

Bilag 7 – Analyse af alternative statistiske modeller til DEA

Dette bilag er en kort beskrivelse af Forsyningssekretariatets valg af DEA-modellen.

| | |
|---|----------|
| INDLEDNING | 3 |
| SDEA | 3 |
| Fastsættelse af effektiviseringspotentialer igennem netvolumenmålet | 3 |
| Et alternativ til netvolumenmålet | 4 |
| SDEA vs. DEA | 4 |
| COLS | 4 |
| Fastsættelse af effektiviseringspotentialer igennem netvolumenmålet | 4 |
| Et alternativ til netvolumenmålet | 7 |
| COLS vs. DEA | 8 |
| SFA | 8 |
| Fastsættelse af effektiviseringspotentialer igennem netvolumenmålet | 8 |
| SFA vs. DEA | 9 |
| KONKLUSION | 9 |

Indledning

Dette bilag er en kort beskrivelse af Forsyningssekretariatets valg af DEA-modellen som udgangspunkt for beregningen af forsyningernes effektiviseringspotentialer.

Forsyningssekretariatet havde indledningsvis i benchmarkingprocessen flere overvejelser omkring valget af benchmark-model. Der var til at begynde med fire modeller inde i overvejelserne: DEA-, SDEA-, COLS- og SFA-modellen. Forsyningssekretariatet har valgt at benytte en DEA-model til at vurdere forsyningernes effektivitet. Dette bilag har til formål at illustrere, hvorfor Forsyningssekretariatet har valgt at benytte DEA-modellen til benchmarking af vandsektoren frem for de andre modeller.

Bilaget forudsætter, at man har en nogenlunde forståelse af intuitionen i de fire modeller. Dette bilag indeholder dermed ikke en forklaring af teknikken i modellerne eller en generel forklaring af modellerne. Dette bilag har kun til formål at påpege de mest relevante problemstillinger ved SDEA-, COLS- og SFA-modellen set i forhold til både netvolumenmålet og den senere brug af DEA-modellen.

SDEA

SDEA-modellen fungerer efter samme princip som DEA-modellen. Der henvises derfor til bilag 4, hvor DEA-modellen beskrives. Forskellen på de to modeller er, at SDEA-modellen har et forsigtighedshensyn indbygget i beregningen af effektiviteten for de enkelte forsyninger. Det indbyggede forsigtighedshensyn har til formål at sikre, at tilfældige udsving i omkostningerne ikke tilskrives ineffektivitet. Forsigtighedshensynet beregnes ud fra gennemsnittet af omkostningerne for de mest effektive forsyninger set over en årrække. SDEA-modellen er derfor mere kompleks end den i forvejen komplicerede DEA-model og derfor mindre gennemskuelig.

Fastsættelse af effektiviseringspotentialer igennem netvolumenmålet

Der er to problemstillinger forbundet med SDEA-modellen.

For det første kræver SDEA-modellen, at vi har data for en vis årrække. Dette vil medføre, at forsyningerne skal indberette oplysninger om omkostningsoplysninger for flere år tilbage i tiden. Dette vil forøge forsyningernes administrative byrde væsentligt.

For det andet har Forsyningssekretariatet vurderet, at det er mest retvisende i første omgang af benchmarkingen, at Forsyningssekretariatet selv fastsætter et forsigtighedshensyn udenfor modellen i stedet for at lade data bestemme usikkerheden ved modellen. Forsigtighedshensynet, som Forsyningssekretariatet fastsætter, vil bestå i en vurdering af hvor usikkert et

grundlag benchmarkingen bliver foretaget på. Dermed vil forsyningssekretariatet primært vurdere usikkerheden omkring præcisionen i beregningen af omkostningsækvivalenterne og præcisionen i modellen som helhed, dvs. hvorvidt modellen indeholder en beskrivelse af de relevante forhold for forsyningernes omkostninger.

Et alternativ til netvolumenmålet

Det er ligesom i DEA-modellen en mulighed at lade alle costdrivere og de samlede omkostninger indgå i SDEA-modellen. På denne måde ville det udelukkende være SDEA-modellen, der bestemte den sammenhæng, der er imellem omkostninger og costdrivere. Problemet med denne metode er, ligesom med den almindelige DEA-model, at der kræves et lavt forhold imellem antallet af inputs og outputs og antallet af forsyninger. Hvis forholdet er for højt vil modellen overvurdere mange forsyningers effektivitet. Derfor er det ikke muligt at benytte SDEA-modellen til dette.

