

Bilaga 2

Lennart Hjalmarsson

Juni 27, 2010

Metoder i forskning om produktivitet och effektivitet

Sammanfattning

I en ESO-rapport, Hjalmarsson (1991), gavs en översikt över den mikrodatabaserade produktivitets- och effektivitetsforskningen. Syftet med denna rapport är att översiktligt följa upp utvecklingen efter 1991 inom detta område.

Generellt kan konstateras att utvecklingen på metodsidan under 1990- och 2000-talen har haft karaktären av vidareutveckling och finputsning av de framsteg som gjordes under 1970- och 1980-talen – med ett undantag, nämligen framväxten av metoder för beräkning av konfidensintervall för effektivitets- och produktivitetsmått såväl för deterministiska som stokastiska fronter. Dessutom har ett mycket stort antal empiriska tillämpningar publicerats. Gradvis har det vuxit fram två huvudalternativ för empiriska analyser, nämligen Data Envelopment Analysis, DEA, och stokastisk frontanalys, SFA. De bidrag till litteraturen som framförallt kommer att belysas här är:

- Metoder för beräkning av konfidensintervall för effektivitetsmått
- Utvecklingen av paneldatamodeller med tidsvarierande effektivitet
- Metoder för identifiering av extremobservationer (outliers)
- Stordriftsfördelar, optimal skala och skaleffektivitet

Vid sidan av frontlitteraturen, inom ramen för new empirical industrial organisation, NEIO, och med rötter i neoklassisk teori håller det också på att växa fram ett fält för produktivitetsstudier. Fokus ligger här på att fånga den dynamiska utvecklingen inom en sektor med entry och exit av företag och med skiftande efterfrågan och identifiera effekterna på produktiviteten av dessa faktorer. Fältet är metodmässigt mycket

avancerat och domineras ännu så länge av en liten grupp forskare på den internationella forskningsfronten.

1. Inledning

I en ESO-rapport, Hjalmarsson (1991), gavs en översikt över den mikrodatabaserade produktivets- och effektivitetsforskningen. Syftet med denna rapport är att översiktligt följa upp utvecklingen efter 1991 inom detta område.

Rapporten har en utformning och nivå som förutsätter att läsaren är väl förtrogen med innehållet i den tidigare ESO-rapporten. Detta innebär att denna rapport inte i någon högre grad överlappar med ESO-rapporten, utan kan ses som en direkt fortsättning på denna. För grundläggande definitioner och begrepp hänvisas till ESO-rapporten.

Produktivitet och produktionseffektivitet analyseras normalt inom ramen för den nationalekonomiska produktionsteorin. Den klassiska modellen är en produktionsfunktion med en enda produkt och ett antal produktionsfaktorer, alternativt en kostnadsfunktion med priser för insatsfaktorerna eller en vinstfunktion. Denna single output – multiple input-modell lämpar sig bra för analys av varuproduktion (och är heller inte speciellt datorkrävande), men är mindre lämpad för serviceproduktion som ofta karakteriseras av multiple output och multiple input och i många fall avsaknad av outputpriser. Under de senaste tre decennierna har det dock skett betydande framsteg, och ett antal metoder har utvecklats för analys av flervaruproduktion.

Dels har analyser av multiple-output och multiple-input-problem utvecklats inom ramen för estimering av kostnads- och vinstfunktioner, dels har metoder utvecklats för direkt skattning av produktionsfunktioner utan krav på prisdata. Intresset har i mycket hög grad varit inriktat på analys av produktionseffektivitet, dvs. produktionsenheternas relativa avstånd till best-practice i branschen – en mera avancerad form av benchmarkinganalys. Nyckelorden är här frontfunktionsestimering och distansfunktioner med idag två huvudmetoder, dels s.k. Data Envelopment Analysis (DEA), dels stokastisk frontanalys (SFA).

