

Arbeidsnotat
3/2008

**Analyseopplegg for å
kunne måle om
reorganisering av
skatteetaten fører til en
mer effektiv ressursbruk**

Finn R. Førstund
Sverre A.C. Kittelsen



*Stiftelsen Frichsenteret for samfunnsøkonomisk forskning
Ragnar Frisch Centre for Economic Research*

Analyseopplegg for å kunne måle om reorganisering av skatteetaten fører til en mer effektiv ressursbruk

Finn R. Førsund
Sverre A.C. Kittelsen

Sammendrag: Arbeidsnotatet er utarbeidet på oppdrag for Skattedirektoratet som et analyseopplegg som kan gi svar på spørsmålet om reorganiseringen av skatteetaten (ROS reformen) har ført til mer effektiv ressursbruk. Notatet går gjennom begrepsbruk og metodisk grunnlag for å kunne måle produktivitetsendringer under og etter reorganiseringen og for hvordan en kan isolere effekten av reformen fra andre endringer som kan påvirke produktivitetsutviklingen. Det legges vekt på metoder som kan gi statistisk utsagnskraft, og metoder som kan anslå om det finnes spesialiseringsgevinster i sektoren. Tilgjengelige data diskuteres, og det legges spesielt vekt på muligheten for å komme anslå produktivitetsutvikling for skatteetaten som helhet når det ikke er samme kontorene/enhetene som finnes før og etter reformen.

Nøkkelord: Produktivitet, Malmquistindeks, Data Envelopment Analysis (DEA), Reformevaluering, Skatteetaten

Kontakt: f.r.forsund@econ.uio.no, www.frisch.uio.no

Forord

Arbeidsnotatet er utarbeidet som svar på konsulentbistand/avrop på rammeavtale 20050036 angående utarbeidelse av analyseopplegg som kan gi svar på spørsmålet om reorganiseringen av skatteetaten (ROS reformen) har ført til mer effektiv ressursbruk. Tilbudet fra Frischsenteret sendt 17/10 om utarbeidelse av et analyseopplegg, ble akseptert av Skattedirektoratet i et innstillingsbrev av 29/10.

Skattedirektoratet peker bl.a. på følgende punkter som analyseopplegget skal skissere metode og databehov for:

- Opplegg for å kunne måle produktivitetsendringer under og etter reorganiseringen og hvordan isolere reformeffekten
- Effekter av spesialisering
- En totalvurdering av produktivitet under og etter reorganiseringen

Innhold

Forord.....	ii
Innhold	iii
Oppsummering.....	iv
1. Bakgrunn.....	1
2. ROS – reformen	3
3. Effektivitet og produktivitet.....	4
4. Analysemetode.....	5
5. Måling av produktivitetsendringer før, under og etter reformen	15
6. Effekter av spesialisering.....	21
7. Produktivitet for skattekontorene samlet	23
Referanser	26

Oppsummering

I samfunnsøkonomiske analyser defineres *produktivitet* som forholdet mellom produksjon og ressursbruk, og *effektivitet* som forholdet mellom faktisk produktivitet og en norm for best mulig produktivitet gitt de begrensningene en virksomhet driver innenfor. Samlet effektivitet i samfunnsøkonomisk forstand kan deles mellom *ytre* og *indre* effektivitet, mellom å gjøre de riktige tingene og å gjøre tingene riktig. Ytre effektivitet er et spørsmål om å avveie verdien av en type offentlig tjeneste for brukerne opp mot verdien av en annen tjeneste, gitt at de samme ressursene medgår. I samfunnsøkonomisk analyser er det et sett av tekniske effektivitetsbegreper som står sentralt. Da det ikke alltid er så lett å finne fram i litteraturen om definisjonene, blir disse gjennomgått forholdsvis grundig. Det sentrale utgangspunktet for empirisk tallfesting er beste observerte praksis. Dette leder til å ta i bruk en transformasjon av ressurser til tjenester som bygger på enheter som er mest effektive. Den mest populære metode, DEA (Data Envelopment Analysis), bygger på en innhylling av datapunkter uten å forutsette en parametrisering av en frontfunksjon. Hvert enkelt kontor som er med som enhet, vil få bestemt sin effektivitetsfaktor. Beregning av produktivitet på kontornivå bygger på å måle relativ effektivitetsutvikling over tid.

Når det gjelder opplegg for å kunne måle produktivitetsendringer under og etter reorganiseringen trengs det data for minst to år. Beregning av produktivitet vil bygge på DEA – metoden. For å kunne ta hensyn til skjevheter som kan komme inn pga. et begrenset antall kontorer vil en spesiell statistisk teknikk, bootstrapping, bli tatt i bruk. Denne metoden gir oss konfidensintervall for både effektivitets- og produktivitetstall for de enkelte kontorer.

Reformeffekten kan isoleres ved å bruke en frontfunksjon fra perioder før reformen, og så se hvor de nye kontorenehetene ligger i forhold til den gamle teknologien. Reformeffekten kan også forsøkes isoleres ved å beregne korrelasjon mellom produktivitetstall etter ny teknologi og forklaringsfaktorer som tempoet i implementeringen av reformen for de enkelte kontorer, innføring av IT, regelendringer, og liknende faktorer.

Spesialiseringsgevinster i gjennomsnitt vil kunne identifiseres hvis det er slik at enheter med spesialisert produksjon gjennomgående ligger nærmere beste praksis – fronten enn andre kontorer. Det kan også testes statistisk (med en ikke – parametrisk test) om gruppen av spesialiserte enheter har høyere effektivitet eller produktivitet enn gjennomsnittet.

For å beregne produktivitet for alle kontorene under ett, både før og etter reformen, må en summere all ressursbruk og all tjenesteproduksjon for etaten for hver tidsperiode/år. Resultatet vil da kun være en observasjon per periode og det vil ikke være informasjon om spredning av produktivitetsutviklingen som kunne nyttiggjøres i å finne årsakssammenhenger mellom reform og produktivitet. En vil likevel kunne si noe om den samlede produktiviteten i alle perioder før og etter reformen dersom en kan løse problemet med sammenveining av tjenestetyper. En metode er å bruke vekter som reflekter relative kostnader målt ved arbeidskraftbruk ved de enkelte tjenester. En annen metode er å la vektene være beregnet av DEA - metoden på grunnlag av observasjonene for de enkelte kontorer, dvs. behandle de årlige observasjonene av sektoren som helhet som vanlige enheter, og så sammenligne produktivitetsutviklingen til disse med fronten som er tallfestet på grunnlag av alle enkeltkontorene.

1. Bakgrunn

Frischsenteret er for tiden engasjert til å evaluere NAV – reformen, og har ansvaret for 2 av de 7 modulene (delprosjektene) evalueringen er delt opp i. Dette evalueringsarbeidet har klare paralleller til evalueringen av ROS – reformen. Formålet med sammenslåingen av trygdekontorer, arbeidskontorer og deler av kommunale sosialkontorer er å få mer sysselsetting og velferd for den samme totale ressursbruk. I det ene delprosjektet, Modul 5, studeres kausale virkninger på sysselsetting og stønadsbehov. I Modul 7 har Frischsenteret ansvaret for å studere om reformen har bidratt til en mer effektiv arbeids- og velferdsforvaltning. Da det er et meget stort prosjekt å finne ut noe om velferdseffektene på kundesiden for hvert enkelt kontor, konsentreres evalueringen om sammenhengen mellom tjenesteproduksjonen og ressursinnsatsen. Dette er i prinsippet samme problemstilling som vi vil legge til grunn for evaluering av ROS – reformen. De overordnede mål for skatteetaten er å sørge for at innbetaling skatter og avgifter skjer i henhold til lover og regler. En studie av hvordan ROS – reformen virker på slike mål vil være av en annen type, og sannsynligvis falle svært mye dyrere, enn å studere virkninger på indre effektivitet, dvs. kostnadseffektiviteten til de tjenester som faktisk produseres.

Frischsenteret har også deltatt i evalueringen av Sykehusreformen av 2002 (Kittelsen et al., 2007). En angrepsvinkel som ble brukt under den evalueringen og også foreslått brukt under NAV – evalueringen, er å se på sammenhengene mellom de individuelle produktivitetstallene og forklaringsfaktorer knyttet til reformen og det tempoet de enkelte kontorer har tilpasset seg denne. Disse analysene er basert på produktivitetstall korrigert for skjevheter ved bruk av bootstrapping. Dernest er sammenhengen mellom produktivetsutviklingen og reformen søkt forklart i en regresjonsmodell der det også kontrolleres for andre faktorer som påvirker produktivitet. I analysen av sykehusreformen fant en positiv reformeffekt på ca 4-5%, identifisert som den produktivetsfremgangen norske sykehus hadde utover produktivetsfremgangen i de andre nordiske landene, alt annet likt.

I sykehusanalysen ble sykehus i de andre nordiske landene tatt med både for å fange opp produktivetsendringer som skyldes felles utviklingstrekk i de nordiske landene (medisinsk

teknologi, organisering, folkehelse, etc.), og for å øke antallet observasjoner. Et faglig problem med testing er at antallet observasjoner kan være for lite til å få signifikante resultater. En mulighet som bør undersøkes for ROS - reformen, er om andre nordiske land kan brukes som kontroller for en produktivitetsutvikling uten reform i likhet med hva som er gjort ved evaluering av sykehusreformen.

Frischsenteret har tidligere analysert effektivitet og produktivitetsutvikling til likningskontorer (Førsund og Lindseth (2004), Førsund et al. (2005), (2006)). Det var fastsettingsfunksjonen som var utgangspunktet for å definere tjenestetyper som det kunne finnes data for, uten å måtte foreta en spesiell ny innsamling. Antall tjenestetyper var 6, alle målt ved antall typer følgende saker: folkeregistrering, kontroll av folkeregistrering, likninger for lønnstakere og pensjonister, klager til nemnd/rettssaker (inkludert saker sendt til nemnd av likningskontorer), likninger for personlig næringsdrivende og likninger for selskaper (AS). Kun en ressurs ble spesifisert: totale utgifter i kroner. Resultatene for effektivitet og produktivitet viste en relativt stor spredning mellom likningskontorene med en tendens til at små kontorer hadde lavest effektivitet. Korrigering for utvalgsskjevhet (Førsund et al., 2006) ga et helt annet inntrykk av hvilke kontorer som var signifikant ineffektive (effektive) eller hadde negativ (positiv) produktivitetsvekst enn ukorrigerede tall. Analyseopplegget for ROS – evalueringen vil bygge på erfaringene vunnet ved disse analysene, og være en videreføring og videreutvikling av de tidligere studier.