SDEA vs. DEA

SDEA-modellen har næsten de samme egenskaber som DEA-modellen. Den største forskel er det indbyggede forsigtighedshensyn. Som beskrevet ovenfor er det Forsyningssekretariatets vurdering, at dette forsigtighedshensyn ikke er en fordel i denne første omgang af benchmarkingen. SDEA-modellen er derfor ikke at foretrække frem for den almindelige DEA-model.

COLS

COLS-modellen fungerer efter parametriske og statistiske principper. En fordel ved COLS-modellen er netop i de statistiske metoder. Herved indgår hele datasættet i bestemmelsen af den form, der skal beskrive sammenhængen mellem omkostninger og costdrivere. Dette adskiller sig fra DEA-modellen, hvor det kun er de effektive virksomheder, der bestemmer sammenhængen mellem omkostninger og costdrivere. Ulempen ved COLS-modellen består omvendt i, at modellen er meget følsom overfor størrelsen af spredningen i datasættet.

Fastsættelse af effektiviseringspotentialer igennem netvolumenmålet

I datasæt med en stor spredning, er der ofte problemer med observationer der har stor indflydelse som f.eks. outliers, og der er større usikkerhed pga. den større variation i denne type datasæt. Dette problem er det samme, som Forsyningssekretariatet har haft i forbindelse med beregningen af omkostningsækvivalenterne.

Det er en mulighed at løse dette problem ved at transformere data på samme måde som Forsyningssekretariatet har gjort for flere af de omkostningsækvivalenter, der indgår i beregningen af netvolumenmålet. Da der allerede er foretaget transformationer i beregningen af netvolumenmålet, er det ikke meningsfyldt at foretage yderligere transformationer. I

vandsektoren er der forholdsvis stor forskel på de enkelte virksomheder, hvilket netop giver et datasæt med meget forskellige observationer og dermed en meget stor spredning.

Selve beregningen af effektiviseringspotentialet i COLS-modellen fungerer ved, at der først findes en gennemsnitlig sammenhæng i data jf. de lyserøde punkter i figur 1. Denne sammenhæng forskydes nedad i diagrammet, så den rammer det eller de forsyninger der passer ind i denne sammenhæng og har de laveste omkostninger. Denne linje kaldes COLS-linjen i figur 1. Denne linje udgør den effektive front i COLS-modellen og effektiviseringspotentialet for samtlige virksomheder beregnes ud fra deres individuelle afstand ned til linjen.

Til sammenligning finder DEA-modellen det laveste forhold mellem netvolumenmålet og omkostninger, og opstiller en front der går igennem dette punkt. Denne linje kaldes DEA-linjen i figur 1. På samme måde som ved COLS-modellen beregnes effektiviseringspotentialet som afstanden fra en forsyning og ned til fronten.

Der opstår dog et problem i COLS-modellens måde at beregne effektiviseringspotentialet på. Effektiviseringspotentialet beregnes i praksis ved at finde den forsyning, som ligger under de lyserøde punkter og har den længste afstand op til den linje, som de lyserøde punkter udgør. Afstanden som denne forsyning har op til linjen udgør grundlaget for effektiviseringspotentialet for alle andre forsyninger, uafhængigt af forsyningernes størrelse i forhold til hinanden.