Efter de grundläggande genombrotten i den mikrodatabaserade produktivitets- och effektivitetsforskningen på 1970-talet utvecklades fältet snabbt på teori- och metodsidan och med mycket omfattande empiriska analyser av olika sektorer och aktiviteter. För en översikt av utvecklingen fram till 1990, se Hjalmarsson (1991). Metodmässigt domineras alltså fältet idag av dels estimering av stokastiska parametriska frontfunktioner (SFA), dels den icke-parametriska deterministiska DEA-ansatsen. Syftet med denna rapport är att ge en icke-teknisk översikt av framstegen inom detta område efter 1990. Eftersom forskningsfältet formligen exploderat under de senaste decennierna blir hänvisningarna till litteraturen mycket selektiva. De bidrag till litteraturen som framförallt kommer att belysas här är:

- Metoder för beräkning av konfidensintervall för effektivitetsmått såväl inom SFA som DEA
- Utvecklingen av paneldatamodeller med tidsvarierande effektivitet
- Metoder för identifiering av extremobservationer (outliers)
- Stordriftsfördelar, optimal skala och skaleffektivitet

Dessutom kommer jag att kort diskutera utvecklingen inom ett par närliggande områden med traditionell neoklassisk inriktning men av relevans för produktivitetsforskningen. Det gäller framförallt fälten:

- Traditionell neoklassisk produktivitetsforskning
- New empirical industrial organisation, NEIO
- Utvecklingen inom reglering och mechanism design, MD

Förutom några få monografier och samlingsvolymen består effektivitetslitteraturen i stor utsträckning av artiklar publicerade dels i nationalekonomiska tidskrifter, speciellt i *Journal of Productivity Analysis* och i *Journal of Econometrics*, dels i operationsanalytiska tidskrifter typ *Management Science* och *European Journal of Operational Research*. Artiklar inom NEIO och traditionell produktivitetsforskning och inom reglering och MD återfinns ofta i mera centrala nationalekonomiska tidskrifter typ *Rand Journal of Economics*, *Review of Economic Studies*, *American Economic Review* och *International Journal of Industrial Organisation*.

2. Nyckeltal och jämförelsekonkurrens

Beräkning av olika typer av nyckeltal eller indikatorer är en naturlig del av verksamheten inom en organisation, antingen det är en affärsdrivande verksamhet eller icke-vinstdrivande organisation. Nyckeltalsjämförelser, sk benchmarking, har traditionellt varit vanlig inom företag, men har under senare decennier också fått vid spridning inom offentlig sektor i olika delar av världen. Metodiken för nyckeltalsjämförelser har utvecklats starkt under senare år mot mera sofistikerade metoder och användarvänliga datorprogram. En orsak till detta är att avregleringen av många marknader har ökat kraven på effektiv reglering av ej konkurrensutsatta men ofta privatägda (mer eller mindre) naturliga monopol, såsom elnät, gasnät, telenät, bannät, etc. Globalisering och ökad konkurrens torde också bidragit till denna utveckling liksom ökade krav på effektivitet inom offentlig sektor. De mycket stora framsteg som gjorts inom forskningen, med DEA och stokastiska fronter som nyckelbegrepp, har varit en förutsättning för denna utveckling.

Ur presentationssynpunkt finns det ett antal olika, men delvis överlappande, sätt att strukturera fältet, såsom med avseende på:

- Komparativa vs absoluta nyckeltal
- Partiella vs totala nyckeltal
- Finansiella vs icke-finansiella nyckeltal
- Effektivitetstal vs produktivitetstal
- Ekonometriskt skattade vs matematiskt beräknade nyckeltal
- Stokastiska vs deterministiska nyckeltal

I princip kan vi skilja mellan det *komparativa* begreppet effektivitet och det *absoluta* begreppet produktivitet. Medan produktivitet definieras som kvoten mellan resursinsats och produktionsresultat definieras effektivitet (inom benchmarking) som avståndet från en produktionsenhet till en referenspunkt. Denna referenspunkt behöver inte vara förutbestämd utan kan utgöra en del av analysen. Ett exempel på förutbestämda referenspunkter var de, mycket detaljerade, jämförelser som, under vinjetten ”bäst i världen”, inom ramen för Produktivitetsdelegationens arbete, gjordes mellan SAS och Singapore Airlines, Electrolux och Whirlpool etc. Ofta bestäms dock referenspunkten

inom ramen för effektivitetsanalysen genom beräkningen av en best-practice-front som sedan ligger till grund för avståndsmätningen.

