts parametrene er tendensen tydelig. Parametrene bliver i gennemsnit imellem 13.000 € og 25.000 € større gående fra 1995-2003 til 2007-2009. De eneste variable som afviger ift. denne tendens er Arbejdskraft og Økonomisk størrelse. For den Økonomiske størrelse er det i springet fra tidsperioden 2004-2006 til

2007-2009, hvor der for standardafvigelsen registreres et fald. Det betyder, at spredningen er blevet mindre, da værdierne for den Økonomiske størrelse er kommet tættere på gennemsnittet. For variabelen Arbejdskraft observerer jeg et fald i både middelværdi og spredning fra tidsperioden 1995-2003 til 2004-2006. Variablen Arbejdskraft er faktisk den eneste variabel, som ikke registrerer en stigning i både middelværdi og spredning fra tidsperioden 1995-2003 til 2007-2009. Her optræder der nemlig for både middelværdi og spredning et fald. En mulig forklaring på at arbejdskraften falder, kunne være at EU15-landene i større grad vælger at substituere fra arbejdskraft til maskiner.

I ovenstående tabel ses en klar trend, der tyder på at landbrugene fra 1995 til 2009 i gennemsnit bliver større og større og producerer mere og mere. Denne udvikling har jeg valgt at se nærmere på ved at kigge på udviklingen i antallet af bedrifter i EU15. Denne udvikling er angivet på nedenstående figur.

Figur 3.3: Udviklingen i antal bedrifter per år i EU15-landene



Kilde: http://ec.europa.eu/agriculture/rca/database/database_en.cfm og egen tilblivelse

Figuren viser, at antallet af bedrifter i EU15-landene er klart aftagende i den periode mine datapunkter strækker sig over. Antallet af bedrifter går fra omkring 3,7 millioner bedrifter i 1995 til

ca. 2,9 millioner bedrifter i 2009, hvilket er en reduktion på 0,8 millioner på knap 15 år. I perioden fra 1995 til 2000 er antallet af bedrifter ellers rimeligt stabilt placeret mellem 3,7 og 3,9 millioner, men i både 2001 og 2002 falder antallet af bedrifter drastisk. I perioden fra 2003 til 2009 er antallet af bedrifter igen stabilt dog anes en svag nedadgående tendens. Sammenholder jeg denne figur med middelværdierne og standardafvigelseerne for EU15, må det konkluderes, at der i EU15-landene i 2009 er færre landbrug ift. 1995, men at de i gennemsnit er blevet større og større.

4. Resultater

I dette kapitel vil jeg præsentere resultaterne af DEA- og Malmquist-analyserne. Alle resultater er fremkommet via beregninger i Excel ved overvejende brug af Solver funktionen og programmeringssproget VBA i henhold til den gennemgåede teori i kapitel 2.

Som tidligere skrevet består mit datasæt af 15 års data, hvor hvert år indeholder mellem 120 og 133 observationer. Mine beregninger involverer derfor løsninger af et stort antal lineære programmeringsproblemer, hvilket giver tusindvis af stykker af information om efficiensscorer og peers/benchmarks for de enkelte regioner hvert år. På baggrund af den store resultatomængde har jeg været yderst selektiv i, hvilke resultater jeg har valgt at præsentere i dette afsnit, samt hvordan disse præsenteres. For overskuelighedens skyld har jeg valgt at angive resultaterne på landebasis (pga. det store antal af regioner). Derudover skal det klarlægges, at når jeg i dette kapitel fremhæver resultater i form af tal eller lignende, der ikke umiddelbart fremgår af tabeller eller figurer, er det resultater, som er fundet i forbindelse med bearbejdning af datasættet.

Dette kapitel indledes med en præsentation af efficiensscorerne fundet ved DEA-analyse på hvert år for sig, hvorefter jeg ser på efficiensscorerne fra en DEA-analyse på hele datasættet på en gang. Kapitlet afrundes med et afsnit for mål af efficiensforandring (Catch-up indekset), randeffekt (Technical change indekset) og TFP ændring (Malmquist indekset) for alle regioner i EU15-gruppen i hvert par af tilstødende år. De lineære programmer tilhørende de tre analyser er opskrevet i hhv. bilag 1, 2 og 3.

4.1. DEA-analyse på hvert år for sig

Denne sektion indledes med en præsentation af resultaterne for efficiensscorerne, når jeg foretager en DEA-analyse på hvert enkelt år for sig. DEA resultaterne for en sådan analyse klarlægger, hvordan en region præsterer ift. de andre regioner i samme år. Resultaterne for ændringerne i efficiens fra år til år giver en ide om, hvordan Catch-up effekten bidrager til TFP væksten (Malmquist indekset), hvilket jeg kommer nærmere ind på senere i dette afsnit. Denne sektion afsluttes med en undersøgelse af, hvilke landes regioner der er benchmark i den nærværende analyse. Efficiensscorerne for de 27 lande, som er med i analysen, er for perioden 1995-2009 præsenteret nedenfor.

Tabel 4.1: Gennemsnitlige årlige efficiensscorer på landebasis, 1995-2009 (hvert år for sig)

Lande	Gruppe	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	Middel
(BEL) Belgium	EU15	1,00	1,00	1,00	1,00	1,00	1,00	0,97	0,87	0,93	0,96	0,94	0,91	0,86	0,87	0,87	0,95
(DAN) Denmark	EU15	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
(DEU) Germany	EU15	0,90	0,86	0,86	0,82	0,85	0,88	0,87	0,81	0,80	0,85	0,81	0,85	0,84	0,87	0,88	0,85
(ELL) Greece	EU15	0,96	0,90	0,89	0,82	0,88	0,87	0,87	0,89	0,91	0,89	0,94	0,87	0,85	0,92	0,86	0,89
(ESP) Spain	EU15	0,96	0,97	0,96	0,94	0,94	0,97	0,97	0,95	0,95	0,94	0,93	0,92	0,90	0,87	0,84	0,93
(FRA) France	EU15	0,96	0,95	0,92	0,88	0,90	0,94	0,93	0,95	0,95	0,93	0,92	0,93	0,92	0,94	0,93	0,93
(IRE) Ireland	EU15	0,94	1,00	1,00	0,96	0,92	1,00	1,00	0,87	0,92	1,00	0,95	0,97	0,94	0,88	0,74	0,94
(ITA) Italy	EU15	0,92	0,93	0,90	0,88	0,86	0,86	0,92	0,92	0,94	0,94	0,93	0,94	0,92	0,91	0,90	0,91
(LUX) Luxembourg	EU15	1,00	1,00	0,94	1,00	1,00	0,99	0,99	0,90	0,93	0,91	0,89	1,00	0,92	0,96	0,86	0,95
(NED) Netherlands	EU15	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
(OST) Austria	EU15	0,98	1,00	1,00	1,00	1,00	1,00	1,00	0,98	0,89	0,94	0,91	0,95	0,90	1,00	0,85	0,96
(POR) Portugal	EU15	0,70	0,66	0,66	0,62	0,71	0,70	0,71	0,75	0,74	0,77	0,75	0,83	0,68	0,82	0,78	0,73
(SUO) Finland	EU15	0,82	0,79	0,75	0,84	0,84	0,80	0,80	0,83	0,86	0,79	0,72	0,77	0,77	0,83	0,86	0,80
(SVE) Sweden	EU15	0,90	0,89	0,90	0,85	0,84	0,84	0,80	0,81	0,83	0,97	0,90	0,96	0,88	0,95	0,87	0,88
(UKI) United Kingdom	EU15	0,92	0,91	0,88	0,80	0,83	0,90	0,88	0,86	0,90	0,85	0,84	0,91	0,85	0,85	0,92	0,87
(CYP) Cyprus	EU10										1,00	0,96	0,98	0,93	1,00	1,00	0,98
(CZE) Czech Republic	EU10										0,82	0,83	0,84	0,59	0,79	0,88	0,79
(EST) Estonia	EU10										0,83	1,00	0,89	0,71	0,82	1,00	0,88
(HUN) Hungary	EU10										0,96	0,95	0,93	0,77	0,95	0,90	0,91
(LTU) Lithuania	EU10										1,00	1,00	0,75	0,57	0,98	0,94	0,87
(LVA) Latvia	EU10										0,80	0,86	0,81	0,63	0,72	0,83	0,77
(MLT) Malta	EU10										1,00	1,00	1,00	1,00	1,00	1,00	1,00
(POL) Poland	EU10										0,86	0,86	0,89	0,72	0,82	0,78	0,82
(SVK) Slovakia	EU10										0,72	0,76	0,86	0,60	0,91	0,88	0,79
(SVN) Slovenia	EU10										0,91	0,86	0,82	0,53	0,74	0,79	0,78
(BGR) Bulgaria	EU2													0,57	0,82	0,76	0,72
(ROU) Romania	EU2													0,77	0,90	0,92	0,86
Middel EU15		0,93	0,92	0,90	0,87	0,88	0,90	0,91	0,90	0,91	0,91	0,89	0,91	0,88	0,90	0,89	0,90
Middel EU10											0,91	0,91	0,90	0,73	0,89	0,88	0,87
Middel EU2														0,68	0,87	0,85	0,80
Middel		0,93	0,92	0,90	0,87	0,88	0,90	0,91	0,90	0,91	0,91	0,89	0,91	0,84	0,89	0,88	0,89

Kilde: Egne beregninger fra data

Efficiensscorerne i tabellen ovenfor er angivet som et gennemsnit over regionerne i et givent land eller gruppe. I søjlen længst til højre har jeg for hvert land angivet den gennemsnitlige værdi (middelværdien) for de årlige efficiensscorer, og i bunden af tabellen er der for hver gruppe, og for hele datasættet angivet middelværdierne for de årlige efficiensscorer, hvor der er taget højde for, hvor mange regioner hvert land består af. Middelværdierne nederst i tabellen er altså angivet som et

gennemsnit for en region hhv. tilhørende en given gruppe og for det samlede datasæt. Det skal dog fastslås, at det kan være problematisk direkte at sammenligne efficiensscorer på tværs af årene, da landbrugsresultaterne kan være skabt ud fra forskellige forudsætninger (f.eks. er vejret ikke ens hvert år). Det giver f.eks. ikke mening at sammenligne middelværdierne, som er angivet i søjlen længst til højre, fordi der i middelværdierne ikke tages højde for, hvordan landbrugsresultaterne for hvert år er ift. hinanden, dvs. om de er blevet dårligere eller bedre. Til gengæld giver middelværdierne for hver gruppe per år (angivet i bunden af tabellen) gode fortolkningsmuligheder i den forstand, at de giver et billede af, hvordan Catch-up effekten bevæger sig fra år til år.

Inden jeg går i dybden med efficiensscorerne i tabellen, vil jeg vende blikket imod regionen Baleares (i Spanien), der som tidligere skrevet har datapunkter fremstillet ved lineær interpolation (i 1998 og 1999). Regionen har efficiensscorer på 1 i både 1998 og 1999, hvilket medfører, at regionen ligger på den efficiente rand i både 1998 og 1999. I 1998 og 1999 er Baleares benchmark for hhv. 6 og 8 andre regioner, hvilket svarer til hhv. 6 % og 8 % af regionerne i hvert af de pågældende år. Jeg er udmærket klar over, at det faktum at den fiktive region Baleares både er efficient og benchmark for andre regioner (hvor datapunkterne er reelle landbrugsresultater) er en problematik. Jeg har på baggrund af denne problematik ikke valgt at fjerne regionen fra datasættet, da årene 1998 og 1999 følger den generelle tendens for Baleares. Med dette mener jeg, at det ikke er en uvant situation, at Baleares ligger på den efficiente rand. Faktisk er Baleares i gennemsnit benchmark 5,4 gange om året, hvor de 40 % af gangene sker i 2000. Det betyder, at regionen i 1998 og 1999 optræder som benchmark ca. lige så mange, som gennemsnittet foreskriver.

Efficiensscorerne i tabellen ovenfor spænder fra et gennemsnit på 0,72 for en region i Bulgarien til maksimum 1,00 med et gennemsnit på 0,89. Den gennemsnitlige efficiensscore indebærer, at de lande/regioner som er med i denne undersøgelse i gennemsnit producerer 89 % af det output, der potentielt kunne være fremstillet vha. de anvendte/aktuelle input mængder ift. det år regionen befinder sig i. Ifølge undersøgelsen udviser Danmark, Holland og Malta perfekte efficiensscorer. Danmark og Holland ligger på den efficiente rand i hele tidsperioden, hvorimod Malta er på den efficiente rand i perioden fra 2004 til 2009. Den perfekte efficiens gør, at jeg senere i dette kapitel med sikkerhed vil se en Catch-up effekt på 1 i hele perioden for både Danmark og Holland, når jeg beregner Malmquist indekset for landene i EU15. Catch-up indekset findes ved at tage forholdet mellem efficiensscorerne i to tilstødende år, og f.eks. er Catch-up indekset ca. 0,96 for Irland i

skridtet fra 1998 til 1999 ($\frac{\text{efficiens i år 1999}}{\text{efficiens i år 1998}} = \frac{0,92}{0,96} = 0,96$)¹⁸. Værdien på 0,96 (< 1) angiver, at Irland (består af én region) er blevet mindre efficient i 1999 ift. året inden (afstanden til den efficiente rand er blevet større).

I EU15-gruppen er det udover Danmark og Holland, Belgien, Irland, Luxemburg og Østrig, som i gennemsnit per år klarer sig bedst. Belgien befinder sig faktisk på den efficiente rand i perioden fra 1995 til 2000, men herefter er den generelle tendens, at Belgien præsterer dårligere og dårligere ift. de andre regioner (i samme år) i EU efterhånden, som årene går. Det største "skift" i efficiens for Belgien findes i 2002, hvor Belgien har en efficiens på 0,87 ift. 0,97 året før. Både i 2001 og 2002 er regionen Lombardia i Italien det bærende benchmark for Belgien (dvs. det benchmark der indgår med den største vægt (λ)), og regionen indgår med en andel på hhv. 1,00 og 0,95 af dens egne datapunkter. Forklaringen på den store ændring i efficiens for Belgien findes derfor ved at kigge på datapunkterne for Lombardia i 2001 og 2002. Generelt er (alle) inputvariablene i 2002 steget en smule ift. i 2001, men til gengæld er det samlede udbytte blevet 21.000 € større svarende til en stigning på ca. 21 %. Grunden til at Belgien har lavere efficiens i 2002 end i 2001, er altså at dens benchmark Lombardia er blevet markant stærkere i 2002.

Portugal og Finland er med gennemsnitlige efficiensscorer på hhv. 0,73 og 0,80 de to lande, som i gennemsnit klarer sig dårligst i EU15. Efficiensscorerne i Finland svinger meget op og ned i hele tidsperioden uden at antage en speciel tendens, hvilket gør det svært at sige noget om værdien af det gennemsnitlige Catch-up indeks for Finland. For Portugal anes der en svag opadgående tendens i efficiensscorerne, og disse ligger i første del af tidsperioden omkring 0,62 - 0,71, hvorimod de i anden del bevæger sig op omkring 0,74 - 0,83 (dog med en enkelt outlier i 2007). Udviklingen i efficiensscorerne for Portugal siger helt konkret, at Portugal har et positivt gennemsnitligt årligt Catch-up indeks, hvis jeg tager hele tidsperioden i betragtning.

I EU10-gruppen er det Cypern, Estland og Litauen, som udover Malta ligger på den efficiente rand i løbet af tidsperioden. Litauen har efficiensscorer på 1 i både 2004 og 2005. I årene lige efter (2006 og 2007) er der store fald i landets efficiens, inden der igen præsteres på et efficiensniveau over 0,90 i 2008 og 2009. Faldet i 2006 kan forklares ved at se på forskellen i datapunkterne for hhv. Litauen i 2005 (hvor Litauen er efficient) og dens benchmark i 2006. I 2006 er Madrid i Spanien,

¹⁸ Catch-up indekset for lande som består af mere end en region kan ikke findes på en helt tilsvarende måde. Disse findes ud fra et geometrisk gennemsnit af efficiens ændringerne for alle de regioner, som et land består af.