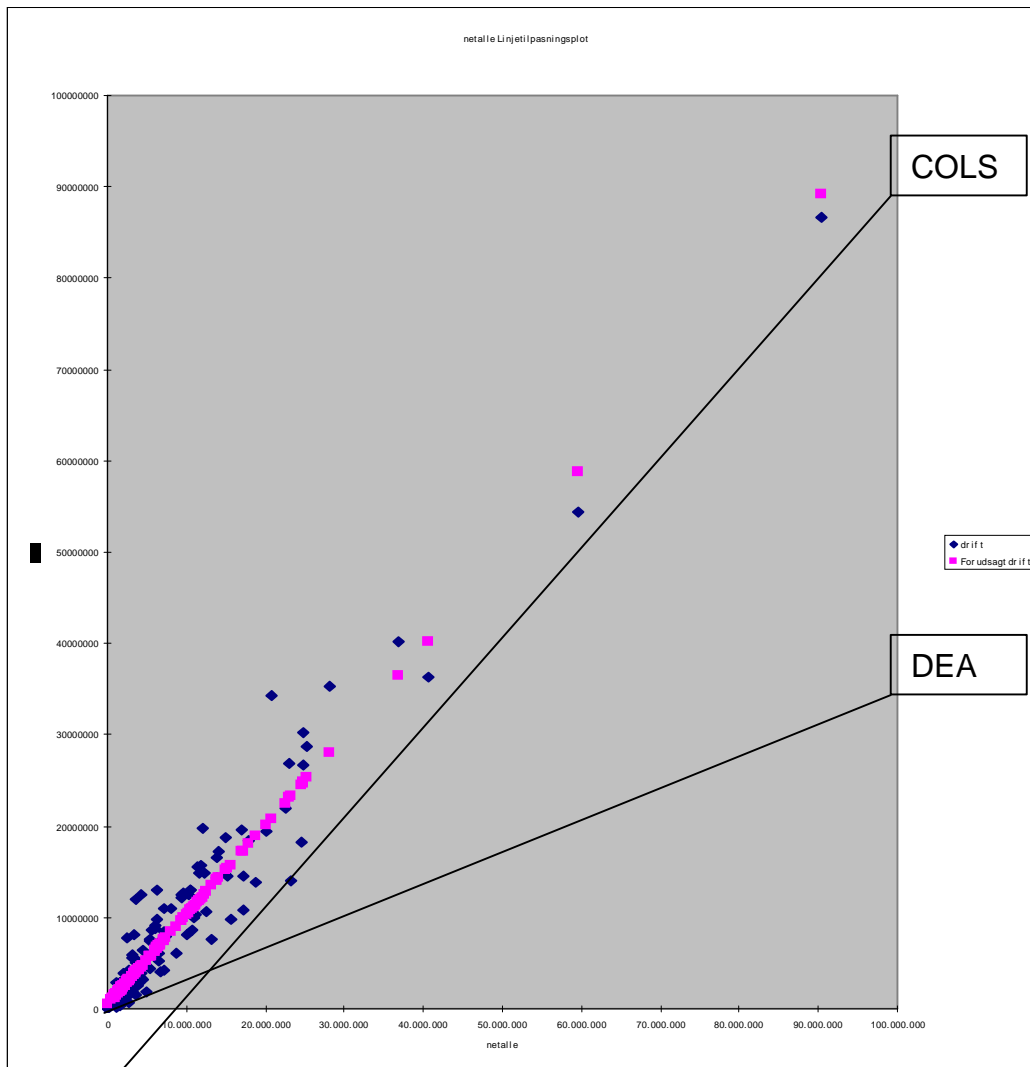
Som et eksempel kunne afstanden være 10 mio. kr. Det betyder, at COLS-modellen vil pålægge alle at forbedre sig 10 mio. kr. i forhold til den gennemsnitlige sammenhæng illustreret af de lyserøde punkter i figur 1. For en forsyning, der ligger over gennemsnittet, vil det betyde, at denne forsyning skal sænke sine omkostninger ned til gennemsnittet og dernæst med de 10 mio. kr., som er afstanden mellem gennemsnittet (de lyserøde punkter) og den bedste forsyning (COLS-linjen). Forsyninger, der ligger under gennemsnittet, kan reducere de 10 mio. kr. med det beløb som de er bedre end gennemsnittet.

Forsyningssekretariatet vurderer ikke, at dette er rimeligt overfor de mindre forsyninger, som ikke har omkostninger i denne størrelsesorden, og derfor ikke har muligheden for at lave en sådan effektivisering. Dette illustreres også af figur 1, hvoraf det fremgår at flere af de små forsyninger ville være nødt til at sænke deres driftsomkostninger til under nul kr., hvis de skal nå ned på den effektive front. Dette giver ikke mening i praksis.

Problemet med den store spredning i datasæt skyldes den store forskel mellem forsyningernes størrelse og dermed deres omkostninger. Den forholdsvis store variation i omkostningerne, for de store forsyninger, ender

med at straffe de mindre virksomheder i en COLS-model. Dette vurderer Forsyningssekretariatet som urimeligt.

Figur 1: Grafisk sammenligning af COLS-modellen og DEA-modellen



Figur 1 viser den effektive front for hhv. COLS- og DEA-modellen for de forsyninger der indberettede medio december 2010. Den vertikale akse beskriver de samlede driftsomkostninger og den horisontale akse beskriver netvolumenmålet.