Att välja en speciell enhet eller organisation som referenspunkt i en parvis jämförelse är inte utan problem. ”Bäst-i-världen”- jämförelserna illustrerade problemet med att jämföra två enheter ur en lång rad olika aspekter. Sådana mångfacetterade jämförelser kan säkert vara av stort värde internt inom ett företag, men för en utomstående betraktare skymms skogen för alla träd. Ur extern policysynvinkel är resultaten av mindre värde. I princip kan ett aggregerat effektivitetsmått beräknas, genom att väga samman de olika komponenterna, men valet av vikter i ett sådant mått blir sannolikt relativt godtyckligt. Det är svårt att sätta ”prislappar” på många komponenter när inte marknadspriser finns att tillgå. I en frontfunktionsanalys undviks detta problem eftersom vikterna bestäms av teknologin.

En typ av nyckeltal är sk specifika åtgångstal, dvs. resursinsats dividerad med produktionsresultat i vid mening, dvs. inversen av produktivitet. Specifika åtgångstal är speciellt vanliga inom miljö- och energiområdet, med utsläpp av föroreningar eller energiåtgång per producerad enhet, där producerad enhet också kan vara en kylskåps- eller TV-timme eller bostadsyta.

Partiella produktivitetsmått går ofta under beteckningen nyckeltal för produktivitet (performance indicators). Svagheten med sådana nyckeltal är flera:

- De är ofta ad hoc-mässigt utvalda utan förankring i någon underliggande modell för produktionsprocessen. Man skulle kunna karakterisera detta sätt att mäta produktivitet som metodfritt.
- De är deterministiska utan hänsynstagande till slumpmässiga faktorer.
- De är partiella också i den betydelsen att någon kontroll ej sker för vad som endogen eller exogen påverkar resursförbrukningen.
- De är ofta inriktade på resursåtgång och processer och med mindre tonvikt på produktionsresultat.

- Vissa finansiella nyckeltal, t ex den andel av kostnaderna i en viss verksamhet som täcks med avgifter, kan lika gärna vara ett uttryck för monopolmakt som för effektivitet.
- De fokuserar på en speciell produktionsfaktor eller komponent utan att ta hänsyn till övriga produktionsfaktorer. Nyckelordet i resurshushållning är optimering (av samtliga resurser) – inte minimering av en viss resurs.

Jag återvänder till problemen med denna typ av nyckeltal i samband med diskussionen av partiella produktivitetsmått i avsnitt 3.

En central indikator är givetvis kostnaden per producerad enhet inom en viss typ av verksamhet. När flera varor eller tjänster produceras inom en verksamhet uppstår som regel problem att fördela fasta kostnader mellan olika tjänster och produkter vid denna typ av nyckeltalsberäkning, eftersom det inte existerar någon objektiv eller ”korrekt” metod att fördela sådana kostnader. Detta hindrar inte att jämförelser av enhetskostnader kan vara av stort värde som konkurrensinstrument i reglerade sektorer. Sådan jämförelsekonkurrens var av stor betydelse i regleringen av den ”gamla” elmarknaden, se Hjalmarsson (1996), men torde fortfarande vara av betydelse inom delar av offentlig sektor. Media publicerar regelbundet jämförelser av soptaxor, vatten- och avloppsavgifter etc.

Finansiella nyckeltal, vinst, intäkt, soliditet etc. är av central betydelse i näringslivet, men är svårtolkade och av mera tvivelaktig relevans ur samhällsekonomisk synvinkel. Vinsten i en verksamhet kan kanske tyckas utgöra ett relevant nyckeltal och spegla verksamhetens effektivitet. Paradoxalt nog är det snarast omvänt. En hög vinst kan faktiskt vara tecken på låg effektivitet. För att förstå detta måste vi gå till en ekonomi som fungerar under perfekt konkurrens och där teknologin karakteriseras av konstant skalavkastning, dvs. inga smådrifts- eller stordriftsfördelar. Priserna i en sådan ekonomi är korrekta i den betydelsen att de speglar den relativa knappheten på olika resurser och ingen marknadsmakt finns inbakad priserna. I en sådan perfekt konkurrensökonomi är effektiviteten den högsta möjliga och vinsten noll. Kostnaden per producerad enhet

(inklusive "normal" avkastning på kapitalet) är exakt lika med marknadspriset per producerad enhet. På oligopol- och monopolmarknader innehåller priserna också marknadsmakt och vinsterna är i allmänhet positiva. Kombinationen av marknadsmakt och smådrifts- eller stordriftsfördelar i produktionen av olika varor och tjänster medför stora svårigheter att spåra något klart samband mellan vinstnivå och effektivitet.