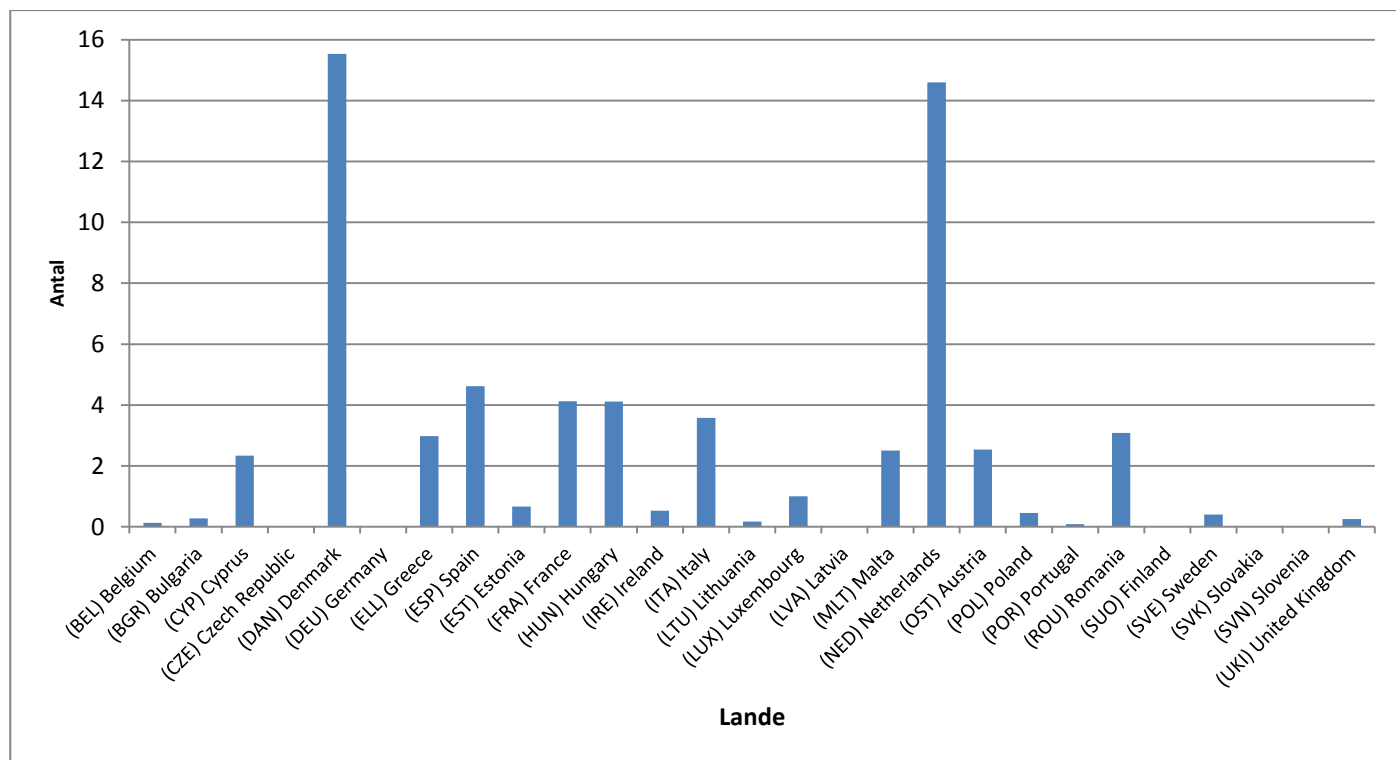
Corse i Frankrig, Valle d'Aoste i Italien og Malopolska and Pogórze i Polen benchmark for Litauen, og regionerne indgår med vægte på hhv. 6 %, 5 %, 30 % og 53 %. Det samlede benchmark for Litauen i 2006, som består af de 4 regioner, præsterer bedre end Litauen i 2005, hvilket medfører et fald i efficiens. I 2007 er efficiensscoren 0,57, som er en af de laveste værdier i tabellen. I 2007 indgår Valle d'Aoste i Italien ligesom i 2006 i referencemængden med en stor andel (vægt på 26 %). Selvom regionen ikke udgør den samme procentdel i de to år, kan jeg ved at se på datapunkterne for denne region få en ide om, hvordan Litauens store efficiensfald kan forklares. Inputvariablene er på samme niveau i de to år, men der ses en stigning på over 100.000 € for outputvariablen Udbytte af afgrøder, hvilket betyder, at det samlede udbytte stiger til over det tredobbelte. Denne markante stigning giver en forklaring på Litauens iøjefaldende efficiensnedgang i 2007.

Som afsluttende bemærkning til Tabel 4.1 skal det noteres, at middelværdien for EU15-gruppen ligger rimeligt konstant i intervallet mellem 0,87 og 0,93. Middelværdierne giver et indblik i, hvilke værdier Catch-up leddet i Malmquist-analysen antager. F.eks. kan jeg af tabellen se, at der i EU15-gruppen er efficienstilbagegang i 1996, 1997, 1998, 2002, 2004¹⁹, 2005, 2007 og 2009, samt fremgang i 1999, 2000, 2001, 2003, 2006 og 2008.

For at give et billede af hvilke landes regioner, som hyppigst optræder som benchmark for andre regioner, har jeg på nedenstående figur illustreret et søjlediagram, der viser hvor mange gange et lands regioner i gennemsnit er benchmark per år.

¹⁹ Der er i både 2003 og 2004 efficiens lig 0,91, og jeg har derfor været nødt til at se på de eksakte værdier for at bestemme, om der er tilbagegang eller fremgang.

Figur 4.1: Det gennemsnitlige antal gange et lands regioner er benchmark per år (hvert år for sig)



Kilde: Egne beregninger fra data

Det ses ikke overraskende, at Danmark og Holland er de lande, hvor en region i gennemsnit per år optræder flest antal gange som benchmark. Både Danmark og Holland består af netop én region, hvilket betyder, at Danmark og Holland i gennemsnit optræder hhv. 16 og 15 gange som benchmark for de inefficiente regioner per år. Danmark optræder i perioden 2004-2009 klart flest gange som benchmark, hvor der i 2006 toppes med 35 gange. Holland ligger i hele tidsperioden meget konstant omkring middelværdien på de 15 gange.

Det sidste land som udviste perfekt efficiens i hele dens tidsperiode er Malta. Malta optræder som benchmark 2 gange i 2004, 4 gange i 2007, 2 gange i 2008 og 7 gange i 2009, dvs. i alt 15 gange. Da Malta kun består af en region og er del af datasættet i 6 år, indgår Malta i gennemsnit som benchmark 2,5 gange ($\frac{15}{6} = 2,5$) om året. Det har dog intet med landets størrelse at gøre, at Malta slet ikke er i nærheden af i gennemsnit at være benchmark i så høj grad som både Danmark og Holland. For selvom Malta (pga. landets størrelse) har helt unikke niveauer af input og output,

bliver de større regioner under DEA-analysen (under antagelse af CRS) sammenlignet med Malta via nedskalerede versioner af dem selv.

Til sidst i dette afsnit skal det bemærkes, at regioner i de europæiske stormagter Spanien, Frankrig og Italien i gennemsnit er benchmark ca. 3-5 gange årligt. Derudover optræder regionerne i store lande som Tyskland og Storbritannien i gennemsnit hhv. 0,01 og 0,26 gange som benchmark om året, hvilket forekommer, da disse to lande indeholder meget få efficiente regioner i løbet af tidsperioden. I Storbritannien er Nordirland den region, som er benchmark flest gange. Nordirland er benchmark 12 gange i tidsperioden, hvilket er et gennemsnit på 0,8 gange om året.

4.2. DEA-analyse på det poolede datasæt

Dette afsnit omhandler resultater, hvor jeg har foretaget en DEA-analyse på hele datasættet på en gang. DEA resultaterne i denne analyse beskriver, hvordan et lands regioner i gennemsnit præsterer ift. de regioner, som klarer sig bedst i hele datasættet (dvs. uanset hvilket år de er fra). Analysen er en nødvendighed for direkte at kunne sammenligne efficiensscorer på tværs af årene (uden at bruge Malmquist indekset endnu). På baggrund af analysen er det muligt at se, hvordan efficiensscorene for landene i EU15 har udviklet sig i løbet af hele tidsperioden, dvs. om de blevet mere efficiente. I slutningen af afsnittet fokuseres igen på, hvilke regioner som er benchmark i den denne analyse.

Når jeg foretager en DEA-analyse på det poolede datasæt, optræder hver enkelt region per år som en observation. Dvs., at Tyskland, som består af 13 regioner og er med i datasættet i hele tidsperioden, i denne analyse indgår med 195 ($13 \cdot 15 = 195$) observationer. Samlet består hele datasættet, som tidligere skrevet, af i alt 1656 observationer²⁰.

²⁰ Rent praktisk kan DEA ikke løses for hele datasættet på en gang i Excel pga. antallet af justerbare celler (λ -værdierne), når Solveren anvendes. I Solveren kan der maksimalt justeres 200 celler ad gangen, og da datasættet består af 1656 observationer overstiges denne begrænsning klart. Denne udfordring har jeg omgået ved at løse små dele af datasættet ad gangen, idet jeg udnyttede, at de regioner som er inefficente, da jeg beregnede DEA på hvert år for sig, naturligvis må have vægte lig 0 uanset hvilken region, som evalueres. Hermed reduceres antallet af de justerbare celler markant. Jeg startede med at beregne perioden 2005-2009, hvorefter jeg påførte flere og flere år, igen uden at det maksimale antal justerbare celler blev overgået. Dette gentog jeg indtil hele tidsperioden var dækket, hvilket betød at benchmarks kun bestod af efficiente regioner.

Samlet set er 77 observationer efficiente, når jeg foretager en DEA-analyse på de i alt 1656 observationer. De 77 observationer består af 31 forskellige regioner i 11 forskellige lande, hvilket betyder, at der er regioner, som er efficiente på forskellige tidspunkter. Canarias (i Spanien), Bretagne (i Frankrig) og Holland (i Holland) er de regioner, som er efficiente flest gange, og de optræder på den efficiente rand hhv. 9, 6 og 6 gange hver. EU15-gruppen står for 70 af de 77 observationer, som er efficiente, hvilket selvfølgelig hænger sammen med, at denne gruppe står for langt den største del af de 1656 observationer, men også at det er i denne gruppe, hvor regionerne i gennemsnit per år har de største efficiensscorer (kan ses i tabellen nedenfor). I øvrigt skal det tilføjes, at grupperne EU10 og EU2 har hhv. 6 og 1 efficiente observationer på den efficiente rand, samt at det for alle 3 grupper gælder, at de har flest efficiente regioner i 2007. 2007 og 2008 er de år, som er repræsenteret med flest efficiente regioner med hhv. 13 og 10 regioner. I perioden 1998-2001 er der færrest efficiente regioner med omkring 1-2 regioner per år.

Efficiensscorerne for de 27 EU-lande, når jeg foretager en DEA-analyse på det poolede datasæt, er for perioden 1995-2009 præsenteret nedenfor.

Tabel 4.2: Gennemsnitlige årlige efficiensscorer på landebasis, 1995-2009 (hele datasættet på en gang)

Lande	Gruppe	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	Middel
(BEL) Belgium	EU15	0,81	0,83	0,78	0,75	0,76	0,84	0,81	0,73	0,76	0,78	0,78	0,78	0,83	0,82	0,76	0,79
(DAN) Denmark	EU15	0,86	0,90	0,90	0,78	0,84	0,95	0,98	0,88	0,88	0,92	1,00	1,00	1,00	1,00	0,97	0,92
(DEU) Germany	EU15	0,73	0,71	0,73	0,69	0,73	0,77	0,76	0,71	0,70	0,75	0,74	0,74	0,80	0,77	0,73	0,74
(ELL) Greece	EU15	0,75	0,71	0,73	0,67	0,68	0,70	0,72	0,74	0,75	0,73	0,78	0,74	0,70	0,79	0,72	0,73
(ESP) Spain	EU15	0,80	0,78	0,82	0,81	0,81	0,82	0,84	0,80	0,83	0,82	0,81	0,80	0,81	0,78	0,72	0,80
(FRA) France	EU15	0,76	0,75	0,76	0,76	0,77	0,78	0,76	0,83	0,81	0,82	0,83	0,84	0,88	0,86	0,78	0,80
(IRE) Ireland	EU15	0,92	0,88	0,96	0,85	0,82	0,92	0,91	0,82	0,83	0,87	0,87	0,87	0,91	0,78	0,63	0,85
(ITA) Italy	EU15	0,70	0,72	0,71	0,71	0,68	0,68	0,74	0,74	0,74	0,74	0,74	0,72	0,77	0,76	0,77	0,73
(LUX) Luxembourg	EU15	0,94	0,91	0,87	0,93	0,92	0,91	0,92	0,82	0,85	0,83	0,88	0,74	0,88	0,85	0,74	0,87
(NED) Netherlands	EU15	0,98	1,00	1,00	0,90	0,92	1,00	0,99	0,93	0,94	0,94	0,97	1,00	1,00	1,00	0,97	0,97
(OST) Austria	EU15	0,86	0,91	0,91	0,88	0,90	0,91	0,97	0,90	0,81	0,85	0,90	0,90	0,84	0,88	0,79	0,88
(POR) Portugal	EU15	0,53	0,51	0,52	0,50	0,56	0,55	0,57	0,59	0,59	0,60	0,64	0,67	0,63	0,63	0,61	0,58
(SUO) Finland	EU15	0,72	0,67	0,65	0,68	0,68	0,69	0,72	0,70	0,73	0,72	0,69	0,69	0,73	0,80	0,71	0,71
(SVE) Sweden	EU15	0,80	0,78	0,80	0,74	0,72	0,76	0,74	0,71	0,68	0,73	0,72	0,73	0,80	0,84	0,64	0,74
(UKI) United Kingdom	EU15	0,80	0,80	0,77	0,68	0,69	0,74	0,78	0,76	0,77	0,74	0,75	0,76	0,83	0,77	0,72	0,76
(CYP) Cyprus	EU10										0,78	0,85	0,77	0,84	0,91	1,00	0,86
(CZE) Czech Republic	EU10										0,51	0,54	0,52	0,54	0,57	0,50	0,53
(EST) Estonia	EU10										0,67	0,73	0,59	0,64	0,69	0,58	0,65
(HUN) Hungary	EU10										0,57	0,59	0,55	0,66	0,74	0,52	0,60
(LTU) Lithuania	EU10										0,47	0,49	0,51	0,56	0,55	0,48	0,51
(LVA) Latvia	EU10										0,54	0,62	0,56	0,56	0,61	0,53	0,57
(MLT) Malta	EU10										1,00	0,90	1,00	1,00	0,98	1,00	0,98
(POL) Poland	EU10										0,66	0,68	0,63	0,68	0,68	0,59	0,65
(SVK) Slovakia	EU10										0,44	0,46	0,47	0,51	0,53	0,40	0,47
(SVN) Slovenia	EU10										0,80	0,76	0,77	0,51	0,51	0,48	0,64
(BGR) Bulgaria	EU2													0,53	0,55	0,50	0,53
(ROU) Romania	EU2													0,72	0,63	0,65	0,67
Middel EU15		0,75	0,74	0,75	0,73	0,74	0,75	0,77	0,76	0,77	0,77	0,77	0,77	0,81	0,80	0,74	0,76
Middel EU10											0,62	0,64	0,61	0,66	0,69	0,58	0,63
Middel EU2														0,64	0,59	0,59	0,61
Middel		0,75	0,74	0,75	0,73	0,74	0,75	0,77	0,76	0,77	0,75	0,75	0,75	0,77	0,76	0,70	0,75

Kilde: Egne beregninger fra data

Ovenstående tabel er opbygget på præcis samme måde som Tabel 4.1, eneste forskel er som tidligere skrevet, at regionerne holdes op imod de bedst præsterende regioner i hele datasættet i stedet for at de sammenlignes med regioner fra samme år. For middelværdierne nederst i tabellen er der igen taget højde for, hvor mange regioner hvert land består af.

Igen vender jeg fokus mod regionen Baleares (i Spanien), hvor datapunkterne i 1998 og 1999 er fremstillet ved en lineær interpolation. Baleares har efficiensscorer på 0,8542 i 1998 og 0,9419 i 1999, dvs. at regionen i begge år er inefficent, og dermed ikke er benchmark for andre regioner. At Baleares ikke ligger på den efficiente rand, medfører at problematikken fra det forrige afsnit ikke bliver et tema i denne analyse. Hvis Baleares havde optrådt på den efficiente rand, ville jeg være nødsaget til at revurdere, om regionen skulle fjernes fra datasættet, da den derved havde haft en endnu større indflydelse på de andre regioner end først forudset.

Ift. Tabel 4.1 er efficiensscorerne generelt blevet mindre, da regionerne nu bliver sammenlignet med de allerbedste i datasættet (effekten er metodisk). Efficiensscorerne i tabellen ovenfor spænder fra et gennemsnit på 0,47 for Slovakiet til maksimum på 0,98 for Malta med en gennemsnitlig årlig efficiens for en region i datasættet angivet til at være 0,75. Man skal være lidt forsigtig med fortolkningen af den gennemsnitlige efficiens, da denne indeholder en sammenligning over forskellige år, og benchmarks måske ikke er opnåelige (i et år med dårligt landbrugsvejr er det umuligt at nå de samme resultater som i andre år). Tilsvarende hvis teknologien har ændret sig over tid, så kan man f.eks. få produktionsfaktorer med højere output (forædlede kornsorter, nye dyreracer osv.).