Frischsenteret ønsker å basere analysen på faglig aksepterte metoder for samfunnsøkonomiske analyser. Spesielt ønsker vi at konklusjoner skal bygge på statistiske metoder og være etterprøvbare. Man må da være oppmerksom på at slike analysemåter nødvendigvis vil ha sine begrensninger når det gjelder å kunne trekke klare konklusjoner. Det er ikke sikkert at skarpe konklusjoner om effekten av en reform kan trekkes. Men det kan være av verdi i seg selv å få demonstrert at man på grunnlag av det tilgjengelige datamateriale eventuelt ikke kan trekke sikre konklusjoner, og at man kan få fram de faglige begrensningene som i prinsippet kan ligge i evalueringen. Det er dermed ikke sagt at verdifull innsikt om virkningen av reformen ikke kan skaffes på annen måte enn den vi vil foreslå, for eksempel ved å foreta et begrenset antall intervjuer med nøkkelpersoner. Men konklusjoner man kan trekke på et slikt eller likende

grunnlag, kan aldri bli mer enn tentative, og vil vanskelig aksepteres av utenforstående (for eksempel Finansdepartementet) som utsagnskraftige konklusjoner.

Vi vil først nevne de vesentlige endringer som ROS – reformen innebærer og deretter definere og forklare de sentrale begreper i en analyse av effektivitet og produktivitet før vi presenterer forslag til hvordan hovedspørsmålene stilt av Skattedirektoratet om evaluering av ROS – reformen kan besvares. Avslutningsvis vil det legges opp en tidsplan hvor de forskjellige trinn av analysen spesifiseres.

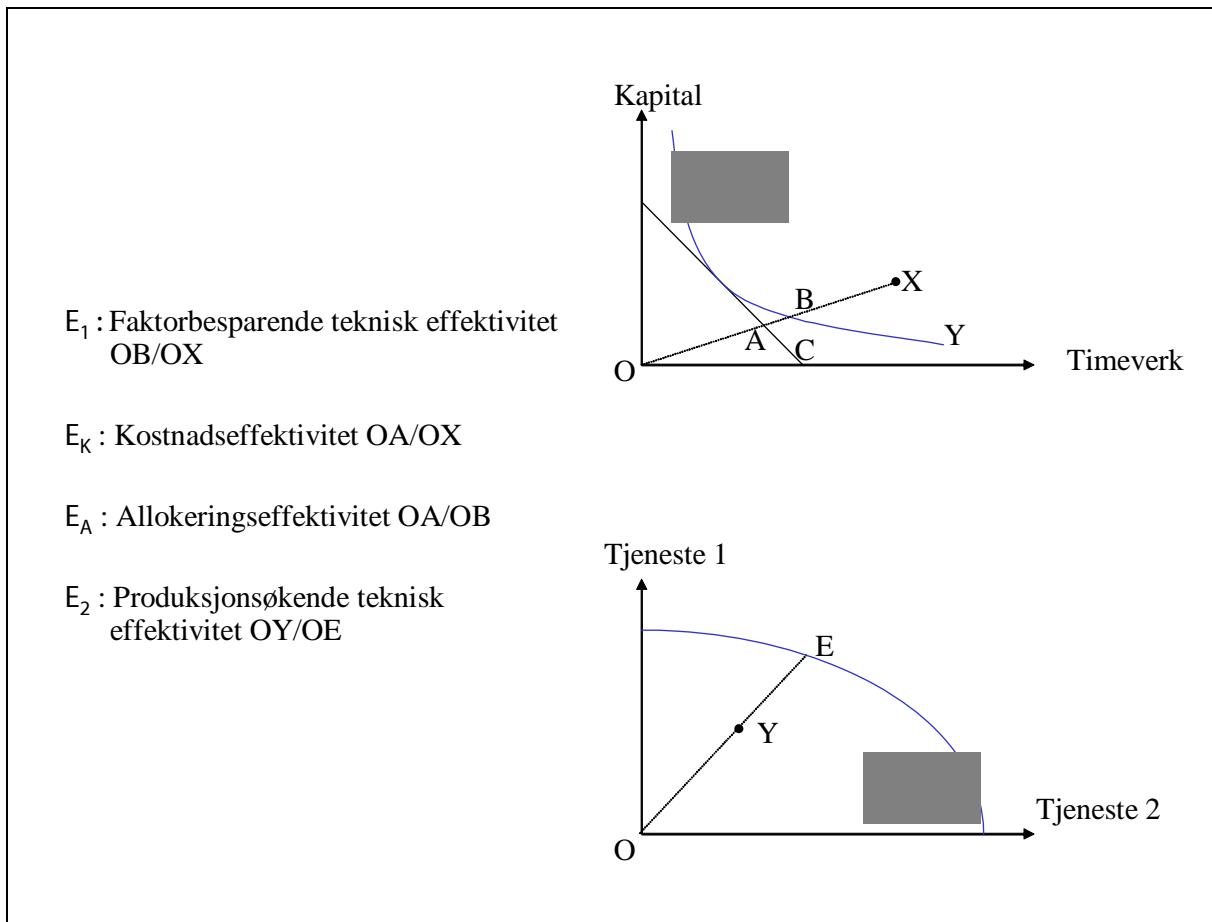
2. ROS – reformen

ROS – reformen innebærer at det innføres 5 regioner som hver er en formell enhet; Skattekontoret (se St. prp. Nr. 1, 2006–2007). I hver region vil det være mange fysiske kontorsteder som kan sammenliknes med tidligere likningskontorer, fylkesskattekontor og skattefogdkontor. Av spesiell interesse er det at de regionale skattekontorene får riksdekkende myndighet, slik at oppgaver til et skattekontor kan overføres til et kontor i en annen region. Dette gir en mulighet for en spesialiseringsgevinst. Arbeidsoppgavene til kontorstedene er delt inn i 5 aktivitetsområder; veiledning, fastsetting, kontroll og rettsanvendelse, skattekrim og innkreving. Disse aktivitetene vil være utgangspunktet for å definere tjenester som produseres. Av spesiell interesse for oss er at fastsetting og innkreving skal være operativt atskilt. Det antas i St prp. Nr.1, 2006-2007, med henvisning til en analyse (uten nærmere referanse), at ”omorganiseringen vil gi betydelige positive nytteeffekter både for borgere og for etaten. Nyttene vil blant annet bli økt kvalitet i saksbehandlingen, økt likebehandling, kortere saksbehandlingstider og økt service og veiledning til brukerne. Videre heter det ”Ny organisering vil også bidra til økt avdekking av skatteunndragelser og bedre etterlevelse av regelverket, og dermed til riktigere skatteinngang.” Vi vil komme tilbake til noen av disse punktene senere, men her bare peke på at i henhold til en intern brosjyre i Skattedirektoratet sies det at et av formålene med ROS – reformen er å bedre ressursanvendelsen i etaten, og å bedre kvalitet og effektivitet. Vi vil starte med å gi en presisering av de sentrale begrepene effektivitet og produktivitet.

3. Effektivitet og produktivitet

Begrepene produktivitet og effektivitet brukes ulikt mellom fagmiljøer, men også blant samfunnsøkonomer vil definisjonene variere. Med vårt utgangspunkt i mikroøkonomisk produksjonsteori vil vi definere *produktivitet* som forholdet mellom produksjon og ressursbruk, og *effektivitet* som forholdet mellom faktisk produktivitet og en norm for best mulig produktivitet gitt de begrensningene en virksomhet driver innenfor. Slike begrensninger kan være geografiske, juridiske, organisasjonsmessige e.l., og effektivitetsbegrepet kan derfor gis ulikt innhold ettersom en ser det på kort eller lang sikt, fra ulike aktørers ståsted osv.

Samlet effektivitet i samfunnsøkonomisk forstand kan deles mellom *ytre* og *indre* effektivitet, mellom å gjøre de riktige tingene og å gjøre tingene riktig. Ytre effektivitet er et spørsmål om å avveie verdien av en type offentlig tjeneste for brukerne opp mot verdien av en annen tjeneste, gitt at de samme ressursene medgår. I mangel av eksplisitte markedspriser for offentlig produserte tjenester vil slike avveininger kunne være et spørsmål om politiske prioriteringer mellom ulike formål, og ytre effektivitet vil også kalles *prioriteringseffektivitet*. Indre effektivitet er derimot et spørsmål om hvorvidt selve produksjonen foregår ved minst mulig bruk av ressurser, og kalles derfor også *produksjonsøkonomisk effektivitet*.



Figur 1. Farrells effektivitetsbegreper i innsatsfaktorrommet og i produktrommet

4. Analysemetode

Farrell's effektivitetsbegreper

Farrell (1957) foreslo et sett med effektivitetsbegreper til bruk i samfunnsøkonomiske analyser som løser problemet med manglende priser, gitt at en kjenner normen for det som er fysisk mulig å produsere ved en gitt ressursbruk, eller motsatt hva som er nødvendig ressursbruk for å oppnå en gitt produksjon. Denne normen er det vi kjenner som produktfunksjonen, eller fronten til produksjonsmulighetsområdet. Grunntanken kan illustreres ved hjelp av figur 1. I øverste panel i figur 1 er det tegnet kombinasjoner av to innsatsfaktorer, kapitalbruk og arbeidsinnsats målt i timeverk. Anta at en ineffektiv enhet produserer en bestemt mengde av sine tjenester ved bruk av

en kombinasjon av areal og timeverk som er representert ved punktet X i figuren. Skal en produsere denne mengden kunne en imidlertid ha klart seg med noe mindre av hver innsatsfaktor, men en kunne også ha brukt betydelig mer av hver ressurs uten at det hadde gått ut over tjenesteproduksjonen. I figuren er *mulighetsområdet* representert ved området til høyre og over kurven Y. Denne kurven (isokvanten) representerer fronten av mulighetsområdet, karakterisert ved at det ikke er mulig å redusere bruken av noen innsatsfaktor uten samtidig å produsere mindre enn kvantumet Y. Punktene på fronten er derfor teknisk effektive, mens alle tilpasninger inne i mulighetsområdet er teknisk ineffektive. Det er vanlig å anta at fronten krummer vekk fra aksekrysset, blant annet fordi det er vanskelig å tenke seg noen produksjon uten litt av hver av innsatsfaktorene. Farrells første mål for teknisk effektivitet (E_1) for denne enheten representert ved punktet X er den andel av ressursbruken som er nødvendig for å produsere den observerte mengden Y. En foretar en *proporsjonal reduksjon* av alle innsatsfaktorer (langs en rett linje fra X til aksekrysset O) så lenge en fortsatt befinner seg i mulighetsområdet. En tilpasning i punktet B har dermed samme relative sammensetning av innsatsfaktorene som i X, hver innsatsfaktor i B har samme prosentvise andel av faktorbruken i X. Effektivitetsmålet E_1 er dermed avstanden OB delt på avstanden OX, dvs nødvendig faktorbruk delt på observert faktorbruk, og kalles derfor *faktorbesparende teknisk effektivitet*. I likhet med de andre effektivitetsmålene vi skal drøfte, er dette et tall mellom 0 og 1, der verdien 1 (=100 prosent) tilsier at enheten er effektiv. Ofte uttrykkes resultatene i effektiviseringspotensialet utregnet som $1-E$, som uttrykker hvilken andel av ressursene en kunne ha spart dersom alle enhetene var effektive. Effektivitetsmål kan beregnes for hver enhet, eller aggregeres til grupper og sektoren samlet.