Det generella problemet med de flesta indikatorer och nyckeltal är aggregeringsproblemet – att på ett korrekt, i betydelsen vetenskapligt grundat, sätt sammanfatta ett antal olika komponenter i ett enda tal. För en lösning på detta problem är vi hänvisade till den nationalekonomiska (eller operationsanalytiska) produktionsteorin.

3. Produktivitet, effektivitet och teknisk utveckling

Begreppen produktivitet och effektivitet har en skiftande innebörd och en närmare precisering av dessa begrepp kräver en modell. I denna rapport är den grundläggande modellen produktionsteoretisk. Det innebär att produktivitet definieras som kvoten mellan ett produktionsindex (outputindex) och ett produktionsfaktorindex (inputindex), dvs. kvoten mellan produktion och resursförbrukning, medan effektivitet definieras som avståndet från en individuell produktionsenhet till en best-practice-front, dvs. som kvoten mellan faktisk produktivitet och en norm för bästa möjliga produktivitet, givet de restriktioner inom vilka en verksamhet bedrivs. Teknisk utveckling definieras som en förskjutning över tiden i en produktionsfunktion.

Produktivitet

Vid enfaktor-envaruproduktion har produktivitet en klar och entydig innebörd. Vid två eller flera produktionsfaktorer, och/ eller flera produkter, blir produktivetsbegreppet, multifaktor- eller totalfaktorproduktivitet, TFP, mera komplicerat och mindre entydigt.

Nu är det inte själva kvoten mellan produktion och resursförbrukning för en viss produktionsenhet som är av primärt intresse, utan det vi nästan alltid är intresserade av är en jämförelse över tiden eller rummet av produktiviteten, dvs. produktivitetsutvecklingen.

Med termen TFP avses oftast förändringen i TFP, och jag kommer av bekvämlighet i fortsättningen att med TFP avse förändringen i detta mått när ingen risk för missförstånd föreligger. TFP avser i allmänhet en jämförelse mellan två tidsperioder eller regioner för en produktionsenhet (företag, sektor). För att erhålla ett operationellt mått på TFP måste aggregeringen till ett index för produktionsfaktorerna specificeras. En sådan aggregering är produktionsfunktionen.

I den enkla produktionsteoretiska modellen sammanfaller begreppen teknisk utveckling och produktivitetsutveckling. Den tekniska utvecklingen eller produktivitetsutvecklingen är det vi "får över" av produktionsresultatet när produktionsfaktorerna fått "sina andelar" värderade till sina respektive bidrag på marginalen. Mera stringent formulerat är

TFP = Produktionsresultat – den vägda summan av produktionsfaktorerna

där vikterna består av respektive faktors marginalelasticitet. (Värdet på marginalelasticiteten anger hur många procent produktionen ökar när produktionsfaktorn ökar med 1 procent, när alla andra produktionsfaktorer hålls konstanta.) Detta mått kallas ofta det primala måttet på total faktorproduktivitet. TFP i form av residualfaktorn eller restposten är det som återstår att förklara när vi från tillväxten i produktionsresultatet dragit ökningen i arbetskraft, energi, råvaror och kapital vägda med respektive marginalelasticiteter.

På grund av den sk dualiteten mellan produktionsfunktioner och kostnadsfunktioner kan det duala TFP-måttet beräknas. Med kostnadsfunktionen som bas definieras således (det duala måttet på) TFP som takten i kostnadsreduktionen vid given nivå på faktorpriser och produktion. Vid konstant skalavkastning sammanfaller de båda måtten, men vid smådrifts- och stordriftsfördelar skiljer sig måtten något åt.

En viktig fråga är om det är möjligt att mäta TFP, utan aggregering av output och input genom skattning av produktionsfunktioner, enbart genom beräkning av index som endast bygger på observerade data i form av priser och kvantiteter. I en beräkning av ett TFP-index ingår alltid tre grundläggande uppgifter:

- * identifiera och mäta output och input
- * välja vikter för aggregeringen
- * bestämma formen på indexet

Svaret på frågan ovan är att under vissa förutsättningar kan produktivitetsindex beräknas. Vid fri konkurrens och antagande om kostnadsminimering kan vi utnyttja villkoret att värdet av en faktors marginalproduktivitet är lika med dess faktorpris, vilket också innebär att marginalelasticiteten för en produktionsfaktor är lika med dess kostnad som andel av total intäkt. TFP mäts i detta fall som skillnaden mellan produktionsökning och summan av produktionsfaktorernas relativa ökningstakter vägda med utgiftsandelarna som vikter. Detta TFP-index kallas för Divisia produktionsfaktorindex. Detta index kräver ingen kunskap från skattade parametrar i en produktionsfunktion utan bygger enbart på i princip observerbara, men kontinuerligt genererade data.