I ovenstående tabel er der ingen lande/regioner, som udviser perfekt efficiens igennem hele tidsperioden, hvilket også ville have været særdeles opsigtsvækkende. Det overraskende ligger i, at regionerne i denne analyse bliver sammenlignet med sig selv (og mange andre) på forskellige tidspunkter, og da tidsperioden strækker sig over 15 år, vil der være år som f.eks. er præget af dårligt landbrugsvejr. Resultaterne for landbrugsproduktionen (datapunkterne) vil selvfølgelig blive påvirket af et år med en dårlig høst, og da der for hver region findes unikke sammenlignelige enheder, er det praktisk talt umuligt at optræde på den efficiente rand i hele tidsperioden. Danmark, Holland og Malta er ifølge tabellen igen de lande, som hyppigst ligger på den efficiente rand. Udover disse er det faktisk kun Cypern, som (i 2009) har en efficiensscore på 1.

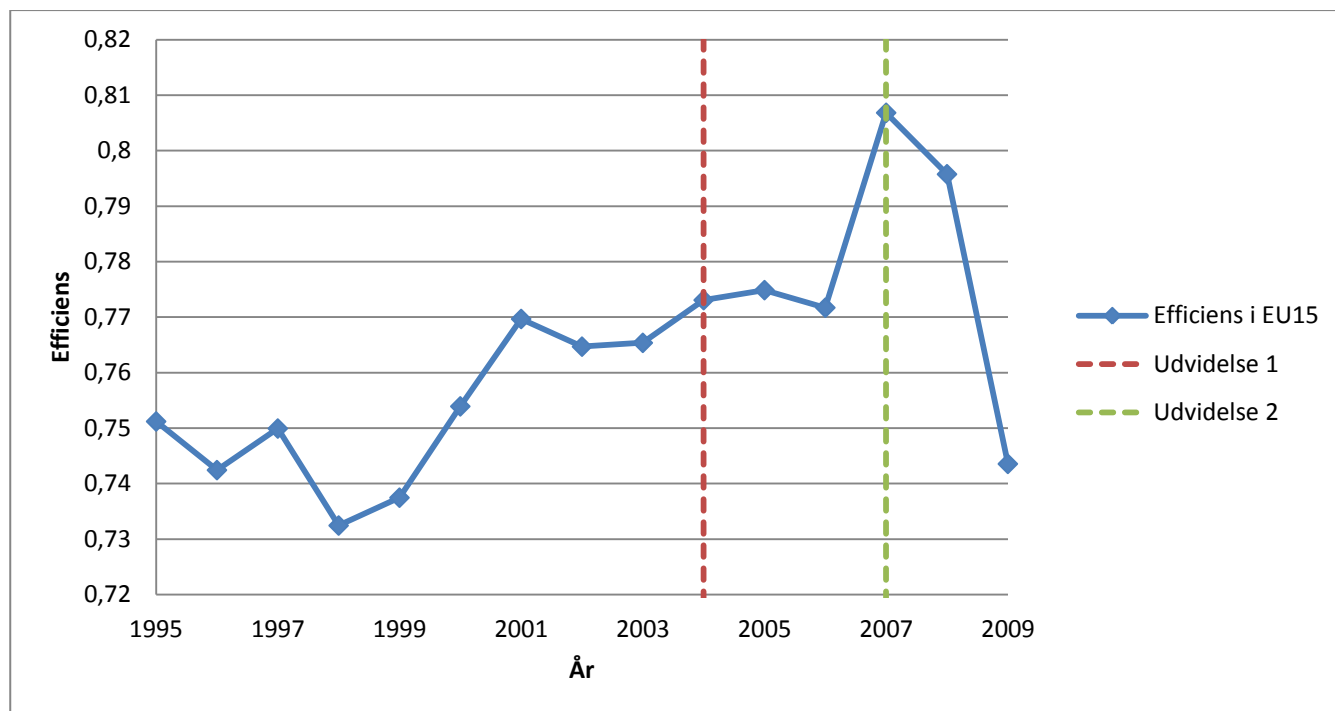
I EU15-gruppen er det Danmark og Holland, som klarer sig markant bedre end de andre lande. Danmark og Holland har gennemsnitlige efficiensscorer på hhv. 0,92 og 0,97, hvilket er et pænt stykke foran de lande, som ligger nærmest (Østrig og Luxembourg). Hollands høje gennemsnitlige

score forstærkes af, at Holland faktisk er det land i EU15, som præsterer bedst (nogle år deles denne status med andre lande) i alle år på nær i 1998 og 2005, hvor hhv. Luxembourg og Danmark har højere efficiensscorer. Portugal og Finland er, ligesom ved DEA-analysen på hvert år for sig, de to lande, som har de laveste gennemsnitlige efficiensscorer i EU15-gruppen på hhv. 0,58 og 0,71.

I grupperne EU10 og EU2 er det kun Malta (tilhørende EU10) og Cypern (tilhørende EU10), der har efficiensscorer over gennemsnittet på 0,75. Malta er faktisk det land i datasættet, som klarer sig bedst af alle lande med en gennemsnitlig efficiensscore på 0,98 per år. I det hele taget præsterer disse to grupper ikke specielt overbevisende i denne analyse. De årlige efficiens gennemsnit for EU10 ligger et i interval fra 0,58 til 0,69 (med et samlet gennemsnit på 0,63), hvor de maksimale værdier forekommer i 2007 og 2008. EU2 har et samlet gennemsnit på 0,61, og den største efficiens findes i 2007 med 0,64.

Mit hovedfokus i denne afhandling er at se på produktivitetsudviklingen i landbruget for de gamle medlemslande i den Europæiske Union. Derfor har jeg, for overskuelighedens skyld, illustreret udviklingen i de gennemsnitlige efficiensscorer for EU15-gruppen på nedenstående figur. De nøjagtige værdier kan ses i Tabel 4.2.

Figur 4.2: Udviklingen i gennemsnitlig efficiens for EU15



Kilde: Egne beregninger fra data

På figuren er de 2 udvidelser af EU (i 2004 og 2007) angivet med stiplede linjer. På trods af en efficienstilbagegang i årene 1996, 1998 og 2002 ses der en klar opadgående tendens i tidsperioden før den 1. udvidelse i 2004. Den gennemsnitlige efficiensscore for regionerne i EU15 ligger i 1995 på 0,75 og bevæger sig stille og roligt op i et leje omkring 0,77 i 2003. Omkring udvidelsen i 2004 og i årene frem til den 2. udvidelse ligger den gennemsnitlige efficiensscore stabilt omkring de 0,77 uden nævneværdige udsving. Den opadgående tendens som anes inden 1. udvidelse bliver vedligeholdt. I forbindelse med udvidelsen i 2007 ses den højeste gennemsnitlige efficiens på 0,81, hvor regionerne i EU15 i gennemsnit er blevet ca. 0,04 mere efficient end året før. Kurven toppe i 2007, hvilket ikke kommer som den store overraskelse, idet 2007 også indeholder klart flest regioner på den efficiente rand af alle årene i datasættet. Det maksimale punkt i 2007 betyder, at at værdien af landbrugsproduktionen, ift. det anvendte inputniveau, set over hele tidsperioden var bedst i 2007. I årene herefter bliver efficiensscorene markant mindre og ender i 2009 på 0,74. Middelværdien af de gennemsnitlige efficiensscore i EU15 for perioden inden 1. udvidelse (1995-2003) er 0,75, for perioden mellem 1. og 2. udvidelse (2004-2006) er middelværdien 0,77 og for perioden efter 2. udvidelse (2007-2009) er middelværdien 0,78. Middelværdierne viser, at landene i

EU15 faktisk er blevet mere effiente i takt med, at EU er blevet større. For at fastslå om disse forskelle er signifikante, har jeg foretaget en sammenligning af efficiensen i de 3 tidsperioder. Jeg har anvendt Mann-Whitneys U test²¹ til dette formål, da DEA-scorer ikke har pæne statistiske egenskaber, og det derved ikke kan foretages med et almindeligt t-test på gennemsnittene. Mann-Whitney-testen er en ikke-parametrisk (her i den forstand at det ikke bygger på en normalfordelingsantagelse) udgave af et "t-test med to uparrede stikprøver", og fremføres ved at lave en fælles rangordning af observationerne i to tidsperioder. Jeg bruger Mann-Whitneys U test, til at sammenligne middelværdierne (af effizienz) i to tidsperioder. Hypotesen (nulhypotesen) som testes er, at de to grupper har ens middelværdier, mens modhypotesen siger, at middelværdierne er forskellige. Forskellen i effizienz for EU15-gruppen for perioden inden første udvidelse og perioden mellem første og anden udvidelse er ifølge testen signifikant²² (signifikansniveau på 0,05). Derimod er forskellen i effizienz for perioden mellem de to udvidelser og for perioden efter anden udvidelse ikke signifikant²³ (signifikansniveau på 0,05). En af grundene til at effizienzgraden for EU15 generelt er steget i løbet af tidsperioden kan være, at der i tidsperioden har været generelle tekniske forbedringer. Sådanne forbedringer vil have gjort regionerne mere effektive, og dermed også mere produktive, men dette vil jeg tage fat på i Malmquist-analysen.

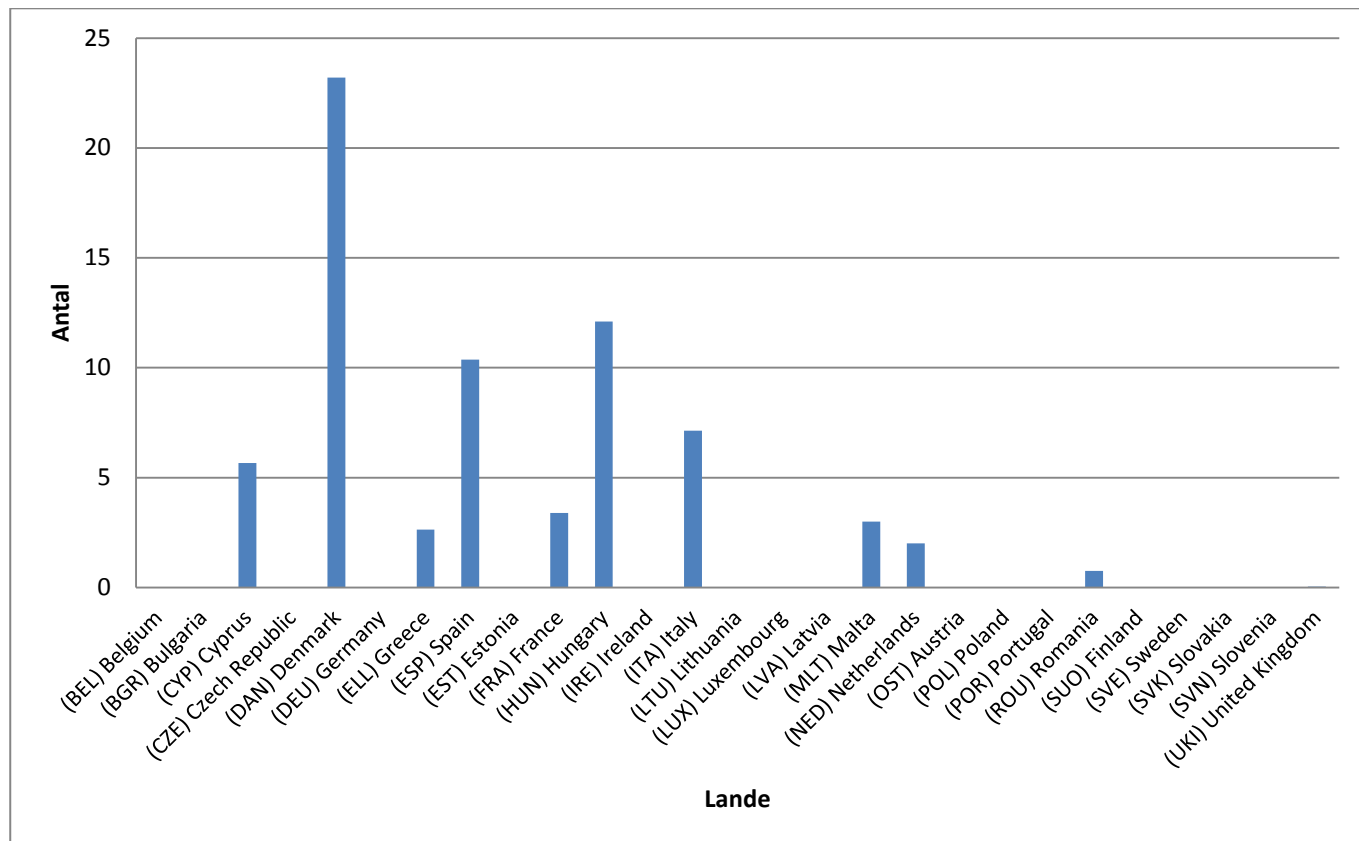
For at afrunde denne sektion har jeg, som i det forrige afsnit, illustreret et søjlediagram på nedenstående figur for at give et billede af hvilke landes regioner, som hyppigst optræder som benchmark for andre regioner. Søjlerne viser, hvor mange gange et lands regioner total set er benchmark per år.

²¹ Kilde til Mann-Whitney-testen: <http://statnoter.biolyt.dk/index.php?pageID=81>. Se bilag 4 for en mere detaljeret gennemgang/beskrivelse af Mann-Whitney-testen.

²² Jeg får en z-værdi på -2,35, og hypotesen forkastes, da værdien er numerisk større end 1,96.

²³ Jeg får en z-værdi på -1,27, og hypotesen accepteres, da værdien er numerisk mindre end 1,96.

Figur 4.3: Det samlede antal gange et lands regioner er benchmark per år (hele datasættet på en gang)



Kilde: Egne beregninger fra data

I ovenstående diagram er 11 lande repræsenteret med værdier større end 0, hvilket betyder, at 16 lande udelukkende indeholder inefficiente regioner, og dermed ikke har regioner, der optræder som peers for andre regioner. De 11 lande er fordelt med hhv. 7 lande i EU15 (Danmark, Grækenland, Spanien, Frankrig, Italien, Holland og Storbritannien), 3 lande i EU10 (Cypern, Ungarn og Malta) og 1 land i EU2 (Rumænien).

Som tidligere skrevet er det i 2007, der er flest effektive regioner, og faktisk udgør de effektive regioner i 2007 omkring 31 % af det samlede antal gange en region indgår som benchmark. Topscoren (på landebasis) er ligesom, da jeg så på benchmark i forbindelse med DEA-analysen på hvert år for sig, Danmark, som total set er benchmark flest antal gange per år (23 gange per år). Danmark (består af 1 region) er effektiv og benchmark i tidsperioden 2005-2008. En af de største

forskelle på ovenstående figur og Figur 4.1 er, at Holland slet ikke i samme grad indgår som benchmark, og faktisk optræder Holland samlet set som benchmark færre gange end Malta. Grunden til at Holland i denne analyse ikke længere er så vigtig som benchmark hører sammen med, at ingen af de 6 datapunkter (de år hvor Holland er efficient) i særlig høj grad bliver brugt som benchmark. Faktisk er det kun i 2008, at Holland er benchmark for over 10 regioner (16 gange benchmark). I det poolede datasæt er det Ungarn og Spanien, som har overtaget Hollands rolle, og disse optræder total set som benchmark hhv. 8 og 5 gange yderligere om året (ift. Figur 4.1). I Ungarn er Közép-Magyarország den eneste region, som er benchmark, og faktisk er stort set hele Ungarns hyppighed skabt af Közép-Magyaróságs resultater i 2008, hvor regionen er benchmark for 507 regioner (Közép-Magyarország er benchmark 508 gange i alt). I Spanien er Canarias den region, som er benchmark flest gange, men ift. Ungarn er Spaniens høje hyppighed i antallet af benchmark skabt på baggrund af mange forskellige efficiente regioner (10 i alt).

4.3. Malmquist indekset

I dette afsnit vil jeg fremvise resultaterne for Malmquist indekset. Malmquist indekset belyser den tidsmæssige udvikling i produktiviteten, hvor der vælges et bestemt år, *basisåret*, som sammenlignings år²⁴. Motivationen bag anvendelsen af Malmquist indekset er bl.a., at indekset kan opsplittes i 2 dele. Opsplitningen gør det nemlig muligt at se på, om produktivitsændringerne er en effekt af efficiensændring (Catch-up effekt) eller teknologisk ændring (Technical change/randeffekt).