Ovenstående illustrerer, at der er væsentlig forskel i resultaterne af beregningen af effektiviseringspotentialerne i hhv. DEA- og COLS-modellen. Ydermere er der stor forskel på forsyningernes indbyrdes placering i de to modeller. Dvs. der er stor forskel på, hvilke forsyninger der får høje og lave krav i de to modeller. Dette gælder for både vand- og spildevandsforsyninger jf. tabel 1.

Tabel 1: Væsentlige forskelle mellem DEA-modellen og COLS-modellen

| Gennemsnitligt effektiviseringspotentiale | | | |
|--|-------|------------------------|-------|
| Vandforsyninger | | Spildevandsforsyninger | |
| DEA | COLS | DEA | COLS |
| 73 % | 413 % | 50 % | 123 % |
| Gennemsnitlig ændring af placering | | | |
| Vandforsyninger | | Spildevandsforsyninger | |
| 76 | | 31 | |
| Antal i alt | | Antal i alt | |
| 177 | | 90 | |

Tabel 1 viser, at COLS-modellen beregner de gennemsnitlige effektiviseringspotentialer til over 100 % for både vand- og spildevandsforsyningerne. Dette giver ikke mening, idet en forsyning ikke kan have negative driftsomkostninger.

Desuden viser tabel 1 den gennemsnitlige ændring i placering, der er ved at skifte fra DEA-modellen til COLS-modellen. Den gennemsnitlige ændring i placering er for vandforsyninger 76. Det vil sige, en forsyning kan i DEA-modellen være den 3. bedste forsyning, mens COLS-modellen kan vurdere at der, for samme forsyning, er 79 forsyninger der er mere effektive. Forsyningen vil derfor få et lavt krav i DEA-modellen, men et højt krav i COLS-modellen.

Et alternativ til netvolumenmålet

COLS-modellen kan også bruges til at vægte de enkelte parametres betydning for omkostningerne. Dvs. modellen kan afgøre, hvor store omkostninger, der er forbundet med de enkelte costdrivere. COLS-modellen kunne derfor være brugt i stedet for at beregne et netvolumenmål, således at de samlede omkostninger og alle costdriver-oplysningerne kunne indgå i en COLS-model.

Forsyningssekretariatet vurderer ikke, at det har noget formål at benytte COLS-modellen til dette. Dette skyldes den store korrelation eller sammenhæng, der er mellem de forskellige costdrivere. Denne sammenhæng kan forklares ved, at der er en tendens til, at hvis en forsyning pumper meget vand op af sine borer, pumper det også meget vand ud af sit vandværk, har mange kilometer ledning og dermed også høje omkostninger etc.

Den høje sammenhæng mellem de forskellige costdrivere udelukker, at alle disse forhold kan indgå i den samme COLS-model. Som en konsekvens ville det medføre, at modellen bliver noget mere simpel og dermed mindre præcis end netvolumenmålet.

Netvolumenmålet er baseret på en kombination af, at forsyningerne selv har foretaget en grov fordeling af deres omkostninger ud på de enkelte costdrivere, og at Forsyningssekretariatet dernæst har benyttet statistiske

metoder, til at lave en finere opdeling af omkostningerne i form af omkostningsækvivalenterne.

Forsyningssekretariatet vurderer, at denne tilgang til omkostningsfordelingen er mere robust og præcis, da forsyningernes egen vurdering af omkostningsfordelingen indgår samtidig med, at det er muligt at lave en forholdsvis præcis model ved brug af de statistiske metoder.

COLS vs. DEA

De fordele der er forbundet med COLS-modellen udnytter Forsyningssekretariatet allerede i beregningen af netvolumenmålet. Dermed kan COLS-modellen ikke bidrage med noget yderligere her. Derudover gør følsomheden overfor den store spredning i datasættet, jf. figur 1 og tabel 1, at det er praktisk umuligt at benytte COLS-modellen til benchmarking i vandsektoren.

På baggrund af ovenstående vurderer Forsyningssekretariatet ikke, at der er nogen umiddelbare fordele ved at benytte COLS-modellen i stedet for DEA-modellen til benchmarking af vandsektoren. Forsyningssekretariatet har derfor fravalgt COLS-modellen til fordel for DEA-modellen.

SFA

SFA-modellen fungerer stort set efter samme principper som COLS-modellen. SFA-modellen har dog, ligesom SDEA-modellen, et indbygget forsigtighedshensyn. Dette hensyn afhænger af den indbyrdes variation forsyningerne imellem, i modsætning til SDEA-modellen, hvor forsigtighedshensynet afhænger af variationen over tid for de enkelte forsyninger på den effektive front.