Problemet med Divisiaindex är att det bygger på kontinuerlig tid medan mätningar av produktivitet oftast sker i diskret tid. Samtidigt gäller att ett diskret index inte är oberoende av approximeringen till diskret tid. Det visar sig då lämpligt att approximera Divisiaindexet med ett Törnqvistindex mellan tidpunkterna t och $t+1$. Vikterna utgörs då av genomsnitten av respektive kostnadsandelar mellan t och $t+1$. Törnqvistindexet torde vara ett av de mest utnyttjade TFP-indexen.

Ett annat index som endast kräver information om input och output är det sk Malmquistindexet till vilket jag återkommer i avsnitt 5.

Produktivitetjämförelser går vanligen ut på jämförelser av produktionsnivåer när vi håller nivån på resursinsatsen konstant. Dessa index kallas produktionsbaserade index (output based index). Om däremot produktionsnivån hålls konstant medan jämförelsen avser skillnader i resursinsats erhålls produktionsfaktorbaserade index (input based index). Skillnaden mellan dessa index är en faktor som speglar graden av skalavkastning.

Produktivitetsindex är ingen ny företeelse utan har en lång tradition i tillämpad forskning. Det som är nytt är däremot den förankring av indexteorin i ekonomisk teori som har växt fram under de senaste decennierna genom upptäckten av att egenskaper hos indextal direkt kan relateras till egenskaper hos de underliggande aggregeringsfunktioner som indexen representerar - produktionsfunktioner, nyttofunktioner etc.

Upptäckten att ett antal indexformler explicit kan härledas från speciella aggregeringsfunktioner har gett en ny teoretisk bas för val mellan olika index. Istället för att fundera över valet mellan ett antal till synes lämpliga produktivitetsindex kan man starta med att specificera en produktionsfunktion med önskvärda egenskaper och härleda den korresponderande indexformeln. Det index som då erhålles kallas exakt för den valda aggregeringsfunktionen.

Diewert (1976) argumenterar kraftfullt för att begränsa övervägandena till aggregeringsfunktioner av typen flexibla former, dvs de som utgör en andra ordningens approximation till en godtycklig aggregeringsfunktion. Diewert kallar de index, som är exakta för flexibla aggregeringsfunktioner, superlativa. Två av de mest uppmärksammade superlativa indexen är

- Fishers ideala index som utgör det geometriska medelvärdet av Laspeyre- och Paascheindexen. Detta index är ett sk idealt index vilket innebär att det uppfyller ett antal krav av konsistensstyp som är rimliga att ställa på ett index. Begreppet myntades av Fisher (1922) som i en omfattande genomgång av olika index fann att det geometriska genomsnittet av Laspeyre och Paasche var ett idealt index.
- Törnqvists translogindex

Det första indexet är exakt för en viss kvadratisk flexibel form och det andra för en homogen translog produktionsfunktion.

Möjligheten att utnyttja TFP-index istället för skattning av produktionsfunktioner beror på antagandet om kostnadsminimering samt förutsättningen om konstant skalavkastning. Vid variabel skalavkastning undgår vi inte problemen med att skatta en produktionsfunktion för att få ett mått på graden av stordriftsfördelar.

TFP-index är trots allt relativt krävande att beräkna och det som ofta lyftes fram i den allmänna debatten är partiella produktiviteter, dvs. förhållandet mellan produktionsresultat och en specifik produktionsfaktor. Speciellt arbetsproduktivitet men även energiproduktivitet har varit mycket populära som produktivitetsmått under senare år eller, som diskuterades i avsnitt 2, inversen av energiproduktiviteten, det specifika åtgångstalet eller inputkoefficienten.