Jeg har i dette afsnit valgt at se nærmere på landene i EU15 og på selve gruppen EU15 i al almindelighed, hvilket har betydet, at jeg kun har fundet Malmquist indekset og dets komponenter for disse lande²⁵. Denne beslutning er yderligere truffet ud fra, at data for landene i grupperne EU10 og EU2 ikke er tilgængeligt i hele tidsperioden, og det derfor ikke vil give mening at finde Malmquist indeks for disse lande i de få år.

Dette afsnit er indledt i 2 sektioner. Første del er med hovedfokus på produktivitsudviklingen i hvert land for sig, hvorimod jeg i anden del beskæftiger mig med gruppen som helhed.

²⁴ Når jeg i dette afsnit angiver årstal i forbindelse med Malmquist resultater er årstallet et udtryk for det år, der skiftes til. Dvs. produktivitsændringen fra 1995 til 1996, er angivet som 1996.

²⁵ Observationerne for EU10 og EU2 indgår dog stadig i analysen ved at være en del af de observationer, der sammenlignes med.

Som tidligere nævnt er de efficiente regioner i Malmquist-analysen de samme, som jeg beskæftigede mig med i afsnittet DEA-analyse på hvert år for sig. Derudover skal det præciseres at alle resultater, som er fremstillet i dette afsnit er fundet vha. et geometriske gennemsnit over antallet af observationer (regioner) i det pågældende land eller gruppe. Det geometriske gennemsnit er gunstigt at anvende, når man ser på procentvise stigninger for indeks, og det kommer i dette kapitel til sin ret, da jeg beregner de gennemsnitlige årlige produktivitetsændringer og dets komponenter.

I forbindelse med udregningerne af Catch-up, Technical change og Malmquist indekset for de 100 regioner i EU15 for samtlige af tilstødende år, stødte jeg på et problem. Som jeg noterede tidligere i Statistisk beskrivelse af data afsnittet skaber regionen Canarias i Spanien problemer i skiftet mellem 1999 og 2000, hvilket sker da både Dyr (inputvariabel) og Udbytte af dyr (outputvariabel) i 1999 har værdier lig 0, mens Canarias i 2000 har positive værdier for begge variable. Ændringen i datapunkterne gør, at resultatet af det lineære problem i Malmquist, hvor Canarias i 1999 holdes op imod randen i 2000 bliver 0. Den reciprokke værdi af denne løsning indgår (jf. definitionen af Malmquist indekset) i nævneren, når Malmquist indekset beregnes, hvilket gør, at udregningen af Malmquist indekset ikke kan lade sig gøre. Problematikken (med regionen Canarias) har jeg omgået ved manuelt at lave værdierne for hhv. Catch-up, Technical change og Malmquist Indeks blanke (fjerne værdierne) i tidsperioden 1996-2000 (den periode hvor der optræder 0-værdier), så regionen ikke indgår i gennemsnittet for hhv. Spanien og EU15 i disse år. Jeg har ikke valgt at fjerne regionen fra datasættet, da Canarias er peer i hele tidsperioden, når jeg foretager DEA-analysen på hvert år for sig²⁶, og den tilmed indgår som benchmark i alle årene undtagen 2009. Canarias indgår i gennemsnit som benchmark lidt over 17 gange per år, hvilket betyder, at regionen er en særdeles vigtig sammenlignelig enhed for andre regioner.

For at se på om der er sammenfald mellem de 2 udvidelser og produktiviteten i EU15-landene, har jeg opdelt gruppen i 3 tidsperioder. 1. tidsperiode indeholder årene inden 1. udvidelse (1996-2003), den 2. tidsperiode består af årene efter 1., men inden 2. udvidelse (2004-2006) og den 3. tidsperiode indbefatter perioden efter 2. udvidelse (2007-2009). I det følgende har jeg fremstillet tabeller for de 3 tidsperioder, der viser de gennemsnitlige årlige værdier for efficiensforandring, randeffekt og TFP ændring. I tabellerne er landene præsenteret i alfabetisk rækkefølge alt efter deres initialer, og for middelværdierne (gennemsnitlige årlige ændringer) angivet i bunden af hvert tabel, er der taget

²⁶ Faktisk er Canarias også peer i 1997, 1998, 2001, 2002, 2003, 2004, 2005, 2007 og 2008, når jeg foretager DEA-analyse på det poolede datasæt. Derudover er regionen benchmark i alle årene på nær i 2002 og 2003.

højde, for hvor mange regioner hvert land består af (gennemsnit bliver på regionsbasis). I alle 3 tabeller er Catch-up indekset lig 1 for både Danmark og Holland, hvilket stemmer overens med, at både Danmark og Holland ligger på den efficiente rand i hele tidsperioden i den første DEA-analyse. Tabellen for den 1. tidsperiode ser ud som følger.

Tabel 4.3: Gennemsnitlig årlig produktivitet og dets komponenter for lande tilhørende EU15 i tidsperioden 1996-2003

Land	Catch-up	Technical change	Malmquist Indeks
(BEL) Belgium	0,9905	1,0092	0,9996
(DAN) Denmark	1,0000	1,0153	1,0153
(DEU) Germany	0,9855	1,0071	0,9925
(ELL) Greece	0,9925	1,0206	1,0130
(ESP) Spain	0,9977	1,0276	1,0252
(FRA) France	0,9989	1,0065	1,0054
(IRE) Ireland	0,9965	0,9884	0,9850
(ITA) Italy	1,0033	1,0110	1,0143
(LUX) Luxembourg	0,9920	0,9958	0,9879
(NED) Netherlands	1,0000	1,0066	1,0066
(OST) Austria	0,9880	1,0059	0,9938
(POR) Portugal	1,0069	1,0055	1,0124
(SUO) Finland	1,0059	1,0044	1,0104
(SVE) Sweden	0,9902	1,0027	0,9929
(UKI) United Kingdom	0,9968	0,9989	0,9957
Middel	0,9975	1,0107	1,0082

Kilde: Egne beregninger fra data

Tabellen viser, at regionerne i Spanien og Danmark, i perioden 1996-2003, er de regioner med den gennemsnitlig største TFP vækst. Spanien har en årlig TFP vækst på 2,52 %, hvilket ene og alene domineres af en stigning i randeffekten på 2,76 %. Den gennemsnitlige årlige TFP vækst er 0,82 % i 1. tidsperiode, på trods af at Catch-up indekset er mindre end 1. Derudover ses der en positiv randeffekt for hele perioden, selvom der for nogle individuelle lande (Irland, Luxembourg og Storbritannien) er bevis for tilbagegang. Randeffekten siger noget om, hvordan randen har bevæget

sig, og at der både optræder negative og positive randeffekter betyder, at dele af randen har forbedret sig, mens andre dele er blevet forringet. Rent matematisk indebærer det, at randen i et år krydser randen i det næste år, hvilket betyder, at der ikke registreres en generel teknisk forbedring for alle regioner.

En af grundene til at der i tabellen fremgår et gennemsnitligt årligt Catch-up indeks på under 1 kan bl.a. forklares via de betydelige fald i efficiens i Storbritannien. Disse ses ikke lige umiddelbart i ovenstående tabel, men de fremkommer tydeligt i mine resultater i perioden 1996-1998.

Efficienstabset er korreleret med udbredelsen af sygdommene kogalskab (BSE)²⁷ og mund- og klovsygen (FMD)²⁸ i Storbritannien (USDA, 2001). I forbindelse med forekomsten af sygdommene i Storbritannien i midten af 1990'erne bredte der sig en forvirring blandt befolkningen i Storbritannien om påvirkningen af hhv. BSE og FMD. En effekt heraf kunne i 1996 registreres ved en øjeblikkelig nedgang i salget af oksekødsprodukter på 40 %, og de samlede anslåede omkostninger ved BSE for Storbritannien var omkring 5 mia. £ (USDA, 2001). Sygdommene har haft en direkte effekt på efficiensscorene, hvilket i øvrigt ses på Tabel 4.1. De gennemsnitlige efficiensscorer for Storbritannien er 0,92 i 1995, 0,91 i 1996, 0,88 i 1997 og 0,80 i 1998, hvilket er en klar negativ tendens. Denne tendens gør, at der i årene fra 1996 til 1998 fremkommer et stort gennemsnitligt negativt Catch-up indeks på ca. 5 % per år for Storbritannien. En anden forklaring på at det gennemsnitlige årlige Catch-up indeks < 1 kunne være de negative efficiensændringer på omkring 5-6 %, som registreres i Tyskland i 1996 og 1998. Den forholdsvis høje efficienstilbagegang kan skyldes, at Tyskland stadig i denne tid var under genopbygning efter landets opsplitning. Desuden skal det bemærkes, at et negativt Catch-up ofte ses, når randen forbedres (som i denne tidsperiode), da der går lidt tid inden de inefficiente regioner kan følge med de teknologiske forbedringer.

I det følgende vil jeg gå videre til næste tidsperiode, og tabellen for 2. tidsperiode er præsenteret nedenfor.

²⁷ Kogalskaben er kendt under forkortelsen BSE (*Bovine Spongiform Encephalopathy*).

²⁸ Mund- og klovsygen er kendt under forkortelsen FMD (*Foot-and-mouth disease*).

Tabel 4.4: Gennemsnitlig årlig produktivitet og dets komponenter for lande tilhørende EU15 i tidsperioden 2004-2006

Land	Catch-up	Technical change	Malmquist Indeks
(BEL) Belgium	0,9947	1,0065	1,0011
(DAN) Denmark	1,0000	1,0677	1,0677
(DEU) Germany	1,0201	1,0199	1,0404
(ELL) Greece	0,9835	0,9859	0,9696
(ESP) Spain	0,9915	0,9578	0,9496
(FRA) France	0,9919	1,0228	1,0145
(IRE) Ireland	1,0200	0,9929	1,0128
(ITA) Italy	1,0013	0,9893	0,9905
(LUX) Luxembourg	1,0230	1,0219	1,0454
(NED) Netherlands	1,0000	1,0399	1,0399
(OST) Austria	1,0237	1,0094	1,0333
(POR) Portugal	1,0383	1,0110	1,0498
(SUO) Finland	0,9637	1,0290	0,9917
(SVE) Sweden	1,0519	1,0059	1,0582
(UKI) United Kingdom	1,0033	1,0153	1,0186
Middel	1,0013	1,0014	1,0026

Kilde: Egne beregninger fra data

Ift. tabellen for den 1. tidsperiode ses en positiv tilvækst for alle 3 indeks. Selvom den gennemsnitlige årlige TFP vækst (i 2. tidsperiode) for EU15-landene er positiv, oplever både Grækenland, Spanien, Italien og Finland produktivitetstab. Grækenland og Spanien har tilbagegang i både efficiens og randeffekt, hvorimod Italien og Finland har produktivitetstab udelukkende domineret af hhv. randeffekten og efficiensforandring. I tabellen er Malmquist indekset for hvert enkelt land spredt udover et meget stort interval. TFP ændringerne spænder fra et produktivitetstab på 5 % i Spanien til en produktivitetsfremgang på næsten 7 % i Danmark. Det store produktivitetstab i Spanien fremkommer hovedsageligt pga. et gennemsnitligt tab på 10 % i 2004, som skabes via store produktivitetstab i regionerne Galicia, Asturias og Canarias. Danmarks

fremgang fremkommer, da de gennemsnitlige årlige stigninger i 2004, 2005 og 2006 alle ligger omkring 6-8 %.

Den gennemsnitlige årlige produktivitetsudvikling, efter udvidelsen i 2004 og inden 2. udvidelse i 2007, er på 0,26 %. Værdien er mindre end i 1. tidsperiode (TFP tilvækst på 0,82 %), men faktisk bibeholdes den svagt stigende tendens i den gennemsnitlige årlige produktivitetsudvikling for landene i EU15 på trods af, at data er blevet udvidet med 10 lande.

I det følgende retter jeg fokus mod den 3. tidsperiode, og tabellen for denne periode er præsenteret nedenfor.

Tabel 4.5: Gennemsnitlig årlig produktivitet og dets komponenter for lande tilhørende EU15 i tidsperioden 2007-2009

Land	Catch-up	Technical change	Malmquist Indeks
(BEL) Belgium	0,9858	1,0085	0,9942
(DAN) Denmark	1,0000	0,9964	0,9964
(DEU) Germany	1,0147	0,9838	0,9982
(ELL) Greece	0,9955	1,0064	1,0019
(ESP) Spain	0,9669	1,0060	0,9727
(FRA) France	0,9991	0,9688	0,9679
(IRE) Ireland	0,9146	0,9597	0,8778
(ITA) Italy	0,9840	1,0440	1,0274
(LUX) Luxembourg	0,9524	1,0078	0,9599
(NED) Netherlands	1,0000	0,9979	0,9979
(OST) Austria	0,9631	0,9935	0,9568
(POR) Portugal	0,9783	0,9997	0,9781
(SUO) Finland	1,0407	0,9747	1,0144
(SVE) Sweden	0,9664	0,9966	0,9631
(UKI) United Kingdom	1,0054	0,9922	0,9976
Middel	0,9906	0,9993	0,9898

Kilde: Egne beregninger fra data

I modsætning til tabellen for 2. tidsperiode er alle indeks under 1. TFP tabet stammer hovedsageligt fra Spanien og Frankrig, fordi disse lande påvirker de gennemsnitlige årlige ændringer markant pga. deres store antal af regioner. Produktivitetstab i Spanien og Frankrig kommer fra hhv. Catch-up effekten og randeffekten. I øvrigt skal det siges, at Irland med 0,8778 har det gennemsnitlig laveste Malmquist indeks, men da Irland kun består af én region påvirker denne værdi ikke det årlige produktivitetsgennemsnit nævneværdigt.

I tabellen ovenfor er Italien det land, som påvirker produktivitetssudviklingen mest i den positive retning (2,74 %). Den positive tilvækst for Italien stammer udelukkende fra randeffekten, og faktisk er det kun Grækenland og Finland, som udover Italien har positive produktivitetsændringer. I 3. tidsperiode oplever landene i EU15 i gennemsnit en årlig negativ produktivitetsændring, hvor den i de to foregående tidsperioder har været positiv. Den negative produktivitetsændring fremkommer på trods af en stor produktivitetssfremskud i 2007 på knap 7,5 %. Fremskuddet i 2007 er en effekt af, at alle lande på nær Portugal og Østrig oplever fremskud, og i øvrigt er Sverige det land, som har den største produktivitetssstigning (næsten 13 %). I 2008 og 2009 falder det gennemsnitlige produktivitetsniveau med hhv. 2,4 % og 7,6 %, og det store fald i 2009 fremkommer, da alle lande undtagen Italien oplever en produktivitetssnedgang (Italien har en fremskud på 0,01 %). Disse tendenser fornemmes i øvrigt på Figur 4.2, hvor der i 2007 var en kolossal efficiens forbedring, og faktisk er den gennemsnitlige efficiens (ifølge figuren) også størst i 2007. Ligesom i ovenstående tabel bliver denne forbedring opvejet af en endnu større tilbagegang i 2008 og 2009. Det gennemsnitlige årlige produktivitetstab i 3. tidsperiode er på 1,02 %, hvilket hovedsageligt stammer fra efficienstab.

Produktivitetstabene i ovenstående tabel falder sammen med, at den økonomiske krise er indtruffet i 3. tidsperiode. Mine resultater viser, at nogle af de lande, hvis økonomier har været hårdest ramt af krisen (Frandsen, 2010), også er blandt dem, som har haft den største produktivitetstilbagegang. Dette ses i hhv. Grækenland, Spanien og Irland. Grækenland har et TFP tab på 11 % i 2009, Spanien har et produktivitetstab på 10 % i 2009 og Irland har et tab på 18 % i 2008 og 19 % i 2009. Italien er det eneste land, som på trods af en skrantende økonomi i denne tidsperiode har en positiv produktivitetssudvikling.

Efter at have set på de 3 tidsperioder hver for sig, har jeg med henblik på at skabe et bedre overblik valgt at se på produktivetsprofilerne for samtlige EU15-lande for hele tidsperioden. Jeg har derfor opdelt landene i 5 kategorier alt efter deres TFP ændring, og i tabellen kan det registreres hvilke kilder, der tilskrives denne ændring.