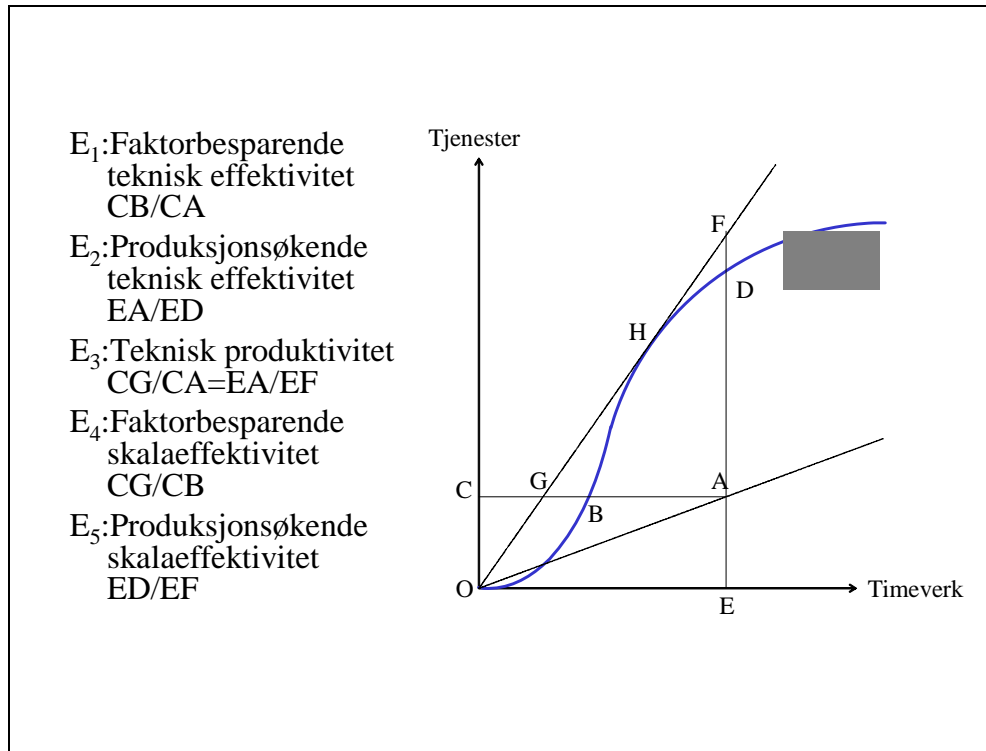
Dersom en kjenner prisene på innsatsfaktorene, vil forholdet mellom dem kunne representeres ved en kostnadslinje som C i figuren. Den ”riktige” sammensetningen av innsatsfaktorene vil da være den som gir lavest kostnader. I figuren er dette tangeringspunktet mellom isokvanten Y og kostnadslinjen C. Alle punkter på kostnadslinjen har samme total kostnad, slik at også punktet A representerer de minste nødvendige kostnadene. Dermed kan *kostnadseffektivitet* for enheten som produserer Y beregnes som avstanden OA delt på OX, dvs minste nødvendige kostnader delt på observerte kostnader. Forholdet mellom kostnadseffektivitet og teknisk effektivitet kalles *allokeringseffektivitet* og er den delen av kostnadseffektiviteten som skyldes ikke-optimal faktorsammensetning.

I nedre panel av figur 1 er den *produksjonsøkende tekniske effektiviteten* (E_2) illustrert. Her avgrenser fronten (den krumme kurven) de kombinasjoner av to produktmengder som det er mulig å produsere ved en gitt faktorbruk. Siden det for en gitt ressursbruk som regel er mulig å produsere færre tjenester enn det maksimale, er mulighetsområdet i dette rommet mellom fronten og aksene. Helningen på fronten tilsier at en effektiv virksomhet ikke kan øke antall av den ene tjenestetypen uten enten å øke ressursbruken eller ved å redusere antall av den andre tjenestetypen. Også den produksjonsøkende effektiviteten er et forholdstall mellom det mulige og det observerte ved en proporsjonal endring av alle produktmengder samtidig. Vi ser på en enhet som – for gitt faktorinnsats – produserer de to tjenestene i et omfang som tilsvarer punktet Y. I dette tilfelle er den produksjonsøkende effektiviteten gitt ved avstanden OY delt på OE, eller den observerte delt på den potensielle produksjonsmengden gitt at sammensetningen av produkter og nivået på ressursbruken ikke endrer seg. Farrell's effektivitetsbegreper kalles ofte radiale effektivitetsmål, fordi de i diagrammer som figur 1 vil måles langs rette linjer fra origo.

Det er ikke opplagt at det er mulig å oppnå samme produktivitet for en liten virksomhet som for en stor, eller omvendt, selv om virksomhetene er effektive. Det generelle tilfellet er tvert om at det finnes en *optimal størrelse* hvor produktiviteten på fronten er høyest. Figur 2 illustrerer forholdet mellom produktivitet og teknisk effektivitet i et forenklet diagram med bare ett produkt og en innsatsfaktor. I figur 2 er mulighetsområdet til høyre og nedenfor fronten, fordi en antar at det alltid er mulig å bruke mer av en faktor ved gitt produksjon, eller produsere mindre av en tjeneste ved gitt faktorbruk. Siden produktivitet er forholdet mellom produksjon og ressursbruk vil et produktivitetsnivå kunne tegnes som en rett linje fra aksekrysset i figur 2. Teknologien vil derimot kunne ha et forløp med stigende, konstant eller fallende skalautbytte. I figuren er det punktet H som har høyest produktivitet blant mulige tilpasninger på eller bak fronten, og en sier da at H har optimal skala (produktiviteten er lavere på begge sider av H). Hvorvidt den optimale skalaen er liten eller stor i forhold til et gjennomsnittlig nivå er et empirisk spørsmål.

En mulig normering av *teknisk produktivitet* (E_3) er å se produktiviteten til en enhet med tilpasning i punktet A i forhold til den høyest mulige produktiviteten i H, med andre ord

forholdet mellom helningene til de stiplede linjene OA og OH. Siden den faktorbesparende tekniske effektiviteten (E_1) i figur 2 er CB delt på CA, kan en definere den *faktorbesparende*



Figur 2. Produktivitet, teknisk effektivitet og skalaeffektivitet i faktor-produktrummet

skalaeffektiviteten (E_4) som forholdet mellom produktivitet og teknisk effektivitet (CG/CB) eller forholdet mellom faktorbruken ved optimal skala og ved en teknisk effektiv enhet med samme størrelse på produksjonen som enheten i punktet A. Tilsvarende kan produksjonsøkende skalaeffektivitetsmål defineres som forholdet mellom E_3 og E_2 . En viktig videreføring av Farrell's radiale effektivitetsbegrep er muligheten til å se på produktivitetsendring over tid.

Ofte er formålet med produksjonen knyttet til effekter eller resultater som kun indirekte påvirkes av de tjenestene eller varene som produseres. Det kan vise seg å være vanskelig å operasjonalisere produkter ut fra formålene med virksomhetene. Innen helsesektoren er det f.eks. bedre helse heller enn antall operasjoner som er det egentlige formålet. Har man gode mål for helseeffekter, knyttet f.eks. til funksjonsdyktighet, arbeidsmarkedstilknytning e.l., vil man

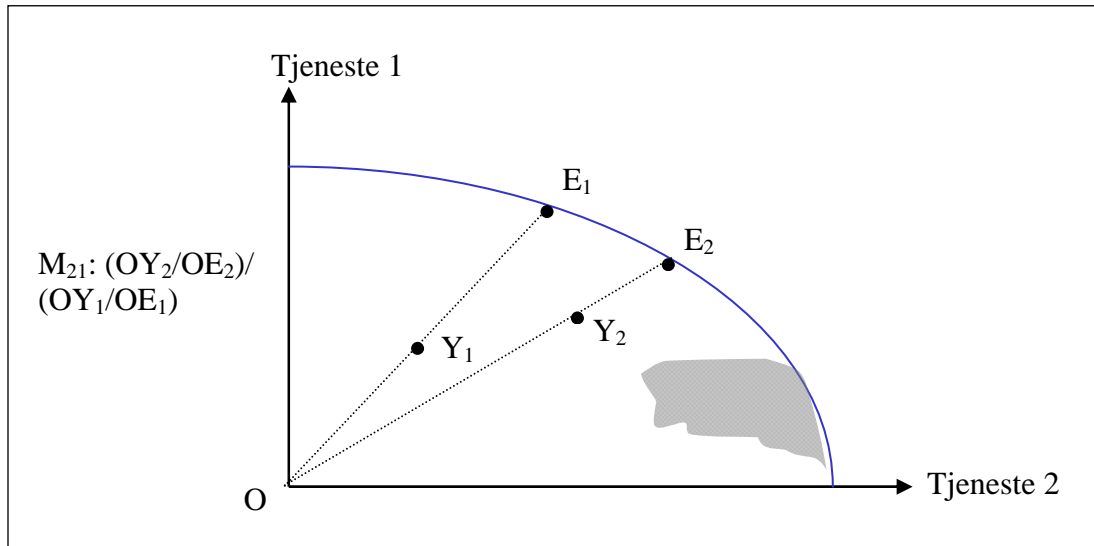
kunne bruke det produksjonsøkonomiske metodeverket til å estimere effektiviteten i måloppfyllelsen (resultateffektivitet eller "outcome efficiency"), gitt at en er i stand til å kontrollere for andre faktorer som også påvirker helsetilstanden. På tilsvarende måte som for skala og kostnader, kan en da også foreta en dekomponering av resultateffektiviteten i en del som skyldes effektiviteten i tjenesteproduksjonen og en del som skyldes effektiviteten i sammensetning av tjenester.

Produktivitet

Et hovedproblem ved måling av produktiviteten til en enhet oppstår når det er *flere* produkter eller tjenester og/eller *flere* innsatsfaktorer eller ressurstyper. For å finne et tallmessig uttrykk for forholdet mellom produksjon og ressursbruk må en veie sammen de ulike produktene og de ulike innsatsfaktorene. Etter samfunnsøkonomisk teori er det riktig å bruke markedsprisene til produktene og innsatsfaktorene som vektor dersom markedene tilfredsstillende betingelser om perfekt informasjon, mangel på markedsrett osv, slik at disse prisene reflekterer betalingsviljen for produktene og alternativverdien til innsatsfaktorene. Problemet er at skatteetaten, i likhet med det meste av offentlig sektor, leverer flere typer av tjenester som ikke omsettes i et (konkurrans)marked og som ikke har *prisinformasjon*. Til forskjell fra markedsbasert og konkurranseutsatt produksjon, får vi derfor ingen "tilbakemelding fra markedet" som gir en sortering av effektive og mindre effektive enheter alt etter overskudd og overlevelsessevne. Nettopp i slike tilfeller vil effektivitetsanalyser basert på produktfunksjoner som omfatter flere produkter, være et viktig hjelpemiddel for vurdering av "bransjens" interne effektivitet, organisering og ressursallokering.