En partiell produktivitet är emellertid svårtolkad av flera olika skäl:

- De erhållna produktivitetsmått är ej konstanter utan variabler eller funktioner som varierar inte bara med insatsen av den studerade faktorn utan med nivån på samtliga produ
- En partiell faktorproduktivitet uppnår sitt maximum då dess marginalelasticitet antar värdet 1, då alla andra produktionsfaktorer hålls konstanta. Samtidigt gäller att totalfaktorproduktiviteten uppnår sitt maximum, i optimal skala, då summan av marginalelasticiteterna antar värdet 1. Detta innebär att när
- När totalfaktorproduktiviteten uppnår sitt maximum så avviker samtliga partiella produktivitet från sina respektive maxima. Vi kan således inte resonera som så att vi först ser till att arbetsproduktiviteten uppnår sitt maximum, sedan energiproduktiviteten, därefter kapitalproduktiviteten osv. eftersom dessa produktivitet inte är oberoende av varandra.
- Vid flervaruproduktion ökar tolkningsproblemen eftersom även produktionen av de olika produkterna är inbördes beroende.
- Det är mycket riskabelt att dra slutsatser om totalproduktivitetens utveckling på basis av partiella produktivitet. Även om samtliga partiella produktivitet visar en förskjutning i produktivitetshöjande riktning kan man härav inte dra slutsatsen att också totalfaktorproduktiviteten förbättrats. Ännu svårare blir det att dra slutsatser när de partiella produktivitetsfördelningarna för olika produkter eller produktionsfaktorer förskjuts i olika riktning.

Produktiviteten kan beräknas på olika nivåer från den nationella nivån ned till enskilda produktionsprocesser. För ekonomin som helhet beräknas ofta mått på arbetsproduktivitet som kvoten mellan produktionsresultat och ett mått på arbetsinsatsen, antal arbetade timmar, antal årsverk, antal sysselsatta osv. Produktion per sysselsatt utnyttjas ofta som mått på levnadsstandard. Energiproduktivitet på nationell nivå beräknas också ofta i samband med energianalyser.

På branschnivå utnyttjas beräkningar av arbetsproduktivitet för jämförelser av olika branschens lönebetalningsförmåga och vid beräkning av vinsterna av att omfördela arbetskraft mellan olika sektorer. Beräkningar av energiproduktivitet används på motsvarande sätt för att analysera konsekvenserna av energiprishöjningar.

På företagsnivå utnyttjas också produktivetsmått för analyser av konkurrenskraft i jämförelse med andra företag och ofta som nyckeltal för jämförelser av olika avdelningar inom ett företag. På företagsnivå karakteriseras ofta ett företag med hög produktivitet som effektivt.

Ofta utnyttjas arbetsproduktivitet, som produktivetsmått med motiveringen att detta skulle bygga på mindre restriktiva förutsättningar eller vara enklare att tolka än totala produktivetsmått. Detta är absolut inte fallet. Partiella produktivetsmått är inte enklare att tolka eller mindre restriktiva ifråga om underliggande antaganden än totala produktivetsmått, men eftersom problemen sopas under mattan kan de lätt framstå som enklare.

Ytterligare ett problem med arbetsproduktivetsmålet (och i viss mån med energiproduktivetsmålet) är att kortsiktiga variationer i detta mått domineras av cykliska faktorer, dvs. arbetsproduktiviteten varierar med kapacitetsutnyttjandet. För att kunna jämföra utvecklingen av arbetsproduktiviteten mellan olika sektorer eller olika länder är det därför nödvändigt att konjunkturjustera måttet. Utan en fullständig dynamisk faktorefterfrågemodell blir justeringar för kapacitetsutnyttjande ad hoc-mässiga.

Produktivitet vs teknisk utveckling vs effektivitet

Effektivitetsbegreppet är mindre väl avgränsat än produktivetsbegreppet. I dagligt tal används termen effektivitet i allmänhet som synonym med produktivitet, och någon klar gränslinje mellan de båda begreppen existerar inte. I läroböcker i nationalekonomi definieras produktivitet ofta som att göra saker på rätt sätt medan effektivitet avser att göra rätt saker, dvs. det som i nationalekonomisk terminologi kallas allokeringseffektivitet eller priseffektivitet. Allokeringseffektivitet råder då $\text{pris} = \text{kortsiktig marginalkostnad}$ på en marknad. Det är då varor och tjänster produceras i rätta volymer dvs. konsumenternas betalningsvilja för de sist konsumerade enheterna sammanfaller med kostnaderna för att producera dessa.