Tabel 4.6: Gennemsnitlig årlig produktivitet, samt produktivetsprofil for hvert land i EU15 i %

Land	Catch-up	Technical change	Malmquist Indeks
<i>Lande som ligger på randen med produktivetsstigning</i>			
(DAN) Denmark	0,0000%	2,2151%	2,2151%
(NED) Netherlands	0,0000%	1,1762%	1,1762%
<i>Lande med produktivetsstigning grundet efficiensforbedring og positiv randeffekt</i>			
(POR) Portugal	0,7300%	0,5429%	1,2769%
(SUO) Finland	0,4019%	0,3150%	0,7182%
(UKI) United Kingdom	0,0025%	0,0954%	0,0979%
<i>Lande med produktivetsstigning grundet positiv randeffekt men efficienstab</i>			
(DEU) Germany	-0,0945%	0,4760%	0,3811%
(ELL) Greece	-0,8772%	1,0005%	0,1146%
(ITA) Italy	-0,1294%	1,3264%	1,1953%
<i>Lande med produktivetsstilbagegang grundet efficienstab</i>			
(BEL) Belgium	-0,9582%	0,8450%	-0,1212%
(ESP) Spain	-1,0342%	0,7656%	-0,2765%
(FRA) France	-0,2580%	0,1751%	-0,0833%
(LUX) Luxembourg	-1,0106%	0,3913%	-0,6233%
(OST) Austria	-0,9897%	0,3967%	-0,5969%
(SVE) Sweden	-0,2080%	0,2065%	-0,0019%
<i>Lande med produktivetsstilbagegang grundet både efficienstab og negativ randeffekt</i>			
(IRE) Ireland	-1,6728%	-1,6838%	-3,3284%

Kilde: Egne beregninger fra data

I de 5 ovenstående grupper er 3 af grupperne betegnet med produktivetsvækst. Halvdelen af landene i EU15 indikerer TFP fremskridt i løbet af perioden, og den anden halvdel oplever produktivitetstab. TFP tabene for Belgien, Spanien, Frankrig, Luxembourg, Østrig og Sverige er

udelukkende domineret af et fald i Catch-up effekten, hvorimod TFP tilbagegangen for Irland er præget af et fald i både Catch-up effekten og randeffekten. Danmark, Holland, Portugal og Italien er de 4 lande, som i gennemsnit årligt har den største produktivitetsvækst. Danmark topper med en årlig vækst på 2,22 %, og herefter følger Portugal, Italien og Holland med gennemsnitlige årlige produktivitetsvækster på hhv. 1,28 %, 1,20 % og 1,18 %. Danmark og Holland ligger på den efficiente rand i hele tidsperiode, dvs. at TFP væksten udelukkende stammer fra randeffekten.

I tabellen udviser Portugal, Finland og Storbritannien en stigning i både Catch-up effekten og randeffekten, hvilket selvfølgelig medfører en produktivitetsvækst i løbet af perioden. Tyskland, Grækenland og Italien viser teknologiske fremskridt, men et fald i Catch-up effekten, som dog samlet fører til en gennemsnitlig årlig produktivitetsvækst for de 3 lande.

For at kunne sige noget om EU15-gruppen som helhed har jeg valgt at se på, hvordan Malmquist indekset og dets komponenter ser ud per år for EU15 som en samlet gruppe. Tabellen nedenfor viser de gennemsnitlige årlige efficiensforandringer, randeffekter og TFP ændringer for de 100 regioner i EU15 over perioden fra 1996 til 2009.

Tabel 4.7: Gennemsnitlig årlig produktivitet og dets komponenter for EU15

År	Catch-up	Technical change	Malmquist Indeks
1996	0,9892	1,0032	0,9924
1997	0,9787	1,0308	1,0088
1998	0,9621	1,0205	0,9818
1999	1,0188	0,9823	1,0008
2000	1,0219	1,0041	1,0262
2001	1,0105	1,0085	1,0191
2002	0,9877	1,0359	1,0232
2003	1,0126	1,0015	1,0141
2004	0,9988	1,0005	0,9993
2005	0,9787	1,0097	0,9881
2006	1,0270	0,9940	1,0208
2007	0,9674	1,1109	1,0747
2008	1,0182	0,9589	0,9764
2009	0,9867	0,9366	0,9242
Middel 1996-2003	0,9975	1,0107	1,0082
Middel 2004-2006	1,0013	1,0014	1,0026
Middel 2007-2009	0,9906	0,9993	0,9898
Middel	0,9968	1,0062	1,0031

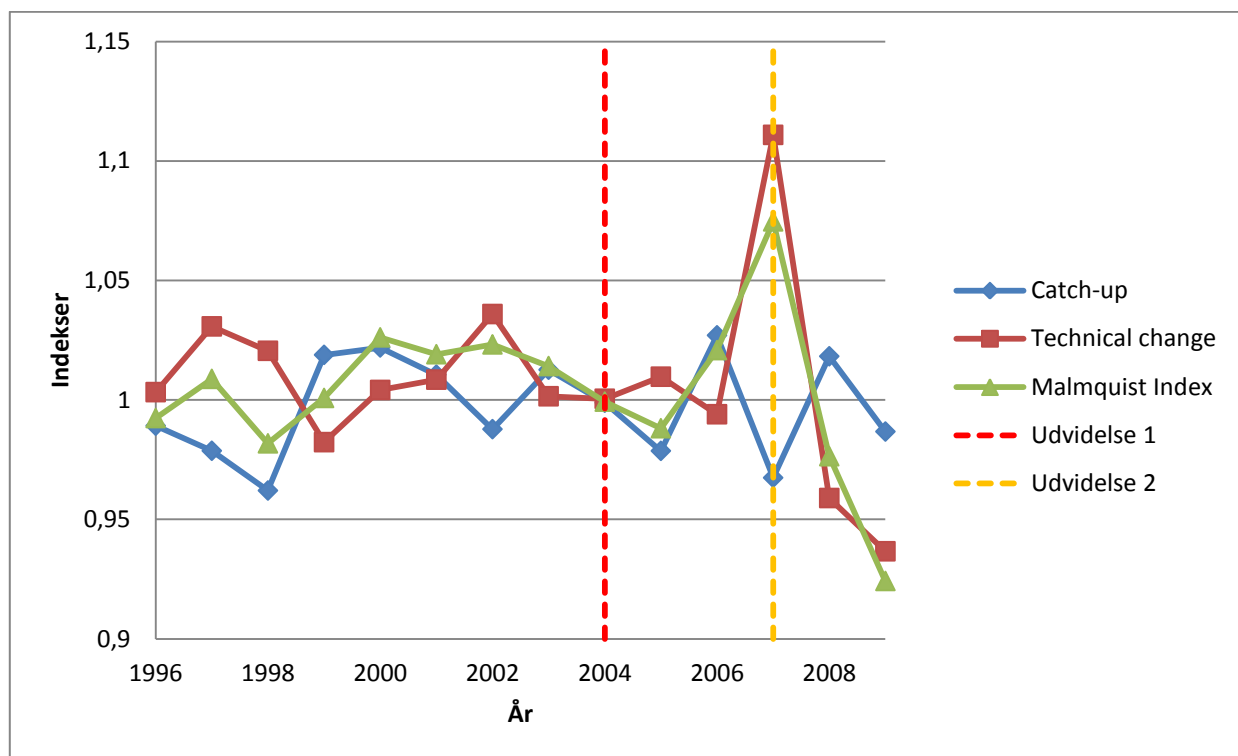
Kilde: Egne beregninger fra data

Det ses straks, at middelværdierne for de 3 tidsperioder er lig middelværdierne fra tabellerne for de 3 tidsperioder, som er præsenteret tidligere i dette afsnit. Desuden stemmer Catch-up værdierne overens med de tendenser, jeg tidligere så i Tabel 4.1.

Den gennemsnitlige årlige produktivitetsudvikling for en region i EU15-gruppen er på 0,31 %, og med en stigning i randeffekten på 0,62 %, er randeffekten alene skyld i TFP væksten. Resultaterne indikerer en maksimal TFP vækst i 2007 på 7,47 %, hvorimod efficiensforandringen viser et markant fald i 1998 og 2007, og en betydelig stigning i 2000 og 2006. Af tabellen ses en generel positiv randeffekt over hele perioden, dog med tilbagefald i årene 1999, 2006, 2008 og 2009.

Figuren nedenfor illustrer plots af udviklingen i Malmquist indekset og dets komponenter over tidsperioden 1996-2009.

Figur 4.4: Malmquist Indekset og dets komponenter for EU15



Kilde: Egne beregninger fra data

På figuren bevæger efficiens komponenten i Malmquist indekset sig modsat af randeffekten, hvilket går helt i tråd med, at Catch-up leddet er modpol til randeffekten jf. definitionen af Malmquist indekset. Derudover ligger grafen for Malmquist indekset generelt imellem graferne for hhv. efficiensforandringen og randeffekten, da Malmquist indekset er et geometrisk gennemsnit af de to komponenter.

Plottet ovenfor viser en stabil landbrugsproduktivitet med produktivitetsændringer på mindre en 3 % i alle årene på nær i 2007 og 2009, hvor der observeres en hhv. positiv og negativ ændring. Malmquist indekset ligger bortset fra i 1996 og 1998 over 1 i alle årene indtil 1. udvidelse. I 2. tidsperiode ligger produktivitsudviklingen meget stabilt omkring 1, men følger den lille positive tendens, som er i 1. tidsperiode. I 3. tidsperiode kan der ikke tales om et decideret mønster.

I 2007 ses den klart største værdi i produktivitetsudviklingen, hvilken udelukkende skabes på baggrund af randeffekten. I årene efter ses et endnu større fald, og i 2009 registreres et produktivitetstab på knap 8 % (Malmquist indeks på 0,92), hvilket er det største produktivitetsfald i hele tidsperioden (skabes via bidrag fra både Catch-up og randeffekten). At grafen topper i 2007 falder sammen med, at det er året inden finanskrisens indtog. Årene efter startede den økonomiske krise, og et sammenfald af denne kan anes ved de store produktivitetstab i 2008 og 2009.

I Tabel 4.7 ses, at middelværdierne for produktiviteten i de 3 tidsperioder bliver lavere i takt med, at EU udvides (anes også på Figur 4.4). For at fastslå om forskellen i produktiviteten mellem de 3 perioder er signifikante, har jeg ligesom under DEA-analysen for det poolede datasæt anvendt Mann-Whitneys U test. Grunden til at jeg igen bruger Mann-Whitney-testen er, at Malmquist bygger på teorien bag DEA, og det derved ikke er muligt at foretage et almindeligt t-test på gennemsnittene. Jeg foretager en fælles rangordning af observationerne i to tidsperioder ved at rangere Malmquists produktivitetsindeks alt efter hvilken region, som har den største produktivitetsmåling. Forskellen i produktiviteten for EU15-gruppen for perioden inden 1. udvidelse og perioden mellem 1. og 2. udvidelse er ifølge testen ikke signifikant²⁹ (signifikansniveau på 0,05). Derimod er forskellen i produktiviteten for perioden mellem de to udvidelser og for perioden efter 2. udvidelse signifikant³⁰ (signifikansniveau på 0,05). Om produktivitetsfaldene i middelværdierne er et sammenfald med udvidelserne af EU, vil jeg tage fat på i det følgende.

Som sidste element i dette kapitel ønsker jeg at intensivere undersøgelsen af, om der er sammenfald mellem de to udvidelser og produktivitetsmålene i EU15-landene. Helt konkret udøves undersøgelsen for at se, hvilken betydning datapunkterne for regionerne i EU10 og EU2 har på produktivitetsmålene og dets komponenter for EU15-landene. Derfor præsenterer jeg nedenfor en tabel, som viser målingerne for produktiviteten og dets komponenter for tidsperioden efter 1. udvidelse (år 2004-2009). Værdierne er præsenteret for datasættet på normalvis (hvor EU10 og EU2 er en del af data), samt hvor jeg udelader de to grupper.

²⁹ Jeg får en z-værdi på -0,34, og hypotesen accepteres, da værdien er numerisk mindre end 1,96.

³⁰ Jeg får en z-værdi på -2,45, og hypotesen forkastes, da værdien er numerisk større end 1,96.

Tabel 4.8: Gennemsnitlig årlig produktivitet og dets komponenter for EU15 hhv. med og uden EU10 og EU2

	Med EU10 og EU2			Uden EU10 og EU2		
År	Catch-up	Technical change	Malmquist Indeks	Catch-up	Technical change	Malmquist Indeks
2004	0,9988	1,0005	0,9993	1,0000	0,9991	0,9992
2005	0,9787	1,0097	0,9881	0,9777	1,0102	0,9877
2006	1,0270	0,9940	1,0208	1,0272	0,9940	1,0210
2007	0,9674	1,1109	1,0747	0,9728	1,1043	1,0743
2008	1,0182	0,9589	0,9764	1,0301	0,9495	0,9781
2009	0,9867	0,9366	0,9242	0,9719	0,9500	0,9233
Middel 2004-2006	1,0013	1,0014	1,0026	1,0014	1,0011	1,0025
Middel 2007-2009	0,9906	0,9993	0,9898	0,9912	0,9987	0,9900

Kilde: Egne beregninger fra data

Når jeg fjerner EU10 og EU2 fra datasættet, betyder det helt konkret, at jeg foretager en Malmquist-analyse udelukkende bestående af data fra landene i EU15 for perioden 2004-2009³¹. Dermed gøres det muligt at identificere sammenfald mellem EU10 og EU2's tilstedeværelse i datasættet og estimerne af produktiviteten for EU15-landene. Den reelle konsekvens ved at fjerne datapunkterne for regionerne i EU10 og EU2 for hvert enkelt år er, at randen i hvert år enten forbliver den samme eller svækkes. I tabellen ovenfor registreres nærmeste ingen ændring i hverken Catch-up effekten, randeffekten og Malmquist indekset for årene lige efter 1. udvidelse (år 2004-2006). Forskellen, på om EU10 er en del af data, er for alle 3 komponenter mindre end 0,2 %. I 2004, 2005 og 2006 er EU10 benchmark for hhv. 6 %, 6 % og 5 % af de samlede antal gange en region indgår som benchmark. Sammenkoblet med at en betydelig andel af de gange, hvor EU10 er benchmark, er for regionerne i netop EU10, betyder at EU10 regionerne i årene 2004-2006 ikke har den store betydning for den gennemsnitlige årlige produktivetsmåling i EU15.

I 2007, 2008 og 2009 observeres der markante sammenfald mellem EU10 og EU2's tilstedeværelse og de årlige værdier af hhv. Catch-up og Technical change for EU15-gruppen. De gennemsnitlige årlige ændringer i bunden af tabellen er dog for denne periode ikke ændret synderligt meget, men ud for de enkelte år ses der store differencer. De mest markante forskelle på fraværet af EU10 og EU2 findes i 2009, hvor randeffekten er 1,33 % større og Catch-up er 1,48 % mindre. EU10 og EU2

³¹ De lineære programmer ser ud som i bilag 3. Eneste forskel er, at det samlede antal observationer (n) er 100 (kun regionerne i EU15) uanset om et år indgår som basisår eller som det efterfølgende år.

udgør sammenlagt i 2008 og 2009 hhv. 26 % og 16 % af de gange en region indgår som benchmark, og en fremtrædende andel af disse gange er i begge år for regionerne i EU15. Effekterne i 2009 er dem der sker fra 2008 til 2009. Derfor må det faktum, at randeffekten i 2009 bliver større skyldes, at der var observationer fra EU10 og EU2 liggende på den efficiente rand i 2008, som gør randen dårligere end de efficiente regioner fra EU10 og EU2 gør i 2009.

At randen bliver stærkere samtidig med at Catch-up bliver mindre er langt fra nogen tilfældighed. Som tidligere skrevet er de to led modpoler til hinanden, hvilket medfører, at når randeffekten bliver stærkere, vil de inefficente regioner ligge længere væk fra den efficiente rand (det omvendte er i øvrigt tilfældet i 2008). I 2009 bliver Tyskland, Finland, Sverige og Storbritannien påvirket mest af, at EU10 og EU2 er taget ud af data. Finland er det land med størst tilbagegang i efficiens med 5,91 %, og faktisk også det land med den største fremgang i randeffekten (6,20 %). Samlet ses konsekvensen af at tage EU10 og EU2 ud af datasættet, at målingen af produktiviteten for Finland bliver 1 % større.

Ud fra ovenstående tabel ser jeg nogle individuelle effekter, men gennemsnitligt set betyder det altså meget lidt for EU15's produktivetsmål, om EU10 og EU2 er med eller ej. EU10 og EU2 bliver generelt ikke brugt ret meget som benchmarks for EU15 regionerne, og ifølge tabellen ændrer det ikke meget på målingen af produktivetsændringer, om de er der. Samlet må det fastslås, at de (regionerne i EU10 og EU2) fra et produktivets/relativt efficiens synspunkt ikke har den store betydning for EU15-landene. Dette betyder, at tilstedeværelsen af EU10 og EU2 i mit datasæt ikke har indvirkning på, at middelværdierne for produktiviteten (i de 3 tidsperioder) i EU15 er nedadgående.

5. Diskussion

I dette kapitel vil jeg sammenligne de resultater og tendenser, jeg registrerede i forrige afsnit med, hvad andre indenfor litteraturen er kommet frem til. I kapitlet vil jeg yderligere kommentere fordele og ulemper ved de anvendte analysemetoder, samt redegøre for ideer til fremtidige undersøgelser.