Malmquist produktivitetsindeks (Caves et al., 1982) bruker endring i teknisk produktivitet (E_3) i forhold til en fast teknologi, og unngår derved det behovet for priser som andre produktivitetsindekser har. Malmquistindexen kan dekomponeres i produktivitetsendring som skyldes teknisk endring eller frontskift, og produktivitetsendring som skyldes endring i effektivitet. Videre dekomponering i skalaendring osv. er også mulig. Figur 3 viser konstruksjonen av Malmquist produktivitetsindeks med utgangspunkt i det nedre panelet i Figur 1 der produksjonsøkende effektivitet illustreres. Den krumme produksjonsmulighetskurven representerer nå den teknologien vi vil bruke ved sammenlikning av to perioder; periode 1 og

periode 2. Ressursmengden som brukes for å kunne være på fronten er gitt. Sammenlikning av produktiviteten til enheten målt ved Y_2 i periode 2 og Y_1 i periode 1 er forholdet mellom Farrell effektivitetsindekser. En indeksverdi større (mindre) enn 1 betyr produktivitetsfremgang (tilbakegang). Vi ser at i figuren så har enheten hatt en produktivitetsfremgang. Produktivitetsfremgangen kan kun skyldes forbedring i effektivitet da



Figur 3. Malmquist produktivitetsindeks

fronten ligger fast for de to perioder. For å kunne se produktivitetsendring av teknisk fremgang må forskjellige fronter for forskjellige perioder brukes.

Tallfesting av produksjonsmulighetene

Det er flere metoder til å tallfeste (estimere) den generelle frontfunksjonen som er vist i figurene på. Et problem er hvordan en anslår hva som er "best mulig", eller fronten av mulighetsområdet. Et enkelt utgangspunkt er å anta at "best mulig" er det samme som "best observert". Dette ligger til grunn for flere av de estimeringsmetodene som brukes i effektivitetsanalyser i dag, hvor de klart vanligste metodene er ulike varianter av DEA (*data envelopment analysis*, eller dataomhyllingsanalyse). Ofte beregnes en best praksis produktfunksjon i fysiske enheter, men kjenner en også faktorpriser er kostnadsfunksjonen lett å avlede. Det finnes flere alternative

metoder som kan være egnet for estimering av produktfunksjoner, se f.eks. Coelli, Rao og Battese (1998). Hovedinnfallsvinklene er vist i figur 4.

Panel a) i figur 4 viser den klassiske tilnærmingen til estimering av produktfunksjoner. Metoden bruker regresjon til å trekke en rett eller krummet funksjon (med forhåndsspesifisert parametriske form) gjennom observasjonssvermen. *Gjennomsnittsestimering* antar at all variasjon skyldes måle- eller utvalgsfeil, og dermed implisitt at variasjonen ikke skyldes teknisk ineffektivitet. I konkurranseutsatte næringer kan dette være en rimelig forutsetning, ettersom ineffektive virksomheter vil bli utkonkurrert med mindre det er innebygd teknisk fremgang. I offentlig sektor er det neppe en like god antagelse.

Panel b) i figuren viser den andre ytterligheten. Den estimerte gjennomsnittsfunksjonen er nå vippet opp til å omhylle alle enhetene, slik at all variasjon tilskrives ineffektivitet og ikke målefeil o.l. Denne metoden kalles *Deterministic Frontier Analysis* (DFA) og utgjør et lite brukt yttertilfelle, fordi den ikke tillater målefeil, samtidig som den parametriske strukturen gjør at den slutter mye dårligere om observasjonene enn DEA-metoden gjør (jfr panel d)). DFA gir de laveste anslagene på effektivitet, og vil generelt kun gi én eller noen få effektive enheter.

Panel c) viser hvordan en kan kombinere muligheten for både stokastikk og ineffektivitet i *Stochastic Frontier Analysis* (SFA). Fronten konstrueres slik at den ligger i overkant av gjennomsnittet i observasjonssvermen, men ikke kun hviler på den beste enheten. virksomhetens produksjon spesifiserer begrenset av en frontfunksjon som er stokastisk i den forstand at den tillates å variere fra virksomhet til virksomhet. Variasjonen skyldes dels forhold utenfor virksomhetenes innflytelse og dels målefeil. Problemet er imidlertid at det må gjøres sterke forutsetninger for å få skilt effektivitet og målefeil. Dersom en har adgang til paneldata, dvs. observasjon av de samme enhetene i flere tidsperioder, kan en komme noe bedre ut ved å anta at effektiviteten til de enkelte enhetene er konstant over tid, men forskjellig mellom enheter.

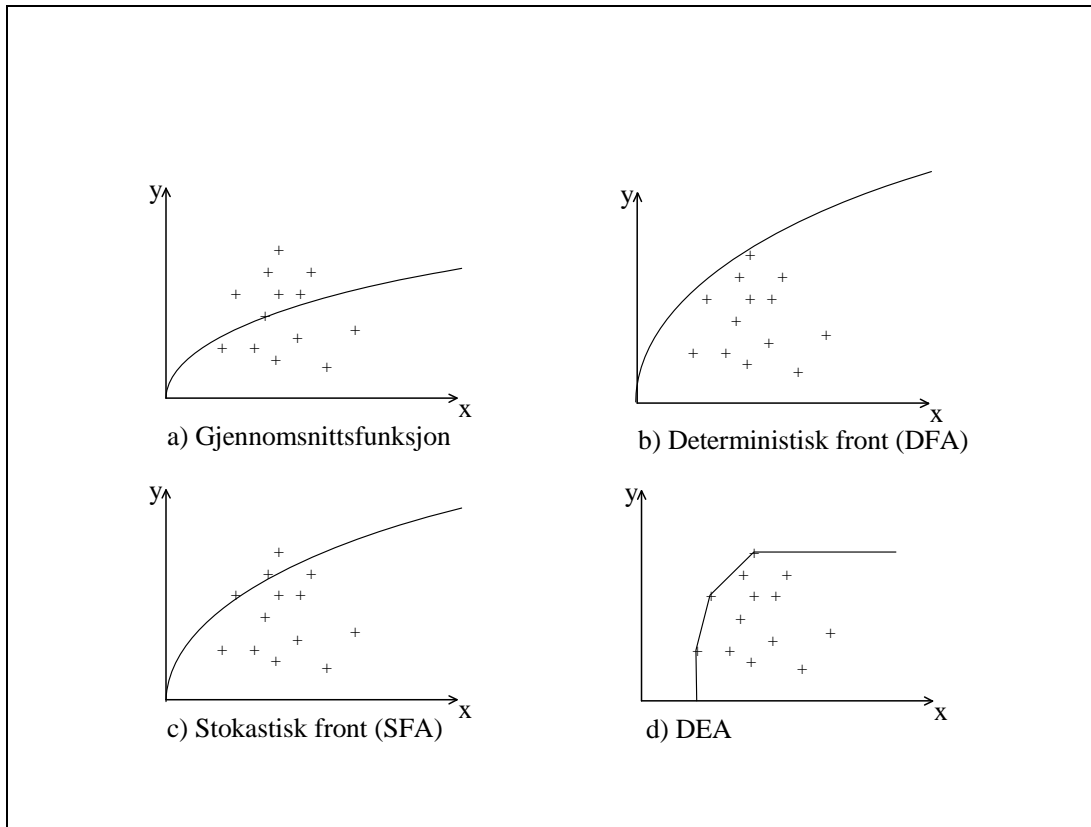
DEA er en ikke-parametriske metode som er illustrert i figur 4d. Igjen er det Farrell (1957) som har foreslått hovedideen med å omhylle observasjonene, mens Charnes et al. (1978) utviklet metoden videre og viste hvordan det generelle tilfellet med flere produkter og konstant

skalautbytte skulle settes opp og løses. Metoden er ikke-parametrisk i den forstand at det ikke pålegges noen bestemt parametrisk struktur på produktfunksjonen. Den bygger på de tre antagelsene a) at de observerte tilpasningene faktisk er mulige, b) *fri avhending*, dvs. at en alltid kan kvitte seg med produkter og innsatsfaktorer, og c) *konveksitet*, som er det samme som å si at dersom en har to mulige tilpasninger, vil også en tilpasning midt i mellom være mulig. I tillegg til effektivitets- og skalatall for hver enkelt enhet gir DEA-metoden også mulighet til å peke ut de enhetene som definerer fronten. Dette kan gjøres for hver enkelt ikke-effektiv enhet ved å navngi de ”referentene” eller ”*læremestrene*” som definerer fronten i det aktuelle området.

Metoden er blitt tatt mer og mer i bruk i offentlig sektor i Norge (se St. meld. nr. 1, Nasjonalbudsjettet 2008, s.168-172.). Sammenlikning av et kontor med beste praksis gir effektivitet, mens utviklingen av effektivitet over tid gir produktivitet.

Svakheten ved DEA-metoden er først og fremst at den antar fravær av målefeil. Imidlertid slutter den DEA-estimerte fronten så tett om observasjonene at det ikke nødvendigvis gir lavere effektivitetsestimater enn SFA, der en må spesifisere en krummet kurve (med forhåndsspesifisert parametrisk form) for frontfunksjonen. Tidligere var det også vanlig å hevde at DEA hadde en fordel fordi SFA ikke kunne håndtere mange produkter og innsatsfaktorer samtidig, mens SFA til gjengjeld hadde tilgjengelig et fullt sett av statistiske tester. Nyere forskning har i stor grad utvisket disse skillene, ved at det tatt i bruk parametriske distansefunksjoner som kan ha mange-dimensjonale produkt- og faktorvektorer, og at det er utviklet statistiske tester og bootstrap - metoder som kan ta hensyn til utvalgsfeil (men ennå ikke målefeil) i DEA-metoden. De statistiske metodene gjør det mulig å teste modellspesifikasjonen, betydningen av ”outliere”, etc, samt å gi konfidensintervaller for de estimerte størrelsene. Valget mellom de ulike metodene må derved dessverre fortsatt gjøres ut fra ekstern informasjon: er forholdene for sektoren slik at en tror at variasjonen i observert produktivitet i hovedsak skyldes tilfeldige avvik som målefeil, eller tror en at den vesentligste delen av variasjonen skyldes ineffektivitet. Selv om eventuelle målefeil vil ha en tendens til å redusere estimert effektivitet i en DEA - modell, vil det også være slik at den tette omhyllingen har en tendens til å øke estimert effektivitet. Hvilken vei eventuelle feil vil påvirke resultatene er usikkert. Ofte vil det være naturlig å bruke både parametriske og

ikke-parametriske metoder til å belyse det samme materialet, men internasjonalt er DEA i dag den mest utbredte metoden i studier av offentlig sektors effektivitet.