Vi kan också definiera begreppet dynamisk effektivitet men i två olika betydelser. Det kan dels avse hur väl investerings- eller kapacitetsexpansionsprocessen fungerar i en bransch, dels avståndet mellan faktisk och optimal volym och sammansättning på forskning och utveckling inom en bransch.

Det effektivitetsbegrepp som står i fokus i denna skrift benämnes ofta produktionseffektivitet (productive efficiency) eller teknisk effektivitet (technical efficiency) och avser avståndet från produktionsenheterna i en bransch till en best-practice-front.

Inom den ekonomiska forskningen har begreppen sin förhistoria och litteratur, och förenklat kan man säga att produktivitetsbegreppet har en stark anknytning till den neoklassiska produktionsteorin och beräkning av produktivitetsutvecklingen för en sektor över tiden, medan effektivitetsbegreppet har en starkare anknytning till operationsanalys och mera grundläggande, delvis axiomatisk, produktionsteori, delvis i kombination med utvecklingen av ekonometriska metoder för skattning av sneda fördelningar.

I denna rapport kommer jag att i viss utsträckning skilja på begreppen och framförallt tala om effektivitet i samband med jämförelser av produktivitet mellan produktionsenheter relativt en teknologifront. Med effektivitet avses då avståndet mellan en produktionsenhet och en referensteknologi eller förhållandet mellan insatta resurser och produktionsresultat för en viss produktionsenhet i förhållande till en referensteknologi. Studier som analyserar spridningen i totalproduktivitet mellan olika produktionsenheter i en sektor relativt en frontteknologi kommer jag därför att benämna effektivitetstudier medan jag med produktivitetsstudier oftast avser analyser av produktivitetsutveckling över tid för en sektor eller direkta jämförelser i rummet mellan produktionsenheter utan relation till en frontproduktionsfunktion. Effektivitet avser således jämförelser av produktionsenheter relativt en teknologifront medan produktivitet avser binära jämförelser mellan produktionsenheter i tiden eller rummet.

Begreppen produktivitetsutveckling och teknisk utveckling är nära förbundna med varandra och även här är terminologin svävande. Vanligtvis definieras teknisk utveckling mera precist som produktionsfunktionens vertikala förskjutning över tiden vid konstant mängd produktionsfaktorer medan totalfaktorproduktivitet definieras som förändringen i förhållandet mellan ett index för produktionsnivån och ett index för resursinsatsen. En positiv teknisk utveckling vid given nivå på resursinsatsen medför alltid en produktivitetsökning och med en produktionsfunktion som bas och icke kapitalbunden teknisk utveckling kan teknisk utveckling identifieras med produktivitetsutveckling. Produktivitet kan dock mätas utan att någon produktionsfunktion införs.

I den ekonomisk-politiska debatten kallas den tekniska utvecklingen ofta för residual-faktorn eller restposten. Terminologin går tillbaka till Solow (1957) som klargjorde skillnaden mellan en förskjutning i produktionsfunktionen och förändringar utmed produktionsfunktionen orsakade av förändringar i insatsen av produktionsfaktorer. Restposten är det som återstår att förklara när vi från tillväxten i produktionsresultatet dragit ökningen i sysselsättning, kapital, energi etc. Den har därför ibland men något oegentligt betecknats som ett mått på vår bristande kunskap. Stor möda har nämligen ägnats åt att förklara och dela upp restposten i olika komponenter sk growth accounting eller åt att utöka den enkla modellen så att en hel del av restposten inkluderas i modellen t ex genom att måttet på arbetskraft även inkluderar arbetskraftens kvalitet. Med en viss tillspetsning skulle man kunna säga att den berömda restposten knappast existerar längre i modern produktivitetsforskning. Som en sammanfattande beteckning på den del av tillväxten som ej är att hänföra till ökad resursinsats har emellertid begreppet fortfarande stor användning.

Med teknisk utveckling avses inom produktivitetsforskningen således den del av en produktivitetstillväxt som ej låter sig förklaras av förändringar i insatser av produktionsfaktorer dvs. förändringen över tiden i produktionen då nivån på alla produktionsfaktorer håll konstant.