5.1. Sammenholdelse af resultaterne i min analyse med andres resultater

For at sætte resultater og tendenser i nærværende afhandling i en kontekst vil jeg sammenligne de resultater, som fremkom i forrige afsnit med resultaterne i artikler af hhv. Rao og Coelli (2003) og Rungsuriyawiboon og Lissitsa (2010). Rao og Coelli undersøger landbrugsproduktiviteten i 93 lande fordelt over hele verden i tidsperioden 1980-2000, hvorimod Rungsuriyawiboon og Lissitsa fokuserer på landbruget i 44 europæiske lande i tidsperioden 1992-2002. Jeg er klar over, at de respektive analyser er foretaget på baggrund af et andet variabelvalg og andre tidsperioder (og derfor andre datapunkter), men jeg vil alligevel undersøge hvilke tendenser og resultater, som er hhv. ens og forskellige ift. mine. I produktivitetssanalyser kan man i en vis udstrækning selv styre efficiensniveauet via antallet af variable, og f.eks. kan observationer kun blive mere efficiente ved tilføjelse af en ekstra variabel. Derfor kan man ikke lave en direkte sammenligning af efficiensscorer og produktivitetsmål, men derimod er en sammenligning af tendenser og indikationer på tværs af studier mulig. Det skal ydermere siges, at mine resultater er et gennemsnit over regionerne i et land, derfor er der ikke en direkte sammenlignelighed mellem mine resultater, og resultaterne i de to artikler jeg refererer til.

I Rao og Coellis artikel sammenholdes landbrugsproduktiviteten i forskellige kontinenter. Europa har i analysen et gennemsnitligt årligt Malmquist indeks på 1,4 %, hvilket kommer af en Catch-up effekt på 0,02 % og en randeffekt på 1,1 %. Den positive produktivitsudvikling ser jeg også i min analyse i tidsperioden 1996-2003, og derudover er randeffekten også den klart dominerende faktor. Danmark, Portugal og Holland er de EU-lande, som ifølge Rao og Coelli klarer sig bedst, og faktisk er disse lande også i min analyse blandt de lande med den største produktivitsudvikling. Desuden udviser Holland også i Rao og Coellis artikel perfekt efficiens i hele tidsperioden (og får altså en Catch-up effekt på 1). Den største forskel på resultaterne i Rao og Coellis analyse kontra min analyse er, at Storbritannien i deres analyse er præget af positive årlige produktivitsændringer, hvorimod Storbritannien i 1. tidsperiode i min analyse er blandt de lande med negative

produktivitetsændringer. En plausibel årsag til denne forskel kunne være, at udbredelsen af FMD og BSE i min analyse stort set præger hele 1. tidsperiode. Malmquist indekset for Storbritannien er i Rao og Coellis analyse taget over en meget større tidsperiode, hvorfor FMD og BSE ikke får så stor negativ indflydelse på den gennemsnitlige årlige produktivitet (perioden 1980-1994 er ikke påvirket af FMD og BSE).

Rungsuriyawiboon og Lissitsas analyse minder mere om min pga. deres hovedfokus på det europæiske landbrug. I analysen observeres nogle af de samme tendenser, som i Rao og Coellis artikel. Danmark er i Rungsuriyawiboon og Lissitsas analyse det land med den højeste gennemsnitlige årlige produktivitetsvækst (3,65 %), og Holland ligger på den efficiente rand i hele tidsperioden. Derudover er Danmark ikke placeret på den efficiente rand på noget tidspunkt i den tidsperiode deres datasættet strækker sig over. Grunden til at Danmark alligevel udviklingsmæssigt performer bedst er pga. en positiv efficiensforbedring, samt et meget dominerende positivt bidrag fra randeffekten. Den gennemsnitlige årlige TFP vækst er på 1,87 % for EU15-landene, og den dominerende bidragsyder er igen randeffekten. Endvidere har lande, som Belgien, Frankrig, Grækenland og Italien et Catch-up indeks på 1 i hele tidsperioden, hvilket er den største forskel ift. analysen i nærværende afhandling. I min analyse er det praktisk talt umuligt for Italien og Frankrig at ligge på den efficiente rand igennem hele tidsperioden, da jeg har angivet data på regionsform (Frankrig og Italien består af hhv. 22 og 21 regioner). F.eks. kræver det for Italien, at alle 21 regioner skal være efficiente i hele tidsperioden, hvilket praktisk talt er umuligt. Det betyder helt konkret, at de lande som består af færrest antal regioner i min analyse umiddelbart har nemmere ved samlet at være efficient (der tages selvfølgelig forbehold for, hvordan datapunkterne ser ud).

5.2. Fordele og ulemper ved de anvendte modeller

Der findes altid fordele og ulemper ved anvendelse af videnskabelige metoder. I dette afsnit redegør jeg for både fordele og ulemper ved hhv. DEA- og Malmquist-analysen, men vil hovedsageligt omtale DEA-analysen, da denne er hovedelement i beregningen af Malmquist indekset. Ift. andre effektivitetsanalyseværktøjer har DEA nogle klare fordele. (I) Under DEA er der ikke et behov for eksplicit at angive en matematisk form for produktionsfunktionen. (II) DEA metoden kan håndtere tilfælde, hvor enheder producerer flere forskellige typer af output og anvender flere forskellige typer af input. (III) Desuden er det under DEA-analysen ikke nødvendigt at tildele vægte til de forskellige input og output, som indgår i analysen. DEA metoden vælger selv vægtene på en sådan

måde, at den enkelte enhed stilles i det bedst mulige lys. Dette gøres ved at vælge vægtene således, at enhederne opnår den højst mulige efficiens ubb., at de valgte vægte ikke får nogle af de andre enheder til at have en efficiens, der er over 1³². (IV) Derudover stilles der ingen krav om kohærens mellem de anvendte variable, og disse er tilladt at have forskellige enheder. Derfor er valg af variable fuldstændigt overladt til den, som ønsker at foretage analysen. Det gælder om, at input indeholder alle ressourcer, som bruges til at producere, og output skal afspejle, det en observation producerer. Denne er i øvrigt også listet over ulemper, men jeg mener, at det primært er en fordel, at man kan konstruere sin egen model, da det giver en større frihed. Inden jeg går videre til næste punkt, skal det noteres, at korrelation mellem input eller mellem output ikke er godt, da variablene vil få en forstærkende effekt. Derimod forventes der implicit korrelation mellem valgte input og output, da en inputvariabel har indvirkning på, hvor meget output der produceres. (V) Ydermere er DEA særdeles brugbar, da de fundne benchmarks beskriver konkrete forbilleder i modsætning til en gennemsnitlig estimeret model, hvor de bagvedliggende enheders betydning er mere uigennemskuelig. DEA gør det herved nemmere at forholde sig til et benchmark, hvilket har stor betydning i en læringsproces. (VI) Derudover er det i DEA, ift. mange andre analysemetoder, nemmere at integrere bløde værdier (hvis de kan kvantificeres) i form af f.eks. miljøkrav. F.eks. kan en udledningsbegrænsning af kvælstofudvaskningen til grundvandet og det omgivende vandmiljø (fra gødning) inden for landbruget være et miljøkrav, som i en produktivitsanalyse (DEA-analyse) kunne indgå som et undesirable (uønsket) output (Ball et. al., 2001). Hvordan man helt konkret håndterer uønskede output er prøvet besvaret i flere undersøgelser. Bl.a. konstruerer Yaisawarng og Klein (1994) en DEA model, som måler effekten af SO₂-kontrol på efficiensændringer i de amerikanske kulkraftværker i 1980'erne. I deres analyse antager de en svag bortskaffelsesmulighed³³ for de uønskede output, som bliver implementeret i DEA modellen. Derudover forsøger Färe et al. (2000) en indeks tilgang, som evaluerer en observations præstation vha. 2 modeller, én i hver fase. Den første fase beregner et indeks baseret på stigninger i de ønskelige output, og den anden fase baseres på fald i de uønskede output. Til slut bliver det endelige efficiensindeks beregnet ud fra de 2 indeks fundet i hver fase.

³² Denne formulering er faktisk for det duale problem (minimeringsproblem) til den outputorienterede DEA model, som jeg har valgt ikke at præsentere i denne afhandling.

³³ Svag bortskaffelsesmulighed for output indebærer, at hvis y kan produceres givet en input vektor, så kan βy også produceres, hvor $0 \leq \beta \leq 1$.

Men DEA har imidlertid også nogle svagheder. (I) Det effektiviseringspotentialer som DEA måler udtrykker kun potentialer målt relativt ift. til andre bedrífers effektivitet. Derfor afslører DEA kun de potentialer, som mindst en anden region allerede har fundet. Det betyder, at hvis den sande produktionsfunktion ligger uden for den estimerede, så finder DEA kun et undervurderet forbedringspotentialer. (II) Derudover beregner DEA hypotetiske enheder som lineare kombinationer af eksisterende enheder, som virkelige enheder sammenlignes med. Det betyder, at der antages at gælde stykkevis linear teknologi (eller med andre ord konveksitet), hvilket i visse tilfælde kan være problematisk. Det kan nemlig diskuteres, om det er rimeligt at antage, at forskellige teknologier (DMU'er) altid kan kombineres til en ny. I denne analyse kan regionerne være præget af forskellige landbrugsmetoder, hvorved det i nogle tilfælde kan være umuligt at opnå det benchmark (blanding af flere potentielle forbilleder), som DEA anviser. (III) Desuden kan man stille sig kritisk overfor antagelsen om fri bortskaffelse af input, som DEA modellen bygger på. Der kan opstå situationer, hvor for mange input kan hindre produktionen af output, f.eks. ved at have for mange medarbejdere således, at de "tramper hinanden over tærne". (IV) Resultaterne i en DEA-analyse er følsomme overfor udvælgelsen af input og output, hvorfor man skal være meget omhyggelig i udvælgelsesprocessen. I forlængelse af dette skal det noteres, at hvis der tages for mange variable med i modellen, bliver for mange observationer efficiente. Tager man for få variable med, kan det føre til, at analysen bliver for upræcis, og observationer får fejlagtigt lave efficiensscorer. Det kan f.eks. ske, hvis nogle af de ressourcer, som bruges til at producere, ikke tages med. (V) I en standard DEA-analyse antages, at alle enheder har samme mulighed for at udnytte teknologien. Dette krav om, at enhederne er sammenlignelige, medfører at benchmarks skal søges i hele mulighedsområdet. Dette er dog ikke altid tilfældet, da der kan være en række forskellige naturgivne eller lovbestemte begrænsninger, der gør, at en sådan sammenligning ikke vil være relevant. Det mest oplagte eksempel ift. denne afhandling er vejret, som ikke er ens for alle regioner, og det derved umiddelbart er nemmere for nogen regioner at få en bedre høst end andre. Rent praktisk er det dog muligt at korrigere for disse forskelle ved at lave restriktioner på hvilke benchmarks, som er mulige (Saraiya, 2005). Det kræver dog, at restriktionen er målbar (som f.eks. vejret). Ift. min analyse vil det betyde restriktioner på kun at tillade benchmark, som har samme eller dårligere vejrbetingelser, som den region der bliver evalueret. (VI) I DEA antages, at der ikke er støj, hvilket er en urealistisk fornægtelse af virkelighedens realiteter, hvor udeladte variable, målefejl, stokastiske chok og anden statistisk støj nærmere er reglen end undtagelsen. Tilstedeværelsen af målefejl og anden støj (f.eks. vejret og strejker) vil påvirke placeringen af

indhylningsoverfladen (og dermed efficiensscorerne), og ekstreme værdier (outliers) kan påvirke resultatet i en stor grad. DEA er deterministisk (ikke-stokastisk), hvilket betyder, at al statistisk støj dømmes til at repræsentere inefficiens. Samlet set kan antagelsen føre til, at både beliggenheden af randen samt andres placering ift. til randen bliver påvirket, da DEA forvirrer stokastisk støj med variationer i inefficiens. (VII) Derudover er det vanskeligt at lave en validering af en DEA-analyses resultater, hvilket normalt undersøges ved supplerende analyser. Som eksempel kan nævnes bootstrapping metoder, som i DEA kan bruges til at korrigere skævheder for efficiensscorer og estimerer konfidensintervaller for dem, hvis man anerkender, at data er underlagt tilfældig støj (Tziogkidis, 2012).

Til sidst i afsnittet vender jeg blikket imod Malmquist-analysen, som ift. andre TFP indeks (f.eks. Törnqvist, Laspeyres, Fisher og Paasche) har den fordel, at der ikke kræves informationer om priser på input og output eller antagelser om teknologi og opførsel for denne. Derudover er fordelene ved Malmquist-analysen, at analysen både er en kvantitativ vurdering af produktiviteten i en specificeret periode, og at denne produktivitetsændring kan nedbrydes for at se, hvor meget af denne der skyldes hhv. Catch-up effekt og gennemførelsen af nye teknologier. Det skal dog bemærkes, at gennemførelsen af nye teknologier også kan skyldes ændrede økonomiske forhold, eller i min analyse nok snarere vejret. Derfor vil der ved et skift i landbrugsvejret fra dårligt til gunstigt, kunne registreres en forbedret rand, og dermed en positiv randeffekt.

Den væsentligste ulempe ved Malmquist indekset er nødvendigheden af at beregne afstandsfunktioner, hvilket i denne afhandling løses vha. DEA. Derfor bygger Malmquist på teorien bag DEA i stedet for at være en selvstændig model. Nedenfor har jeg listet flere kritisable punkter ved Malmquist indekset. (I) Malmquist indekset brugt i denne afhandling er ikke transitivt (Coelli et. al., 1998), da Malmquist indekset ikke afhænger ikke af en fast basisperiode (den ændres fra år til år). Dette hænger sammen med, at komponenterne bliver målt ift. forskellige estimerede rande, hvorfor det ikke er muligt at udregne produktivitetsændringer på tværs af forskellige år. F.eks. er $MI(t, t + 2) \neq MI(t, t + 1) \cdot MI(t + 1, t + 2)$, hvilket betyder, at man ikke kan finde produktivitetsændringen fra periode t til $t + 2$, hvis man har produktivitetsændringerne fra hhv. t til $t + 1$ og $t + 1$ til $t + 2$. Dog skal det siges, at Malmquist indekset er transitivt, hvis der over hele tidsperioden ikke er teknologiske ændringer. (II) Malmquist (og DEA) er desuden sårbar overfor observationer, som et år pludselig klarer sig uhørt godt ift. andre år. Engangsvidundere vil i

Malmquist-analysen føre til, at randen bliver ændret, og de vil derfor også påvirke andre regioners produktivitetsscorer. Konsekvensen kan blive, at en sådan region ender som benchmark for mange andre regioner, selvom denne ikke er et realistisk forbillede. (III) Malmquist er en serie af øjebliksbilleder, hvor der ikke tages højde for tilpasning over tid. F.eks. kan det for nogle regioner give mening at være inefficent i en periode, fordi man er ved at flytte sig et sted hen. Dette kunne i en landbrugsanalyse f.eks. skyldes implementering af nye maskiner, arbejdsmetoder eller strategier.

5.3. Fremtidige undersøgelser

I forlængelse af min analyse er der et par elementer, som er oplagte at tage fat på i fremtidige undersøgelser. F.eks. kan man, hvis data er tilgængeligt for EU10 og EU2 i hele tidsperioden (dvs. i årene 1995-2003 for EU10 og årene 1995-2006 for EU2), undersøge, hvad det (produktivitetsmæssigt) har betydet for de nye lande at blive medlem af EU. Ydermere kan man, ved at udvide data (evt. tage 2010-2012 med i datasættet (hvis data er tilgængeligt)), undersøge, om den tilbagegang i produktivitet mine resultater viser efter 2. udvidelse fortsætter. Det kunne i forlængelse heraf være interessant at undersøge, om den globale økonomiske krise har indflydelse på denne tilbagegang, samt om der evt. er forskel på hvilke lande/landegrupper, som produktivitetsmæssigt rammes hårdest af krisen. I forbindelse med den globale krise blev det dyrere og sværere at låne penge, hvilket medførte, at landmændenes investeringer blev mindre (Wiborg, 2009). I mit datasæt er investeringerne en del af variabelen Økonomisk størrelse og ved at fjerne investeringernes bidrag til denne variabel, vil det være muligt at fastslå om investeringerne (og dermed også den økonomiske krises) har en effekt på produktivitetsestimaterne i min analyse. Dette kan lade sig gøre, da ingen af de andre variable umiddelbart indeholder elementer, som direkte kan forbindes med finanskrisen. Helt specifikt vil man skulle sammenligne de nye produktivitetsestimater (en model uden investeringerne) for EU15 med resultaterne i denne afhandling. Derved vil man kunne udskille effekten af den globale krise, samt identificere hvilke lande/landegrupper, der bliver påvirket mest.