Forsigtighedshensynet i beregningen af effektiviseringspotentialerne baseres på den usikkerhed og tendens der er i data. Effektiviseringspotentialerne i en SFA-model svarer dermed til effektiviseringspotentialerne i en COLS-model med en efterfølgende reduktion i effektiviseringspotentialerne. Denne reduktion er beregnet på baggrund af variationen eller usikkerheden i data og model. Selve beregningen af effektiviseringspotentialerne antager, at hele potentialet, kaldet e , kan opdeles i to dele, kaldet v og u , dvs. effektiviseringspotentiale $e = v + u$. Her betegner v en tilfældig variation i omkostningerne som følge af hændelige uheld som f.eks. ugunstige vejrforhold eller lignende. Som modsætning beskriver u den del af variationen, der skyldes ineffektivitet. Størrelsen på u for hver forsyning er afgørende for den pågældende forsynings effektivitet, mens v er ens for alle.

Fastsættelse af effektiviseringspotentialer igennem netvolumenmålet

Problemstillingerne og fordelene for SFA-modellen er stort set de samme som for COLS-modellen. I SFA-modellen er der dog store problemer med usikkerheden, idet modellen ikke kan kende forskel på den tilfældige

variation i data og den variation der skyldes ineffektivitet. Dette skyldes som før, at der er stor spredning i datasættet. Når de store forsyningers omkostninger, set i forhold til de mindre forsyningers omkostninger, svinger meget medfører det, at modellen ikke kan vurdere, hvad der kommer fra generelle fluktuationer/svingninger i omkostningerne og hvad der kommer fra ineffektivitet.

Dette illustrerer den primære ulempe ved især SFA-modellen, men også COLS-modellen. Problemet er, at datasættet er nødt til at være forholdsvis ensartet for at SFA- eller COLS-modellen kan benyttes. Dette er ikke tilfældet i vandsektoren, hvor den store forskel i størrelse skaber et forholdsvis spredt og uens datasæt. Resultaterne af SFA-modellen anvendt på forsyningernes omkostninger for hhv. vand- og spildevandsforsyningerne overfor netvolumenmålet illustrerer dette problem jf. tabel 2.

Tabel 2: Variation i mål af effektivitet i SFA-modellen

| Spildevandsforsyninger | | |
|-------------------------------|------------------|------------------|
| Variation på "e" | Variation på "v" | Variation på "u" |
| 8.51e+13 | 8.51e+13 | 1797 |
| Vandforsyninger | | |
| Variation på "e" | Variation på "v" | Variation på "u" |
| 1.46e+13 | 1.46e+13 | 54.65 |

Tabel 2 viser, at variationen på det samlede effektiviseringspotentiale e svarer til variationen på den tilfældige del af effektiviseringspotentialet v . Variationen på den del af effektiviseringspotentialet der skyldes ineffektivitet, u , er meget lille. Modellen kan derfor ikke finde nogen nævneværdig variation som kan tilskrives ineffektivitet.

Resultaterne af SFA-analysen skyldes som nævnt den store forskel i variation på forsyningernes omkostninger. Dette problem kunne løses med forskellige transformationer af data, men Forsyningssekretariatet har allerede foretaget en del transformationer i forbindelse med beregningen af netvolumenmålet, og det ville derfor ikke være intuitivt at transformere netvolumenmålet yderligere.

SFA vs. DEA

Det er ikke muligt at beregne effektivitetskrav i SFA-modellen, idet modellen ikke kan skelne mellem den variation, der skyldes uforudsete hændelser og den variation, der skyldes ineffektivitet. Det er derfor at foretrække at benytte DEA-modellen frem for SFA-modellen.

Konklusion

DEA-modellen er klart at foretrække frem for SDEA-, COLS- og SFA-modellen.

Branchen har selv givet udtryk for, at Forsyningssekretariatet selv skal foretage individuelle hensyn og generelle usikkerhedshensyn frem for at lade benchmarkingmodellen og data afgøre dette. Derfor foretrækker Forsyningssekretariatet DEA-modellen frem for SDEA-modellen.

COLS- og SFA-modellen er ikke egnede til brug for forsyninger med store forskelle i størrelse og andre karakteristika. Yderligere er de statistiske fordele ved disse modeller allerede udnyttet af Forsyningssekretariatet i beregningen af netvolumemålet. Forsyningssekretariatet foretrækker derfor DEA-modellen frem for COLS- og SFA-modellen.