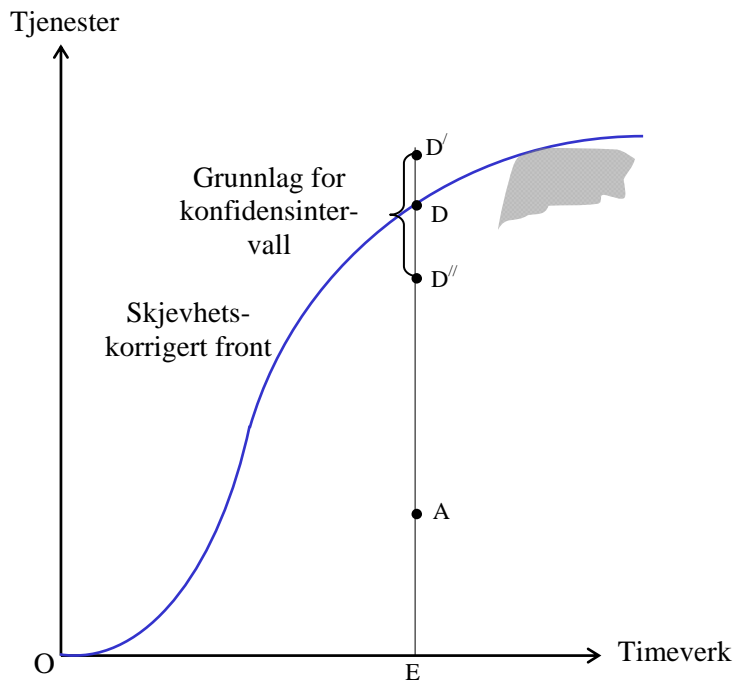
Figur 4. Metoder for å estimere mulighetsområdet illustrert i faktor - produktrommet

Utvalgsskjevhet og statistisk usikkerhet

Det generelle problemet med å kunne trekke sikre konklusjoner ved evaluering av produktivitet og effektivitet er antallet observasjoner man har til rådighet. Det er ofte slik at antall enheter som eksisterer, er for lite til å gi entydige konklusjoner selv om data er av utmerkt kvalitet. I prinsippet kan det tenkes et stort antall andre tilpasninger enn de som er observert i data: poenget er at det rent hypotetisk kan eksistere et ubegrenset antall alternative enheter til det begrensede antall som faktisk eksisterer. I statistisk forstand kan vi da snakke om at det faktisk antallet observasjoner gir grunnlag for utvalgsskjevhet. Dette kan være et problem i alle metoder som baseres på empiriske observasjoner av atferd, men er spesielt et problem i DEA-metoden. I tillegg vil det i alle statistiske metoder være slik at resultatene vil være mindre usikre jo flere

observasjoner de baseres på. Det er utviklet en statistisk metode til å korrigere for utvalgsskjevheten, som samtidig innebærer at vi får beregnet konfidensintervall som mål på den statistiske usikkerheten for den størrelsen vi er interessert i (effektivitetsscore eller produktivitetstall). Metoden kalles bootstrapping og innebærer at det konstrueres kunstige data for hypotetiske enheter, og at det gjennomføres beregninger på hvert sett av kunstige datapunkter (lik det opprinnelige antall enheter) for tilstrekkelig mange kunstige datasett. Antall replikasjoner er gjerne 2000. Den statistiske prosess som har generert de opprinnelige data og ført til effektivitets- og/eller produktivitetsforskjeller må spesifiseres for å kunne generere de kunstige observasjonssett (se f.eks. Førsumd et al., 2006).

Vi kan bruke figur 2 til å gi en skisse av hva slags informasjon bootstrapping gir. I figur 5 kan vi nå tenke på grafen til frontfunksjonen i produksjons-faktorrommet som en skjevhetsskorrigert front. Rent konkret er det punktet D på fronten som er skjevhetsskorrigert, dvs. rensket for beregnede utvalgsskjevheter. Dette D-punktet vil ligge ovenfor punktet D i figur 2. Det produksjonsøkende effektivitetsmål EA/ED er skjevhetsskorrigert. Produktivitetstall som definert i figur 3 kan korrigeres tilsvarende. Metoden gir oss også punktene D' og D'' som bestemmes ved hvilken konfidensgrad vi vil ha ved testing av hypoteser om effektivitet eller produktivitet.



Figur 5. Konfidensintervall for effektivitet

5. Måling av produktivitetssendringer før, under og etter reformen

Databehovet

Som nevnt i innledningen ble effektivitet og produktivitetsutvikling til likningskontorene analysert med utgangspunkt i eksisterende data (Førsund og Lindseth (2004), Førsund et al. (2005), (2006)). Når det gjelder produktvariable for fastsettingsfunksjonen er de 6 tjenestene som ble definert i tidligere analyser de tjenester som man også kan kvantifisere for de nye skattekontorer. Men resultatene viste mindre produksjonsøkingspotensial når det gjaldt variablene kontroll av folkeregister og klager på likning enn de andre 4 (Førsund, 2006). Kontroll av folkeregister er en relativt beskjedent aktivitet, og klage på likning vil også omfatte de saker likningskontorene selv vil prøve fordi rettspraksis ikke er etablert. Det varierende innhold av denne variabelen kan tilsi at denne bør sløyfes i nåværende form.

Det bør vurderes å inkludere skatte - innkrevingsfunksjonen i en ny analyse. Denne vil, som nevnt innledningsvis, bli organisert i egne kontorsteder. Det vil være mulig å gå tilbake å få data for de gamle kontorer før reformen. Ressursen vil være arbeidskraft og tjenesteproduktene kan være innbetalt MVA av sum krav, innbetalt skatt av sum krav og saksbehandlingstid på arveavgift.

Tidligere undersøkelser dekker arbeid med folkeregister og likning for hovedgrupper av skatteyttere. Arbeid med veiledning og informasjon er ikke tatt med. Men dette er et prioritert område under ROS og bør vurderes å komme med. Det vil da være et spørsmål i første omgang om omfanget av publikum - henvendelser til de enkelte kontorer registreres. Det vil nok være meget krevende å finne noe ut om verdien av veiledningen for den enkelte i første omgang og for kvaliteten av utfylte skjemaer i neste omgang. Arbeidet til sentrale "call"- sentere vil være vanskelig å inkludere. Fastsetting av merverdiavgift og arveavgift har heller ikke vært med tidligere. Hvilket statistikkgrunnlag man har, bør undersøkes.

Et viktig mål for ROS – reformen er å oppnå kvalitetsforbedringer (St. prp. Nr.1, 2006-2007). Hvordan kvalitet kan tallfestes for hvert kontor må diskuteres med Skattedirektoratet. Utsagn om kvalitet basert på intervju av ansatte kan neppe gjennomføres annet enn på et aggregert nivå grunnet kostnader. Når det gjelder brukerundersøkelser er det vel også mest realistisk at det kan gjøres undersøkelser på et aggregert nivå og ikke for ”kundene” til hvert enkelt kontor. En mulighet er å knytte kvalitet for det enkelte kontor til kompetanseheving av staben, målt ved erfaring, utdanning, antall gjennomførte målrettede kurs, og lignende.

Når det gjelder kontroller og klager kan slike opplysninger kunne gi noe informasjon om kvalitet, men et problem er mye av innsatsen til kontorene kan gå til å avklare uklarheter i lovtolkninger og regler laget utenfor Skattedirektoratet. Hvis klager på likning skal kunne brukes til å måle kvalitet må opplysning om klage tas til følge eller ikke foreligge. En klagevariable kan være et rimelig godt uttrykk for kvalitet i saksbehandlingen hvis det er publikum som klager innenfor en etablert tolkning av regelverket, og hvis antall klager som blir tatt til følge, blir registrert. Dess færre klager som blir tatt til følge, dess bedre kvalitet.

Et forhold som bør undersøkes er om det er aktiviteter under Kontroll og rettsanvendelse som kan gi indikasjoner på kvalitet.

Sammenliknbarhet av produktene bør vurderes. Når det gjelder f. eks likninger kan ”kundene” fordele seg på en bred skala når det gjelder kompleksiteten og dermed tiden det tar å behandle hvert tilfelle. I tidligere studier er det for eksempel skilt mellom type skatteyster ut fra dette forhold. Men selv om spesialkontorene for store bedrifter, oljeskatt og utenlandssaker tar spesielle saker, så er det vel en betydelig variasjon igjen. Man kan enten tenke seg å etablere en indikator som fanger opp vanskelighetsgrad, eller vurdere en inndeling i flere kategorier kunder.

Måling av produktivitetsendringer før reformen

Ved den nye organiseringen vil bruk av arbeidskraft målt ved årsverk kunne være bedre egnet som mål for ressursbruken enn totale utgifter brukt tidligere. Arbeidsoppgaver kan nå sendes over fra et kontor til et annet, Mens arbeidskraftbruken ved et kontor vil måles ganske nøyaktig er det ikke sikkert at kostnader i kroner vil måles like nøyaktig. Det er mulig å oppdatere

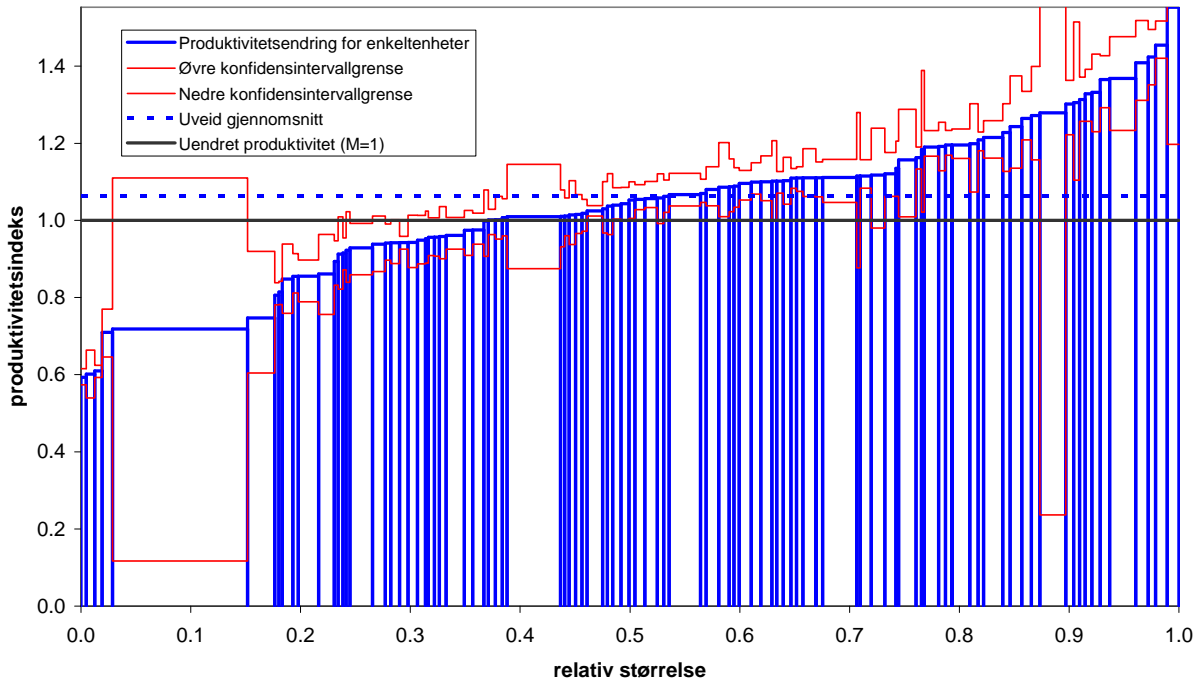
databasen for 2002-2004 og erstatte totale utgifter med arbeidsinnsats. Det bør videre vurderes å utvide databasen med årene 2005-2007 (eller 2006) hvis 2007 er et for unormalt år på grunn av pågående omlegging).