Men hvad har den økonomiske krise betydet for landmændene? I 2008 begyndte den globale økonomiske krise, og i kølvandet på denne fulgte en bekymring over faldende jordpriser, samt et fald i efterspørgslen efter dyrere landbrugsprodukter, som medførte et skift til basisprodukter pga. et fald i husholdningernes disponible indkomst (Swinnen og Van Herek, 2009). Det blev dyrere og sværere for landmændene at låne penge (gældskrise), hvilket var en konsekvens af en lavere

egenkapital i landbruget, og at der var mindre kapital at låne ud af (Wiborg, 2009). Da færre landmænd fik rammer til investeringer, blev det endnu sværere for landmændene at satse på udvikling og implementering af nye teknologier, samt i det hele taget at låne til nødvendige investeringer (Saabye, 2011). Ydermere blev mange landmænd i bl.a. Danmark ramt af at have store lån i schweizerfranc (Jensen, 2011), da kursen på schweizerfranc i perioden efter 2007 steg voldsomt³⁴. Samlet set led den generelle likviditet i landbruget et knæk (Larsen, 2008), hvilket betød at lande, som generelt var præget af en konservativ udlånspolitik til landbruget, også var blandt de lande, som klarede sig bedst gennem krisen (Lind og Zobbe, 2012). I min analyse kunne der være sammenfald mellem krisen og produktiviteten i EU15-landene. I hvert fald ses der på Figur 4.4 markante fald i produktiviteten i år 2008 og 2009, hvilket sker efter, at både Malmquist indekset og randeffekten topper i 2007. For at sige noget mere generelt om den økonomiske krises effekt på landbrugssektoren vil det kræve, at der findes data for flere år end 2009. En analyse af den økonomiske krises effekt på landbruget vil kunne udføres, hvis data for udlån til landbrug kan fremskaffes. Hvis data er tilgængeligt enten for hvert land eller region for hvert år vil man kunne inddrage den økonomiske krise mere direkte. En sådan variabel vil i modellen være en *Environmental variable*, som er en variabel, beslutningstageren ikke har indflydelse på (Asmild og Hougaard, 2006). Da tallene for udlån i regioner/lande er så forskellige, vil man skulle omforme tallene for at kunne sammenligne dem. Dette kunne i givet fald ske ved en eller anden form for indeksering, så udlånsdata vil kunne sammenlignes ud fra hvilke niveauer de ligger i. Variablen vil kunne inkorporeres som en inputvariabel i en produktivitsanalyse.

En anden mulighed i forbindelse med fremtidige undersøgelser kunne være at se på sammensætningen af variable i modellen. Vil det give mening at ændre denne ved enten at påføre eller fjerne variable fra modellen? F.eks. kunne man se, om der modtages tilskud fra regeringer³⁵ eller lign. (f.eks. landbrugsstøtten) i hver region, og i hvilken grad det modtages. Tilskud vil i givet fald skulle indgå som en inputvariabel, og umiddelbart vil det faktum at en region modtager et tilskud kunne føre til overvurdering af dens efficiensscore i min analyse³⁶. I min model er den Økonomiske størrelse den mest utraditionelle variable at bruge i en landbrugsanalyse. Den Økonomiske størrelse bidrager faktisk til, at der opstår en dobbelt counting for variablene Land og

³⁴ Kilde: <http://www.nationalbanken.dk/>

³⁵ Landbrugssektoren er i nogle regioner stærkt prioriteret af regeringer, da den genererer en stor procentdel af bruttonationalproduktet, og er eksistensgrundlag/indtægtskilde for en stor del af en regions befolkning.

³⁶ Det kan selvfølgelig også være, at tilskuddet indgår som en del af det samlede cash flow, hvorved en landbrugsundersøgelse umiddelbart ikke vil blive påvirket.

Dyr, da disse både indgår som selvstændige komponenter i modellen, men også som en del af den Økonomiske størrelse. Det betyder, at Land og Dyr har en større indflydelse på efficiens- og produktivitetsscorer end de egentligt er berettiget. I analysen er effekten af denne dobbelt counting, at regioner med lave værdier af Land og Dyr fremstår mere efficiente (og produktive) end de burde, og regioner med høje værdier af Land og Dyr fremstår mere inefficente end de burde.

Selvom resultaterne i min analyse er ganske plausible og informative, er jeg ganske bevidst om de begrænsninger data indeholder, og behovet for yderligere arbejde på dette område. Det fremtidige arbejde bør omfatte: (I) en undersøgelse af virkningerne af kunstvanding, forskelle i klima (nedbør) og opretholdelse af (kvantificerbare) miljøkrav. En undersøgelse af anvendelsen af kunstvanding kræver, at hver bedrift har tal for, hvor store ressourcer de bruger på kunstvanding. Kunstvanding vil i en landbrugsanalyse indgå som inputvariabel. Man kunne ved brug af DEA skabe en "best practice" rand baseret på de bedrifter, der producerer deres udbytte med den mindste mængde vand. Det vil her betyde at de landbrug, der er i stand til at producere deres outputniveauer ved brug af den mindste mængde vand er bedre vandforvaltere. Disse bedrifter vil fungere som benchmarks, hvor det vil være muligt for de inefficente bedrifter at sammenligne deres forbrug af kunstvanding. I en undersøgelse, hvor der tages højde for forskelle i klima, vil det kræve en viden om, hvilket klima hver bedrift er underlagt. Der vil skulle tages højde for temperaturer, perioder med sol og mængden af nedbør. Det vil kunne integreres i en produktivitetssanalyse, som en environmental variable, hvor der stilles restriktioner for hvilke benchmarks, der er sammenlignelige. Ved at bedrifter sammenlignes med bedrifter, som har samme klima, vil man få et mere realistisk billede af ens forbedringsmuligheder. En opretholdelse af et miljøkrav kræver, at udledningen kan kvantificeres på lande-, regions- eller bedriftsniveau alt afhængigt af, hvad kravet omhandler. Det kan, som tidligere skrevet, implementeres ved at bruge variabelen som et ønsket output. I en sådan model vil det være bedre for bedrifterne at producere mindre af det uønskede output. For miljøkravet vil det i en produktivitetssanalyse, kræve at der indsættes en grænseværdi for udledningen, hvor der ved overtrædelse af denne bliver tildelt en straf. (II) En analyse af kvaliteten af data (man kunne evt. se på indhentningen af data). En analyse af datakvaliteten kræver helt andre metoder end en produktivitetssanalyse, da kvaliteten af data ikke er en størrelse, som kan kvantificeres. For at kunne give et bud på hvordan kvaliteten kan forbedres, kræves en større baggrundsviden, om hvordan indsamlingen af data foretages i praksis. (III) En analyse af effekterne af fusioner og opkøb. En analyse af effekterne af fusioner og opkøb er i en landbrugsterminologi

interessant, da der bl.a. i Danmark er mange landmænd med store produktioner i østlandene³⁷. Grunden til at landmænd i stigende grad vælger at opkøbe jord eller fusionere i lande som Letland, Polen og Slovakiet er bl.a. pga. den billige landbrugsjord, billigere arbejdskraft, samt mere lempelige miljøkrav³⁸. For at analysere effekten af et opkøb eller en fusion i en produktivitetssanalyse er det nødvendigt, at data indeholder årene før og efter opkøbet/fusionen. Dermed vil det via en sammenligning i produktivitetsscorer for de involverede bedrifter være muligt at se, om det har været produktivitetsmæssigt rentabelt (Sufian og Habibullah, 2009). Helt præcist vil man skulle sammenligne efficiens gennemsnittet af de to bedrifter før fusionen/opkøbet med efficiensen efter. Det vil give mest mening at sammenligne efficiensen over tid efter fusionen/opkøbet, så man får et mere realistisk/stabilt billede af, hvordan produktionen efter fusionen/opkøbet er. Man skal dog have sig for øje, at der for produktionen efter fusionen/opkøbet vil skulle indgå et eller andet form for tids-lag, da der kan være forskel i hvilke forhold estimerne er bliver skabt under. (IV) En anvendelse af parametriske afstandsfunktioner til at intensivere undersøgelsen af robustheden i resultaterne. En sådan undersøgelse er bl.a. mulig ved de før omtalte bootstrapping metoder.

³⁷ Kilde: <http://www.business.dk/diverse/danske-landmaend-paa-fremmarch-i-oesteuropa>

³⁸ Kilde: <http://www.business.dk/diverse/danske-landmaend-paa-fremmarch-i-oesteuropa>

6. Konklusion

Denne afhandling indeholder konklusioner og resultater om niveauer og tendenser for landbrugsproduktionen i EU i tidsperioden 1995-2009 og til at foretage undersøgelsen, har jeg valgt efficiensanalysen DEA og Malmquists produktivitetsanalyse.

DEA bliver i denne afhandling brugt til to formål. Først foretager jeg en DEA-analyse på hvert år for sig for at klarlægge, hvordan en region i gennemsnit klarer sig ift. de andre regioner i samme år. Analysen viser, at både Holland, Danmark og Malta er perfekt efficiente, hvilket betyder, at der ifølge DEA ikke findes et forbedringspotentiale for disse. Ydermere giver de gennemsnitlige årlige efficienser for hvert land i denne analyse et billede af, hvordan Catch-up effekten bidrager til TFP væksten under Malmquist-analysen. Derfor kan jeg, ud fra de årlige efficiensscorer for EU15-gruppen (de gamle EU-lande), se, at EU15-gruppen oplever efficienstilbagegang i 1996, 1997, 1998, 2002, 2004, 2005, 2007 og 2009, samt fremgang i 1999, 2000, 2001, 2003, 2006 og 2008. I denne analyse bliver regionerne i Danmark og Holland (begge består dog kun af 1) i gennemsnit anvendt som benchmark flest gange om året, hvilket sker med en hyppighed på hhv. 16 og 15 gange. Det betyder, at en stor del af de andre regioner i EU bruger Danmark og Holland som pejlemærke for, hvor godt de performer på årsbasis.

Desuden foretager jeg en DEA-analyse på hele datasættet på en gang for direkte at kunne sammenligne efficiensscorer på tværs af årene. I undersøgelsen er Danmark, Holland og Malta de lande med de største gennemsnitlige årlige efficiensscorer. Det betyder, at landbrugene i Danmark, Holland og Malta, er dem som i gennemsnit præsterer bedst. Derudover viser denne analyse en generel opadgående efficiens udvikling for de gamle EU-lande fra 1995 til 2007. I 2007 topper den gennemsnitlige efficiens med en værdi på 0,81, hvilket betyder, at værdien af landbrugsproduktionen, ift. det anvendte inputniveau, set over hele tidsperioden var bedst i 2007. Det maksimale punkt i 2007 efterfølges af markante efficiensfald i både 2008 og 2009 på hhv. 2 og 6 %. I denne analyse er det regionerne i Danmark, Ungarn og Spanien, som total set bliver anvendt oftest som benchmark, og det sker med en hyppighed på hhv. 23, 12 og 10 om året.

Endvidere har jeg i denne afhandling anvendt Malmquists TFP indeks, som beregnes ud fra DEA's efficiensscorer, til at se på produktivitetsudviklingen i de 15 gamle EU-lande. I analysen undersøger

jeg ligeledes, om der er sammenfald mellem udvidelserne af EU i hhv. 2004 og 2007 og produktiviteten i de gamle EU-lande. Resultaterne for hele tidsperioden viser en gennemsnitlig årlig produktivitetsvækst i de gamle EU-lande på 0,31 %, hvilket udelukkende er drevet af en stigning i randeffekten på 0,62 %. I perioden før 1. udvidelse er TFP væksten på 0,82 %, selvom der observeres et efficienstab på 0,25 %. I perioden mellem 1. og 2. udvidelse stiger produktiviteten med 0,26 %, hvor Catch-up og Technical change indekset bidrager lige meget. I perioden efter 2. udvidelse registreres der i modsætning til de to andre tidsperioder en produktivitetsnedgang på 1,02 %, hvor et efficienstab på 1 % er den største bidragsyder. De gennemsnitlige årlige produktivitetsestimater for de 3 tidsperioder viser, at EU15-gruppen er blevet mindre produktiv i takt med, at EU er blevet større. Desuden ligger de årlige produktivitsudviklinger for EU15 meget stabilt i intervallet [-3 %; 3 %] helt hen til og med 2006. I 2007 er Malmquist indekset steget med 7,47 %, hvorefter der registreres markante produktivitsnedgange i 2008 og 2009.

I Malmquist-analysen er Danmark det land, som i den undersøgte tidsperiode har den største gennemsnitlige årlige produktivitsfremgang (2,22 %). Derudover viser analysen, at Portugal, Italien og Holland også har høje gennemsnitlige årlige produktivitsfremgange. Dog observeres der produktivitsnedgange for en del lande, og den største nedgang observeres for Irland (-3,33 %). Dvs., at selvom der registreres en samlet gennemsnitlig årlig produktivitsvækst i de gamle EU-lande, er der individuelle lande, som produktivitsmæssigt går tilbage.

I afhandlingen har jeg fundet produktivitscorer for EU15, hvor EU10 og EU2 ikke er en del af datamængden. Sammenligner jeg disse med de oprindelige produktivitscorer, kommer jeg frem til, at tilstedeværelsen af EU10 og EU2 fra et produktivits/relativt efficiens synspunkt ikke har nogen betydning for EU15-landene. Det betyder helt konkret, at de nye lande ikke i særlig høj grad er benchmarks for de gamle. Derfor ændres de efficiente rande ikke markant ved at fjerne EU10 og EU2 fra datamængden, dvs. randeffekten er uændret, og dermed også Catch-up. Men spørgsmålet er, om der er elementer, som jeg ikke har taget højde for i min model? Er der noget, min model ikke fanger? For umiddelbart vil EU10 og EU2 med lavere faktorpriser på f.eks. arbejdskraft og landbrugsjord kunne true de gamle EU-landes konkurrenceevne og produktionsvilkår (Babic, 2007). Ift. de anvendte variable i denne model kunne man argumentere for, om ikke inputvariablen Arbejdskraft (angivet i timer) skulle have været angivet som en prisvariabel (for at få de lavere faktorpriser inkorporeret i analysen). Arbejdskraften er billigere i EU10 og EU2, og da variablen er angivet i timer, fanger modellen ikke den omkostningsforskel, der ligger i anvendelsen af timer.

Forbruget af timer i EU10 og EU2 bliver i modellen håndteret på lige vilkår med EU15, hvilket ikke virker helt rimeligt. Desuden kan det tænkes, at de nye lande pga. de lave lønninger ikke substituerer fra arbejdskraft til maskiner, i samme omfang som i de gamle lande. Derfor vil det i en produktivitetssanalyse være endnu vigtigere for de nye lande, at variabelen Arbejdskraft angives som en prisvariabel, da der for disse lande er en grund til, at antallet af brugte arbejdstimer er højere end i de gamle lande. Konsekvensen kan i denne analyse være, at regionerne i EU10 og EU2 bliver klassificeret mere inefficiente end, hvis variabelen var angivet som en prisvariabel. Dette kan i sidste ende betyde, at regioner som burde ligge på randen ryger væk fra randen, og da det er de efficiente regioner i EU10 og EU2, som kan have indflydelse på produktivitetsestimaterne for EU15, vil det i givet fald føre til, at EU15-landene produktivitmæssigt ikke bliver påvirket af EU10 og EU2's regioner. Det samme kunne man overveje ift. variabelen Land, som også kunne være angivet på samme måde, dvs. som en priskomponent. Analogt med Arbejdskraften er 1 ha landbrugsjord omkostningsmæssigt ikke ens for alle regioner, men derimod meget varierende afhængigt af hvilken gruppe en region tilhører. Denne forskel er dog ikke kun udbredt mellem øst (EU10 og EU2) og vest (EU15), men i det hele taget er højere priser på landbrugsjorden et udtryk for kvalitet. Kvaliteten ligger i at bedre jorder, også kan give højere udbytte. I min analyse angiver jeg variabelen Land ud fra arealet og ikke som en priskomponent, og derfor tager modellen ikke højde for denne forskel.

En anden mulig problemstilling ift. EU10 og EU2's manglende betydning kunne være, om optagelsen i EU i virkeligheden også har gjort det nemmere for EU10 og EU2 at få adgang til hele EU markedet, eller om der stadig er en usynlig barriere? For hvis der stadig er en begrænsning vil den officielle udvidelse ikke have haft nogen indflydelse for konkurrencen i de gamle lande. Derimod vil et frit EU markedet for EU10 og EU2 være med til at forøge det samlede udbud af landbrugsvarer, og hvis forøgelsen i udbuddet er større end forøgelsen i den samlede efterspørgsel vil den større konkurrence kunne presse priserne. Det vil i sidste ende kunne føre til, at produktivitetsestimaterne for EU15-landene bliver mindre.

Samlet set giver denne afhandling et opdateret billede af landbrugsproduktiviteten i EU-landene, hvor det for hvert land er muligt at sammenholde deres regioners produktivitetssudvikling med produktivitetssudviklingen i andre EU-lande. Resultaterne af både DEA- og Malmquist-analyserne viser, at Holland og Danmark er blandt de lande, som efficiens- og produktivitmæssigt klarer sig

bedst. Derudover viser Malmquist-analysen, at de gamle EU-lande, over hele tidsperioden 1995-2009, oplever en gennemsnitlig produktivitetsfremgang.

7. Litteraturliste

7.1. Bøger og artikler

Asmild, M., Hougaard, J.L.; "Economic and environmental efficiency of Danish pig farms"; Agricultural Economics; Vol. 35, s. 171–181; 2006.

Babic, B. (Internationalt Kontor); "EU-udvidelser – status og fremtid"; i Danmarks Nationalbank Kvartalsoversigt, 1. kvartal; 2007.

Ball, V.E., Färe, R., Grosskopf, S., Nehring, R.; "Productivity of the U.S. Agricultural Sector: The Case of Undesirable Outputs"; s. 541 - 586; i Hulten, C.R., Dean, E.R., Harper, M.J.; "New Developments in Productivity Analysis"; University of Chicago Press; 2001.

Banker, R.D., Charnes, A., Cooper, W.W.; "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis"; Management Science; Vol. 30, No. 9, s. 1078-1092; 1984.

Bramsen, J.M., Nielsen, K.; "Interaktiv benchmarking Med eksempler fra landbruget"; Fødevareøkonomisk Institut; Rapport nr. 172; 2004.

Caves, D.W., Christensen, L.R., Diewert, W.E.; "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity"; Econometrica; Vol. 50(6), s.1393-1414; 1982b.

Charnes, A., Cooper, W.W., Rhodes, E.; "Measuring the efficiency of decision making units"; European Journal of Operational Research; Vol. 2, s. 429-444; 1978.

Coelli, T., Rao, D.S.P., Battese, G.E.; "An introduction to efficiency and productivity analysis"; Kluwer Academic Publishers; 1998.

Farrell, M. J.; "The Measurement of Productive Efficiency"; Part of Journal of the Royal Statistical Society. Series A (General); Vol. 120(3), s. 253-290; 1957.

Fulginiti, L.E., Perrin, R.K.; "Ldc Agriculture: Nonparametric Malmquist Productivity Indexes"; Journal of Development Economics; Vol. 53(2), s. 373-390; 1997.

Färe, R., Grosskopf, S., Hernandez-Sancho, F.; "Environmental Performance: An Index Number Approach"; Resource and Energy Economics; Vol. 26(4), s. 343-352; 2000.

Färe, R., Grosskopf, S., Logan, J.; "The Relative Efficiency of Illinois Electric Utilities"; Resources and Energy; Vol. 5(4), s. 349-367; 1983.

Grifell-Tatjé, E., Lovell, C.A.K.; "A note on the Malmquist productivity index"; Economics Letters; Vol. 47(2), s. 169-175; 1995.

Grosskopf, S., Färe, R., Norris, M., Zhang, Z.; "Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries"; The American Economic Review; Vol. 84(1), s. 66-83; 1994.

Hayami, Y., Ruttan, V.W.; "Agricultural Productivity Differences among Countries"; The American Economic Review; Vol. 60(5), s. 895-911; 1970.

Konkurrence- og Forbrugerstyrelsen; "Bilag til Benchmarking af vandsektoren 2011 - Redegørelse for benchmarkingmodeller med fokus på metode og udbredelse"; Forsyningssekretariatet; 2010.

Lau, L.T., Yotopoulos, P.A.; "The Metaproduction Function Approach to Technological Change in World Agriculture"; Journal of Development Economics; Vol. 31, s. 241-269; 1989.

Lind, K.M., Zobbe, H.; "Dansk landbrug og fødevarerindustriens konkurrenceevne og rammevilkår - sammendrag og konklusioner"; Fødevarerøkonomisk Institut; Rapport nr. 210; 2012.

Malmquist, S.; "Index numbers and indifference surfaces"; Trabajos de Estadística; Vol. 4(2), s. 209-242; 1953.

Portela, M.C.S., Thanassoulis, E., Despić, O.; "Data Envelopment Analysis: The Mathematical Programming Approach to Efficiency Analysis"; i Fried, H.O., Lovell, C.A.K., Schmidt, S.S.; "The Measurement of Productive Efficiency and Productivity Change"; Oxford University Press Scholarship Online; 2008.

Rao, D.S.P., Coelli, T.; "Total Factor Productivity Growth in Agriculture: A Malmquist Index Analysis of 93 Countries, 1980-2000"; CEPA Working Paper Series; No. 2; 2003.

Rungsuriyawiboon, S., Lissitsa, A.; "Total factor productivity growth in European agriculture"; <http://www.nesdb.go.th/econsocial/macro/tnce/Download/1/supawat.pdf>; 2010.

Saraiya, D.; "The Impact of Environmental Variables in Efficiency Analysis: A fuzzy clustering-DEA Approach"; Virginia Polytechnic Institute and State University; 2005.

Sufian, F., Habibullah, M.S.; "Do mergers and acquisitions leads to a higher technical and scale efficiency? A counter evidence from Malaysia"; African Journal of Business Management; Vol. 3(8), s. 340-349; 2009.

Swinnen, J.F.M., Van Herek, K.; "The Impact of the Global Economic and Financial Crisis on Food Security and the Agricultural Sector of Eastern Europe and Central Asia"; Executive summary of a report commissioned by the International Labour Organization (ILO); 2009.

Thanassoulis, E.; "Introduction to the Theory and Application of Data Envelopment Analysis A foundation text with integrated software"; Springer; 2001.

Tziogkidis, P.; "Bootstrap DEA and Hypothesis Testing"; Cardiff Economics Working Papers; Working Paper No. E2012/18; 2012.

USDA; "Agricultural Outlook / August 2001"; Economic research service; <http://www.usda.gov/wps/portal/usda/usdahome>; 2001.

Wiborg, T. (Funktionsleder, Kompetencecenter Landbrug - Jyske Bank); "Finanskrisens konsekvenser for landbruget"; Plantekongres; 2009.

Yaisawarng, S., Klein, J.D.; "The effects of sulfur dioxide controls on productivity change in the US electric power industry"; The Review of Economics and Statistics; Vol. 76(3), s. 447-460; 1994.

7.2. Interviews

Frandsen, T.W. (Senior makroanalytiker hos Jyske Bank); <http://finans.tv2.dk/nyheder/article.php/id-33532543:her-er-europas-h%C3%A5rdest-ramte-%C3%B8konomier.html>; 2010.

Jensen, S.J. (Chefanalytiker i Finanshuset Demetra); http://borsen.dk/nyheder/investor/artikel/1/198581/danske_milliardtab_paa_laen_i_schweizerfranc.html; 2011.

Larsen, L.H. (Analysechef i brancheorganisationen Dansk Landbrug); <http://borsen.dk/nyheder/avisen/artikel/12/3071182/artikel.html>; 2008.

Saabye, L. (Direktør i erhvervspolitisk afdeling, Landbrug og Fødevarer); http://www.dr.dk/DR2/deadline2230/Talknuserne/Landbrug_krise.htm; 2011.

7.3. Websites

Berlingske business: <http://www.business.dk/>

Børsen: <http://borsen.dk/>

Databasen FADN: http://ec.europa.eu/agriculture/rca/database/database_en.cfm

Danmarksradio: <http://www.dr.dk/>

EU-oplysningen: <http://www.euo.dk/>

FN's Fødevare- og Landbrugsorganisation (FAO): http://www.fao.org/index_en.htm

Nationalbanken: <http://www.nationalbanken.dk/>

Statistik (Mann-Whitneys U test): <http://statnoter.biolyt.dk/index.php?pageID=81>

TV2: <http://tv2.dk/>

USA's Institut for Landbrug: <http://usda.gov/wps/portal/usda/usdahome>

8. Figurliste

Figur 2.1: Transformation af input til output	12
Figur 2.2: Produktionsteknologi for ét input og ét output	14
Figur 2.3: Udledt mulighedsområde for ét input og ét output	17
Figur 2.4: Farrells input- og outputorienterede efficiensindeks	18
Figur 2.5: Grafisk eksempel på den outputorienterede DEA model	23
Figur 2.6: Referencepunkter for de inefficente DMU'er	24
Figur 2.7: Den efficiente rand i periode 1 og 2	35
Figur 2.8: Malmquists produktivitetsindeks for DMU B	37
Figur 3.1: Kort over udviklingen i EU	43
Figur 3.2: Variable i modellen	47
Figur 3.3: Udviklingen i antal bedrifter per år i EU15-landene	54
Figur 4.1: Det gennemsnitlige antal gange et lands regioner er benchmark per år (hvert år for sig)	61
Figur 4.2: Udviklingen i gennemsnitlig efficiens for EU15	67
Figur 4.3: Det samlede antal gange et lands regioner er benchmark per år (hele datasættet på en gang)	69
Figur 4.4: Malmquist Indekset og dets komponenter for EU15	80

9. Tabelliste

Tabel 2.1: Data for DEA eksempel	22
Tabel 2.2: Resultater for DEA eksempel	26
Tabel 2.3: Data for Malmquist eksempel	34
Tabel 2.4: Resultater for Malmquist eksempel	39
Tabel 3.1: Medlemslande i EU inklusiv deres optagelses år	42
Tabel 3.2: Antal observationer i datasættet per år	49
Tabel 3.3: Antal regioner/observationer i hvert land	50
Tabel 3.4: Beskrivende statistik for input- og outputvariablene for hele datasættet	51
Tabel 3.5: Middelværdi og standardafvigelse for gruppen EU15 angivet pr. år pr. bedrift for 3 tidsperioder	53
Tabel 4.1: Gennemsnitlige årlige efficiensscorer på landebasis, 1995-2009 (hvert år for sig)	57
Tabel 4.2: Gennemsnitlige årlige efficiensscorer på landebasis, 1995-2009 (hele datasættet på en gang)	64
Tabel 4.3: Gennemsnitlig årlig produktivitet og dets komponenter for lande tilhørende EU15 i tidsperioden 1996-2003	72
Tabel 4.4: Gennemsnitlig årlig produktivitet og dets komponenter for lande tilhørende EU15 i tidsperioden 2004-2006	74
Tabel 4.5: Gennemsnitlig årlig produktivitet og dets komponenter for lande tilhørende EU15 i tidsperioden 2007-2009	75
Tabel 4.6: Gennemsnitlig årlig produktivitet, samt produktivitetsprofil for hvert land i EU15 i %	77
Tabel 4.7: Gennemsnitlig årlig produktivitet og dets komponenter for EU15	79
Tabel 4.8: Gennemsnitlig årlig produktivitet og dets komponenter for EU15 hhv. med og uden EU10 og EU2	82

10. Bilag

10.1. Bilag 1 - De lineære programmer for DEA-analysen på hvert år for sig

Gældende for perioden 1995-2003:

$$\max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ubb.

$$\sum_{k=1}^n \lambda^k x_i^k \leq x_i^{k_0}, i = 1, \dots, 6$$

$$\sum_{k=1}^n \lambda^k y_j^k \geq \phi y_j^{k_0}, j = 1, 2$$

$$\lambda^k \geq 0, k = 1, \dots, 100$$

Gældende for perioden 2004-2006:

$$\max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ubb.

$$\sum_{k=1}^n \lambda^k x_i^k \leq x_i^{k_0}, i = 1, \dots, 6$$

$$\sum_{k=1}^n \lambda^k y_j^k \geq \phi y_j^{k_0}, j = 1, 2$$

$$\lambda^k \geq 0, k = 1, \dots, 119$$

Gældende for perioden 2007-2009:

$$\max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ubb.

$$\sum_{k=1}^n \lambda^k x_i^k \leq x_i^{k_0}, i = 1, \dots, 6$$

$$\sum_{k=1}^n \lambda^k y_j^k \geq \phi y_j^{k_0}, j = 1, 2$$

$$\lambda^k \geq 0, k = 1, \dots, 133$$

10.2. Bilag 2 - Det lineære program for DEA-analysen på det poolede datasæt

Gældende for perioden 1995-2009:

$$\max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ub.

$$\sum_{k=1}^n \lambda^k x_i^k \leq x_i^{k_0}, i = 1, \dots, 6$$

$$\sum_{k=1}^n \lambda^k y_j^k \geq \phi y_j^{k_0}, j = 1, 2$$

$$\lambda^k \geq 0, k = 1, \dots, 1656$$

10.3. Bilag 3 - De lineære programmer for Malmquist-analysen på EU15-landene

Generelt for alle nedenstående programmer gælder, at der kun evalueres de samme 100 observationer (regionerne i EU15-landene) hvert år, dvs. $k_0 = 1, \dots, 100$.

Gældende for perioden 1995-2009:

(1)

$$[d_t^{k_0}(y_t, x_t)]^{-1} = \max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ub.

$$\sum_{k=1}^n \lambda_t^k x_{i,t}^k \leq x_{i,t}^{k_0}, i = 1, \dots, 6$$

$$\sum_{k=1}^n \lambda_t^k y_{j,t}^k \geq \phi y_{j,t}^{k_0}, j = 1, \dots, 2$$

$$\lambda_t^k \geq 0, k = 1, \dots, n$$

(2)

$$[d_s^{k_0}(y_s, x_s)]^{-1} = \max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ub.

$$\sum_{k=1}^n \lambda_s^k x_{i,s}^k \leq x_{i,s}^{k_0}, i = 1, \dots, 6$$

$$\sum_{k=1}^n \lambda_s^k y_{j,s}^k \geq \phi y_{j,s}^{k_0}, j = 1, \dots, 2$$

$$\lambda_s^k \geq 0, k = 1, \dots, n$$

(3)

$$[d_t^{k_0}(y_s, x_s)]^{-1} = \max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ubb.

$$\sum_{k=1}^n \lambda_t^k x_{i,t}^k \leq x_{i,s}^{k_0}, i = 1, \dots, 6$$

$$\sum_{k=1}^n \lambda_t^k y_{j,t}^k \geq \phi y_{j,s}^{k_0}, j = 1, \dots, 2$$

$$\lambda_t^k \geq 0, k = 1, \dots, n$$

(4)

$$[d_s^{k_0}(y_t, x_t)]^{-1} = \max_{\phi^{k_0}, \lambda} \phi^{k_0}$$

ubb.

$$\sum_{k=1}^n \lambda_s^k x_{i,s}^k \leq x_{i,t}^{k_0}, i = 1, \dots, 6$$

$$\sum_{k=1}^n \lambda_s^k y_{j,s}^k \geq \phi y_{j,t}^{k_0}, j = 1, \dots, 2$$

$$\lambda_s^k \geq 0, k = 1, \dots, n$$

I ovenstående ligningssystemer antager variablen n forskellige værdier alt afhængigt af hvilke år som evalueres. n er i (1) og (3) tilhørende periode t og i (2) og (4) tilhørende basisperioden (periode s). I nedenstående skema er angivet, hvilken værdi n antager alt afhængigt af, om det pågældende år indgår som basisår eller som den efterfølgende periode:

n	<i>Periode s (Basisår)</i>	<i>Periode t</i>
100	1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003	1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003
119	2004, 2005, 2006	2004, 2005, 2006
133	2007, 2008	2007, 2008, 2009

10.4. Bilag 4 - Mann -Whitney-testen

Mann-Whitneys U test følger følgende antagelser:

- (i) Stikprøverne følger ikke en specifik fordeling
- (ii) Stikprøverne har den samme form
- (iii) Stikprøverne er uafhængige af hinanden.

I testen foretages en fælles rangordning af observationer i to tidsperioder, hvor regionen med den højeste efficiens får rang lig 1. Hvis der er to regioner med lige stor efficiens deler de rangeringen.

De to teststørrelser for de to perioder er defineret som:

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

I ligningerne er n_1 og n_2 antallet af observationer i de to stikprøver, og R_1 og R_2 er summen af rangordningen for de to grupper.

Teststørrelsen er givet ud fra formlen: $U = \min(U_1, U_2)$

Den kritiske værdi som bliver skabt af teststørrelsen, er da:

$$z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

Da stikprøverne i min analyse er relativt store skal denne z -værdi sammenlignes med standardnormalfordelingen, hvor acceptintervallet for et signifikansniveau på 0,05 er $-1,96 < z < 1,96$. Dvs. hvis z ligger indenfor dette interval accepteres nulhypotesen, hvis ikke forkastes den.