Figur 6 viser resultatene av å bytte ut totale justerte kostnader med justerte årsverk, men ellers beholde samme modell som i Førstund et al.(2006). Figuren er sortert etter stigende nivå for produktivitetsendringen 2002-2004. Bredden på hvert histogram er proporsjonal med antall årsverk. Et 95% konfidensintervall er angitt som trappetrinn eller horisontale streker over og under den skjevhetsskorrigerte produktivitetsendring. Vi ser at omtrent 2/3 av enhetene målt ved den relative bruk av årsverk har hatt en produktivitetsfremgang regnet etter størrelse på skjevhetsskorrigert produktivitet. Gjennomsnittlig produktivitetsfremgang vises ved linjen som ligger over 1.0-linjen og er på 1.062, dvs 6.2 %. Rent visuelt kan man lett identifisere store, mellomstore og små enheter. En stor enhet og noen mellomstore har hatt produktivitsnedgang sammen med en del små enheter, mens de fleste mellomstore og flere små enheter har hatt produktivitetsfremgang.

Imidlertid endrer dette bilde seg når vi trekker inn usikkerheten skapt ved utvalgsskjevheten. Bare ca 40% av enhetene målt ved arbeidskraftbruk har signifikant produktivitetsfremgang, dvs. nedre konfidensintervallgrense ligger over linjen som viser 1.0. Av de 5 største enhetene med gjennomsnittlig fremgang er det 2 som ikke har signifikant fremgang.

Motsatt vei er det relativt få enheter som har en signifikant produktivitsnedgang. Dette dreier seg om under halvparten av enhetene (regnet etter relativ bruk av arbeidskraft) som har hatt gjennomsnittlig tilbakegang. Spesielt ser vi at en hypotese om at den største enheten har hatt tilbakegang, kan forkastes på 95% nivå. Dette understreker betydningen av informasjonen vi har om type enhet. Fordi det er så få store enheter vil produktivitetstallene for disse kunne være beheftet med størst usikkerhet.

Resultatene i figur 6 er ganske like resultatene når totale kostnader ble brukt som ressursmål i Førstund et al. (2006). Det skulle derfor ikke være noe i veien rent faglig for å bytte ut kostnader



Figur 6. Produktivitetsendring 2002 – 2004. 95% konfidensintervall

med arbeidskraftbruk. Men vi ser at konfidensintervallene til dels er store. Det må derfor undersøkes nærmere om hvilken modellspesifikasjon som er den mest tilfredsstillende å bruke i det videre arbeidet.

Produktivetsutviklingen før og etter reformen

En første tilnærming til å finne produktivitetsvirkningene av en reform er å sammenlikne enten produktivetsnivået eller produktivitetsendringene før og etter reformen. Generelt vil det da være ønskelig med data for sammenliknbare enheter både før, under (i tilfellet ikke alle enheter endres samtidig, slik forholdet er for NAV – reformen) og etter at reformen er gjennomført. I motsetning til bruk av aggregerte data på fylkes- eller nasjonalt nivå, vil bruken av mikroinformasjon om alle enhetene sikre at vi kan få nok observasjoner til statistisk signifikante utsagn om produktivetsutviklingen. En positiv utvikling i gjennomsnitt for landet kan potensielt skjule at mange har blitt dårligere og mange har blitt bedre. I tillegg trengs den informasjonen om sammenvekting av ulike tjenester som best kan tallfestes ved statiske metoder som beskrevet ovenfor.

ROS – reformen innebærer at tidligere kontorer kan bli reorganisert, slått sammen, etc. til de nye fysiske kontorsteder. Dermed kan vi ikke følge produktivitetsutviklingen for enkeltenheter før og etter reformen. Men det forhold at de samme typer tjenester produseres både før og etter reformen muliggjør likevel en sammenlikning av produktivitetsutviklingen for alle enhetene som finnes før reformen og sammenligne med den tilsvarende utviklingen for alle enhetene som finnes etter reformen. I tillegg vil det være mulig å beregne produktivitetsutvikling på aggregert og nasjonalt nivå over hele tidsperioden. Det er 5 typer tjenester som produseres i hver region; veiledning, fastsetting, kontroll og rettsanvendelse, bekjempelse av skattekrim og innkreving. Vi foreslår at vi konsentrerer oss om fastsetting og innkreving. Fastsettingsfunksjonen kan i prinsippet beskrives ved de samme tjenester som ved den tidligere organiseringen. Det gjenstår som en del av prosjektet å formulere tjenester for andre funksjoner

Ved endring av primærenhetene ved sammenslåing som under NAV – reformen er en mulighet som vurderes til å få isolert virkningen av reformen, å måle de nye kontorenhetene mot teknologifront som gjelder de gamle enheter. Et tilsvarende opplegg kan brukes i evaluering av ROS - reformen. Som beskrevet tidligere vil primærenhetene endre seg etter reformen. Men fordi både produktspekteret og type ressurs som brukes er de samme som før reformen, vil en måte å få fram i alle fall en bruttoeffekt av reformen på, være å utbytte beste praksis estimert ved de gamle enhetene og på data for perioder før reformen. Effektivitet av de nye kontorenhetene kan så måles mot den gamle fronten. Hvis de nye enhetene systematisk ligger over (under) fronten etablert av de gamle enheter, er dette en god indikasjon på at reformeffekten er positiv (negativ). Dess større andel av nye kontorsteder som ligger over den gamle fronten, dess mer positivt har reformeffekten virket rent kvalitativt.

En annen fremgangsmåte er å estimere en ny beste praksis basert på de nye enheter og så sammenlikne beliggenheten av selve frontene. Hvis den nye fronten i sin helhet ligger over den gamle, kan man slutte at rent kvalitativt så har reformen vært positiv for effektiviteten, og *vice versa* hvis den nye fronten i sin helhet ligger under den gamle. Hvis frontene krysser hverandre får man interessante indikasjoner på for hvilke produktkombinasjoner reformen har virket positivt, respektivt negativt.

En tredje fremgangsmåte er å sammenlikne produktivitetstutviklingen før og etter. Det gjennomsnittlige tempoet i produktivitetstutviklingen vil her være av interesse. Hvis formen på produktivitetstutviklingen varierer systematisk med for eksempel en jevn produktivitetstutvikling før reformen, men et skift til en raskere økning etter reformen, vil dette peke på en positiv reformeffekt (og vice versa). Metoden som brukes gir anledning til å følge produktivitetstutviklingen til de enkelte enheter. Det vil da være interessant med selve mønsteret i hvordan utviklingen av de enkelte enheter er før og etter reformen. Et interessant spørsmål vil f. eks. være hvor stor andel enheter med fremgang, stillstand og tilbakegang utgjør før og etter, og et annet er hvordan går det med små og store enheter før og etter.

Isolering av reformeffekten

Et problem er å isolere reformeffekter fra andre forhold som endrer seg over tid. Dette kan være at det stadig kommer bedre IT-løsninger, eller endringer i regelverk (stille aksept), og liknende. For å studere dette er det mulig å dekomponere produktivetsmål i bidrag fra slike variable ved å gjennomføre en totrinnsanalyse. Ved en regresjonsberegning basert på "bootstrapping" vil bidraget fra forklaringsvariable som innføring av ROS, innføring av IT, regelendringer, og lignende dekomponeres. Et tilsvarende opplegg er planlagt brukt til analyser av NAV – reformens resultater.

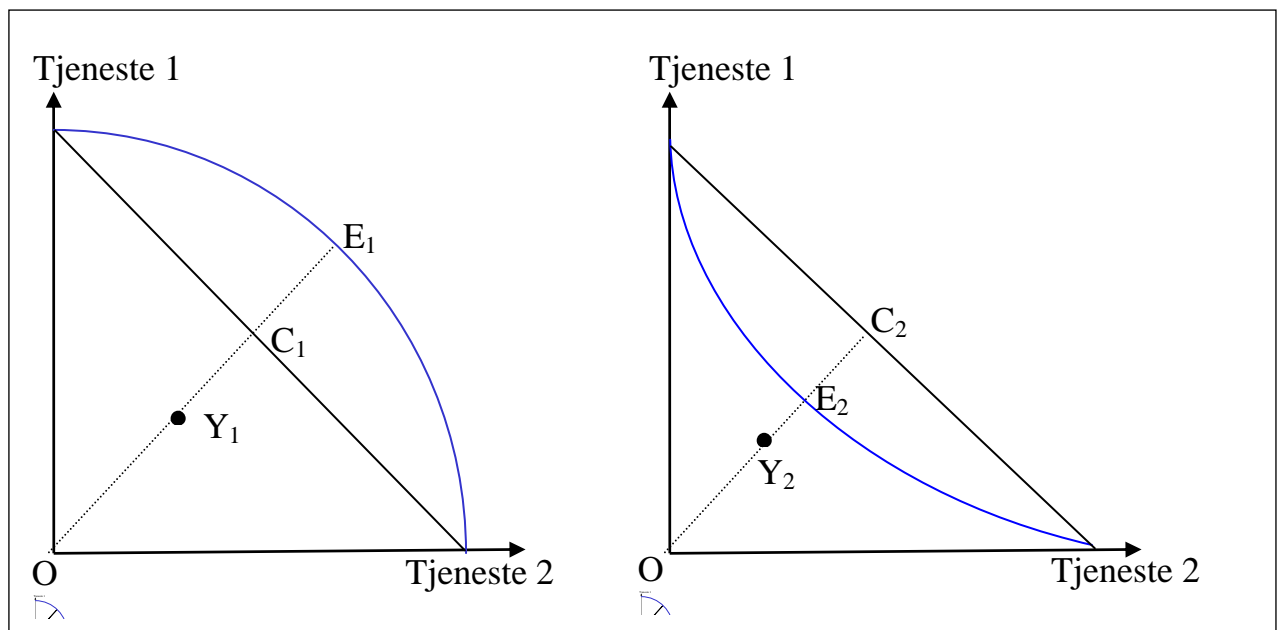
Likevel er det ikke nødvendigvis mulig å fange opp alle andre endringer over tid i variable som kan inkluderes i en totrinnsanalyse. Sterkere resultater kan en få f.eks. ved å sammenligne med en kontrollgruppe som ikke er utsatt for reformen. I den nevnte analysen av sykehusreformen ble de øvrige nordiske landene brukt som kontrollgruppe; antakelsen er da at sykehusene i de andre landene er blir utsatt for de samme øvrige endringene som de norske sykehusene og det er den særnorske produktivitetstutviklingen som hjelper til å identifisere reformeffekten. I NAV-reformen er planen å utnytte det faktum at ikke alle NAV- kontorene blir dannet samtidig; dermed blir de arbeidskontorene og trygdekontorene som venter med å slås sammen en kontrollgruppe til de NAV- kontorene som dannes tidlig. Det gjenstår å se om denne variasjonen er sterk nok til å forbedre anslagene på reformeffekten.

Forklaringer på effektivitet

Analysemetodene kan ikke forklare hvorfor noen enheter er mer effektive enn andre. Men i en tidlig fase med ny struktur vil det være av spesiell interesse å gjøre en nærmere analyse av de kontorsteder som oppnår høye effektivitetstall. Kan man finne andre trekk enn dem som allerede er målt som skiller disse enhetene ut fra andre? Tilsvarende kan man ta utgangspunkt i de enheter som oppnår lav effektivitet; er de faktorer her som skiller disse kontorene fra de andre? Slike analyser kan gi verdifulle innspill til å øke prestasjonene for enheter med relativt lave effektivitetstall.

6. Effekter av spesialisering

Et viktig mål for ROS – reformen er å oppnå spesialiseringsgevinster. Det foreslåtte analyseopplegg er godt egnet til å få fram om spesialiseringsgevinster er til stede, eller om det tvert i mot er samproduksjonsgevinster. Slike analyser er foretatt i Kittelsen og Førsum (1992) for norske domstoler, og Magnussen og Kittelsen (2003) for sykehus. Spesialiseringsevner på



Figur 7. Samdriftsfordeler (til venstre) og spesialiseringsfordeler (til høyre)

fronten vil innebære at krumningen i figurer som f.eks. Figur 3 vil krumme inn mot origo i stedet for vekk fra origo.

Figur 7 viser samdriftsfordeler og spesialiseringsfordeler for en enhet Y som yter en kombinasjon av tjeneste 1 og tjeneste 2. I begge panelene er den blå kurven den sanne teknologifronten, og alle punkter på linjen vil ha samme ressursbruk (kostnader). I det venstre panelet vil en øke nivået av begge tjenester til E_1 uten å bruke flere ressurser, men om en kun hadde enheter som var fullt ut spesialiserte i produksjon av tjeneste 1 eller tjeneste 2 ville en kun hatt mulighet for å øke produksjonen til C_1 . Det er gevinster å hente i å yte begge tjenestene innenfor samme kontor. I det høyre panelet er det derimot spesialiseringsfordeler siden en kan øke tjenestene mer (til C_2) ved å legge aktiviteten til to spesialiserte kontorer enn ved å yte begge tjenestene ved samme kontor (til E_2).

Spesialiseringsevner i gjennomsnitt vil kunne identifiseres hvis det er slik at enheter med spesialisert produksjon gjennomgående ligger nærmere beste praksis – fronten enn andre kontorer. Det kan også testes statistisk (med en ikke-parametrisk test) om gruppen av spesialiserte enheter har høyere effektivitet eller produktivitet enn gjennomsnittet. I analysene av sykehus ble det ikke funnet spesialiseringsevner; tvert om var det slik at de beste enheten var mer produktive dersom de ikke var spesialisert. En slik analyse er ikke avhengig av å følge de samme enhetene over tid. Tvert om er det en fordel å spenne ut mulighetsområdet ved å ha observasjoner som er lite spesialisert i samme datasett som spesialiserte enheter.

Det vil også være av interesse å bringe størrelsen av et kontor eksplisitt inn i en slik type analyse. Er det store kontorer som er spesialisert, eller små kontorer? Da ROS – reformen eksplisitt medfører flytting av arbeid med likninger ut fra hjemstedskommunen til skatteyterne vil det være av spesiell interesse å få data for hvor mange ”fremmede” kunder et kontor behandler og så studere konsekvensene for effektivitet eller produktivitet av en stigende andel fremmede kunder.

7. Produktivitet for skattekontorene samlet

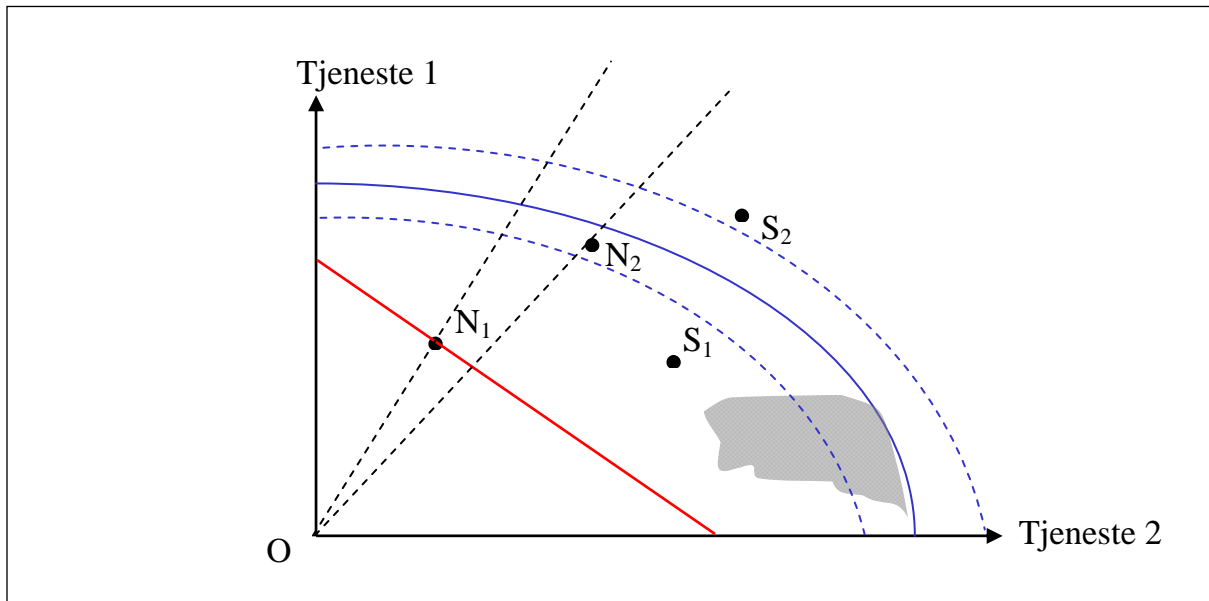
For å beregne produktivitet for alle kontorene under ett, både før og etter reformen, må en summere all ressursbruk og all tjenesteproduksjon for landet for hver tidsperiode/år. Resultatet vil da kun være en observasjon per periode og det vil ikke være informasjon om spredning av produktivitetsutviklingen som kunne nyttiggjøres i å finne årsakssammenhenger mellom reform og produktivitet. En vil likevel kunne si noe om den samlede produktiviteten i alle perioder før og etter reformen dersom en kan løse problemet med sammenveining av tjenestetyper som er beskrevet ovenfor. Det er her to fremgangsmåter som peker seg ut:

i) Finne vektorer fra andre kilder som reflekter relative kostnader. Ved dette opplegget vil vektene for hvordan produkter og ressurser skal veies sammen, være eksogene, dvs. disse må etableres først. Når det gjelder ressurser kan dette forenkles ved bare å se på en ressurs; arbeidskraft målt i hele årsverk. Når det gjelder produkter kan statistikken for tidsbruk nyttiggjøres til å lage "kostnadsvektor" analogt til beregning av aggregert produksjon ved sykehus der behandling av mange hundre typer sykdommer veies sammen ved kostnadsandeler.

ii) La vektene være beregnet av DEA - metoden på grunnlag av observasjonene, dvs. behandle de årlige observasjonene av sektoren som helhet som vanlige enheter, og sammenligne produktivitetsutviklingen til disse med fronten som er tallfestet på grunnlag av alle enkeltkontorene. Dette svarer til Farrell (1957) sitt begrep om struktureffektivitet, og er et mål på hvor godt en sektor som helhet gjør det i forhold til de beste enkeltenhetene. Her vil sammenveiningen av produkter (og eventuelt ressurser hvis det er flere) skje endogen, dvs. dette vil være en del av selve modell-løsningen. De to forskjellige metoder vil således gi en validering av beregningene.

Figur 8 kan illustrere disse metodene. La N_1 være den gjennomsnittlige produksjonen av to tjenester pr enhet arbeidskraft for alle skattekontorene i en periode før reformen, og la N_2 være de tilsvarende tallene for perioden etter reformen. De faste eksogene vektene i metode a) vil kunne være representert ved den røde linjen (for enkelthets skyld kun tegnet gjennom N_1), og produktivitetsframgangen vil være målt ved hvor langt N_2 ligger foran den røde linjen gjennom N_1 målt langs en stråle fra origo. Metode b) vil i stedet være å benytte den fronten vi har funnet ved DEA-metoden, representert ved den blå kurven, til å sammenligne produktiviteten til N_1 og

N_2 . Dette er helt parallelt til hvordan Malmquist- indeksen er beregnet for en enkeltenhet, slik som i figur 3, men i stedet anvendt på skatteetaten som helhet. Figuren illustrerer også konfidensintervallet til denne fronten, slik som i figur 4, og derved hvordan metode b) kan gi oss informasjon om den statistiske usikkerheten til produktivitsutviklingen for sektoren som helhet.



Figur 8. Produktivitsindekser for skatteetaten i ulike land

Nordisk "benchmarking"

En total produktivitsberegning etter begge disse metodene vil kunne gjennomføres i ved å sammenligne de nordiske landene og indikere om det er noe spesifikt skift i norsk utvikling relativt til de andre landene etter ROS reformen. En forutsetning er at det vil være mulig å skaffe data fra andre land på de samme variablene som inngår i den norske studien. I så fall kan en også for de andre landene beregne produktivitsutviklingen over tid ved å bruke de "norske" vektene, enten de er hentet fra eksogene kilder eller beregnet med DEA-metoden på de norske enkeltkontorene. I figur 7 vil f.eks. den svenske skatteetaten kunne representeres ved S_1 og S_2 i de to periodene, og en vil kunne få informasjon om de relative produktivitsendringene.

Selv om en med metode b) vil få et mål for usikkerheten også ved produktivitsanslagene i de andre nordiske landene ved å utnytte konfidensintervallene for fronten som er tallfestet på grunnlag av de norske skattekontorene, vil en ikke få like god statistisk presisjon med aggregerte

data som en kunne ha fått hvis en hadde informasjon om de enkelte skattekontorene i de andre nordiske landene. Hvis et nordisk land viser bedre produktivitsutvikling ved denne metoden enn et annet land, kan det skyldes tilfeldige utslag dersom det er stor spredning mellom enkeltkontorer; hvis det derimot er slik at de aller fleste kontorene i et land gjør det bedre enn i et annet må det sannsynligvis være noe ved landene, f.eks. at de har organisert skattekontorene på en bedre måte. Data av sammenlignbar karakter for de enkelte skattekontorene i flere nordiske land ville gjøre det mer sannsynlig å finne statistisk signifikante forklaringer på forskjeller i produktivitsutvikling. Spesielt ville det være lettere å kunne finne klare sammenhenger mellom ROS - reformen og produktivitsutviklingen dersom en hadde data for enkeltenhetene (eller et utvalg) også i andre nordiske land, slik det bl.a. ble gjort i evalueringen av den norske sykehusreformen

Referanser

- Caves, D.W., Christensen, L.R. and Diewert, E. (1982): 'The economic theory of index numbers |and the measurement of input, output, and productivity,' *Econometrica* 50(6), 1393-1414
- Charnes, A., W.W. Cooper and E. Rhodes (1978): "Measuring the Inefficiencies of Decision Making Units", *European Journal of Operational Research* 2, 429-444.
- Coelli, T., D.S.P. Rao and G.E. Battese (1998): *An Introduction to Efficiency and Productivity Analysis*. Boston/Dordrecht/London, Kluwer Academic Publishers.
- Farrell, M. J. (1957): "The measurement of productive efficiency," *Journal of the Royal Statistical Society, Series A*, 120 (III), 253-281.
- Førsund, F. R. (2006): Sektoranalyser - Gjennomgang av samfunnsøkonomiske analyser av effektiviseringspotensialer for utvalgte sektorer, Rapport 9/2006, Frischsenteret.
- Førsund, F. R. og Lindseth, F. (2004): "Analyse av produktivitet ved Likningskontorene: DEA analyse av produksjonen ved likningskontorene i 2002 og 2003", upublisert notat, Stiftelsen Frischsenteret for samfunnsøkonomisk forskning og Skattedirektoratet.
- Førsund, F. R., Kittelsen, S. A. C. and Lindseth, F. (2005): "Efficiency and productivity of Norwegian tax offices," Memorandum 29/2005 from the Department of Economics, University of Oslo.
- Førsund, F. R., Kittelsen, S.A.C., Lindseth, F. and Edvardsen, D. F. "The tax man cometh - but is he efficient?" *National Institute Economic Review*, 2006, (197), pp. 106-19.
- Kittelsen, S.A.C. and Førsund, F. R. "Empiriske forskningsresultater om effektivitet i offentlig tjenesteproduksjon." *Økonomisk forum*, 2001, (6), pp. 22-29.
- Kittelsen, S.A.C. and Førsund, F. R. "Efficiency analysis of Norwegian district courts." *Journal of Productivity Analysis*, 1992, 3, pp. 277-306.
- Kittelsen, S.A.C., Magnussen, J. and Anthun, Kjartan S. "Sykehusproduktivitet Etter Statlig Overtakelse: En Nordisk Komparativ Analyse," *HERO Working paper. Oslo: Health Economics Research Programme at the University of Oslo (HERO)*, 2007.
- Magnussen, J. and Kittelsen, S.A.C. "Economies of scope in Norwegian hospital production - a DEA analysis," 8, Oslo: Health Economics Research Programme at the University of Oslo (HERO), 2003.
- St. meld. nr. 1, Nasjonalbudsjettet 2008.
- St. prop. Nr.1, 2006-2007.

Publikasjoner fra Frischsenteret

Alle publikasjoner er tilgjengelig i Pdf-format på : www.frisch.uio.no

Rapporter

1/2005	Pliktige elsertifikater	Rolf Golombek, Michael Hoel
2/2005	En empirisk analyse av indekssystemet i det norske legemiddelmarkedet	Tonje Haabeth
3/2005	Formelle og totale skattesatser på inntekt for personer med svak tilknytning til arbeidsmarkedet	Elisabeth Fevang, Morten Nordberg, Knut Røed
1/2006	Finansiering av tros- og livssynsamfunn	Aanund Hylland
2/2006	Optimale strategier i et to-kvotesystem	Rolf Golombek, Cathrine Hagem, Michael Hoel
3/2006	Evaluering av tilskuddsordningen for organisasjoner for personer med nedsatt funksjonsevne	Rolf Golombek, Jo Thori Lind
4/2006	Aetats kvalifiserings- og opplæringstiltak – En empirisk analyse av seleksjon og virkninger	Ines Hardoy, Knut Røed, Tao Zhang
5/2006	Analyse av aldersdifferensiert arbeidsgiveravgift	Gaute Ellingsen, Knut Røed
6/2006	Utfall av yrkesrettet attføring i Norge 1994-2000	Tyra Ekhaugen
7/2006	Inntektsfordeling og inntektsmobilitet – pensjonsgivende inntekt i Norge 1971-2003	Ola Lotherington Vestad
8/2006	Effektiv måloppnåelse En analyse av utvalgte politiske målsetninger	Nils-Henrik M. von der Fehr
9/2006	Sektoranalyser – Gjennomgang av samfunnsøkonomiske analyser av effektiviseringspotensialer for utvalgte sektorer	Finn R. Førund
10/2006	Veien til uføretrygd i Norge	Elisabeth Fevang, Knut Røed
1/2007	Generisk bytte En økonometrisk studie av aktørenes og prisenes betydning for substitusjon	Vivian Almendingen
2/2007	Firm entry and post-entry performance in selected Norwegian industries	Ola Lotherington Vestad

Arbeidsnotater

1/2005	Lifetime earnings	Fedor Iskhakov
--------	--------------------------	----------------

2/2005	Skattefunksjoner i Norge 1990 - 2004	Vivian Almendingen
1/2006	Costs and coverage of occupational pensions	Erik Hernæs, Tao Zhang
2/2006	Inntektsfordelingen i Norge, og forskjellige årsaker til ulikheter i pensjonsgivende inntekt	Ola Lotherington Vestad
3/2006	The Wage Effect of Computer-use in Norway	Fitwi H. Wolday
1/2007	An evaluation of the labour market response of eliminating the retirement earnings test rule	Erik Hernæs, Zhiyang Jia
1/2008		
2/2008		
3/2008	Analyseopplegg for å kunne male om reorganisering av skatteetaten fører til en mer effektiv ressursbruk	Finn R. Førsum, Sverre A.C. Kittelsen

Memoranda

Serien publiseres av Økonomisk institutt, Universitetet i Oslo, i samarbeid med Frischsenteret. Listen under omfatter kun memoranda tilknyttet prosjekter på Frischsenteret. En komplett oversikt over memoranda finnes på <http://www.oekonomi.uio.no/memo/>.

5/2005	The Kyoto agreement and Technology Spillovers	Rolf Golombek, Michael Hoel
6/2005	Labor supply when tax evasion is an option	Øystein Jørgensen, Tone Ognedal, Steinar Strøm
9/2005	The Fear of Exclusion: Individual Effort when Group Formation is Endogenous	Kjell Arne Brekke, Karine Nyborg, Mari Rege
11/2005	Tax evasion and labour supply in Norway in 2003: Structural models versus flexible functional form models	Kari Due-Andresen
16/2005	Prioritizing public health expenditures when there is a private alternative	Michael Hoel
17/2005	Young and Out: An Application of a Prospects-Based Concept of Social Exclusion	Oddbjørn Raau, Jon Rogstad, Knut Røed, Lars Westlie
18/2005	Immigrants on Welfare: Assimilation and Benefit Substitution	Tyra Ekhaugen
19/2005	Time and Causality: A Monte Carlo Assessment of	Simen Gaure, Knut Røed, Tao

	the Timing-of-Events Approach	Zhang
20/2005	Organisational Change, Absenteeism and Welfare Dependency	Knut Røed, Elisabeth Fevang
21/2005	Extracting the causal component from the intergenerational correlation in unemployment	Tyra Ekhaugen
29/2005	Efficiency and productivity of norwegian tax offices	Finn R. Førsund, Sverre A.C. Kittelsen, Frode Lindseth
33/2005	Price regulation and generic competition in the pharmaceutical market	Dag Morten Dalen, Steinar Strøm, Tonje Haabeth
34/2005	American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States	Markus Jäntti, Bernt Bratsberg, Knut Røed, Oddbjørn Raaum, Robin Naylor, Eva Österbacka, Anders Björklund, Tor Eriksson
35/2005	Earnings persistence across generations: Transmission through health?	Tor Eriksson, Bernt Bratsberg, Oddbjørn Raaum
1/2006	The Determinants of Occupational Pensions	Erik Hernæs, John Piggott, Tao Zhang, Steinar Strøm
4/2006	Moving between Welfare Payments. The Case of Sickness Insurance for the Unemployed	Morten Henningsen
6/2006	Justifying Functional Forms in Models for Transitions between Discrete States, with Particular Reference to Employment-Unemployment Dynamics	John Dagsvik
15/2006	Retirement in Non-Cooperative and Cooperative Families	Erik Hernæs, Zhiyang Jia, Steinar Strøm
16/2006	Early Retirement and Company Characteristics	Erik Hernæs, Fedor Iskhakov and Steinar Strøm
20/2006	Simulating labor supply behavior when workers have preferences for job opportunities and face nonlinear budget constraints	John K. Dagsvik, Marilena Locatelli, Steinar Strøm
21/2006	Climate agreements: emission quotas versus technology policies	Rolf Golombek, Michael Hoel
22/2006	The Golden Age of Retirement	Line Smart Bakken
23/2006	Advertising as a Distortion of Social Learning	Kjell Arne Brekke, Mari Rege
24/2006	Advertising as Distortion of Learning in Markets with Network Externalities	Kjell Arne Brekke, Mari Rege
26/2006	Optimal Timing of Environmental Policy; Interaction Between Environmental Taxes and Innovation Externalities	Reyer Gerlagh, Snorre Kverndokk, Knut Einar Rosendahl
3/2007	Corporate investment, cash flow level and market	B. Gabriela Mundaca, Kjell

	imperfections: The case of Norway	Bjørn Nordal
4/2007	Monitoring, liquidity provision and financial crisis risk	B. Gabriela Mundaca
5/2007	Total tax on Labour Income	Morten Nordberg
6/2007	Employment behaviour of marginal workers	Morten Nordberg
10/2007	Long-term Outcomes of Vocational Rehabilitation Programs: Labor Market Transitions and Job Durations for Immigrants	Tyra Ekhaugen
12/2007	Pension Entitlements and Wealth Accumulation	Erik Hernæs, Weizhen Zhu
13/2007	Unemployment Insurance in Welfare States: Soft Constraints and Mild Sanctions	Knut Røed, Lars Westlie
15/2007	Farrell Revisited: Visualising the DEA Production Frontier	Finn R. Førsund, Sverre A. C. Kittelsen, Vladimir E. Krivonozhko
16/2007	Reluctant Recyclers: Social Interaction in Responsibility Ascription	Kjell Arne Brekke , Gorm Kipperberg, Karine Nyborg
17/2007	Marital Sorting, Household Labor Supply, and Intergenerational Earnings Mobility across Countries	O. Raaum, B. Bratsberg, K. Røed, E. Österbacka, T. Eriksson, M. Jäntti, R. Naylor
18/2007	Pennies from heaven - Using exogenous tax variation to identify effects of school resources on pupil achievement	Torbjørn Hægeland, Oddbjørn Raaum and Kjell Gunnar Salvanes
19/2007	Trade-offs between health and absenteeism in welfare states: striking the balance	Simen Markussen



Frischsenteret

Stiftelsen Frischsenteret for samfunnsøkonomisk forskning er en uavhengig stiftelse opprettet av Universitetet i Oslo. Frischsenteret utfører samfunnsøkonomisk forskning i samarbeid med Økonomisk institutt ved Universitetet i Oslo. Forskningsprosjektene er i hovedsak finansiert av Norges forskningsråd, departementer og internasjonale organisasjoner. De fleste prosjektene utføres i samarbeid mellom Frischsenteret og forskere ved andre norske og utenlandske forskningsinstitusjoner.

Frischsenteret
Gaustadalléen 21
0349 Oslo
Tlf: 22958810
Fax: 22958825
frisch@frisch.uio.no
www.frisch.uio.no