På grund av dualiteten mellan produktionsfunktioner, distansfunktioner, kostnadsfunktioner och vinstfunktioner kan dessa definitioner också överföras till de duala funktionerna. Dualiteten innebär att under vissa förutsättningar ger skattning av duala funktioner samma information om teknologin som skattning av produktionsfunktioner. Valet mellan olika ansatser bestäms framförallt av tillgång till data samt ekonometriska överväganden. I denna rapport representeras teknologin av en produktionsfunktion och med best-practice avses bäst observerad teknologi (i vid mening).

Effektivitets- och produktivitetsanalyser är snarare komplement än substitut. Medan effektivitetsanalyser ger information om rationaliseringspotentialen i en sektor vid en viss tidpunkt ger produktivitetsanalyser information om produktivitetsförändringar över tid eller skillnader i produktivitet mellan olika sektorer eller regioner. I den utsträckning som en viss verksamhet har likartad karaktär i olika regioner, t ex bankverksamhet i

olika europeiska länder, kan också effektivitetsanalys utnyttjas för att analysera skillnader mellan länder.

Effektivitet

Utgångspunkten för effektivitetsanalyser är här ett datamaterial för mikroenheterna i en sektor. På basis av dessa data konstrueras en teknologifront som referensteknologi i form av bästa observerad teknologi. Mikroenheter som ligger på denna teknologifront benämnes fullt effektiva enheter. (Som framgår av avsnitt 5 så kan dock vissa fullt effektiva enheter vara mera "fullare" effektiva än andra.) Det finns en lång rad konkurrerande metoder för konstruktion av teknologifronten; se Hjalmarsson (1991).

Med en frontfunktion som benchmark erhålles, med olika sätt att mäta, olika effektivitetsmått och därmed effektivitetsrangordningar av produktionsenheterna inom en sektor. Ett mått ger information om kvoten mellan observerad produktion och den produktion som skulle kunna uppnås med best-practice teknologi, ett alternativt mått ger kvoten mellan produktionsfaktorinsats i best-practice och observerad mängd produktionsfaktorer, givet produktionsnivån. Det första effektivitetsmättet kan benämnas produktionsökningseffektivitet (output-oriented efficiency) medan det senare kan kallas resursbesparingseffektivitet (input-oriented efficiency). Ett tredje effektivitetsmått, skal-effektivitet, mäter kvoten mellan resursåtgång i optimal skala och observerad resursåtgång, dvs. måttet ger information om hur mycket resurser som kan sparas genom att producera med bästa teknik i optimal skala jämfört med att producera under fronten och med en produktionsvolym som avviker från optimal skala.

Frontproduktionsfunktionen ska ligga nära de bästa enheterna i sektorn. Nära är ett relativt begrepp, och det existerar en rad olika modeller och metoder för skattning av frontproduktionsfunktioner. Modellerna kan klassificeras, dels efter hur själva frontproduktionsfunktionen skattas ekonometriskt eller beräknas matematiskt, dels efter vilka antaganden som görs om spridningen i effektivitet mellan anläggningarna i sektorn. Det vi fokuserar på här är radiala mått, dvs. vi ser på proportionella variationer i inputs och outputs. Till skillnad från de icke-radiala måtten har de radiala en klar tolkning i termer av potentiell resursbesparing respektive potentiell produktionsökning.

Eftersom frontfunktionen utgör benchmark för alla produktionsenheter är det av stor betydelse hur denna ”lägges fast”. Det debatten mycket handlat om är inverkan på effektivitetsmått av sk outliers, dvs. olika typ av extremobservationer. Om dessa behandlas som ”äka” eller korrekta observationer (deterministisk front) eller som slumpmässiga mätfel (stokastisk front) har stor betydelse för såväl nivån på effektivitetsmått som spridningen i effektivitet mellan olika enheter i en sektor. En annan skiljelinje, som också påverkar såväl nivån på effektivitetsmått som spridningen i effektivitet, går mellan parametriska och icke-parametriska fronter. En parametrisk front, dvs. en analytisk produktionsfunktion som i SFA, sluter sig mindre tätt om observationerna än en icke-parametrisk front som i DEA.

4. Parametriska stokastiska fronter, SFA

Den ursprungliga parametriska frontproduktionsfunktionsmodellen går nu under beteckningen deterministisk frontfunktion. Att den är deterministisk innebär att alla avvikelser från fronten tolkas som ineffektivitet. Normalt skattas fronten genom att summan av avstånden från enheterna till funktionen minimeras, med restriktionen att ingen produktionsenhet får vara bättre än fronten. Modellen tar ej hänsyn till slumpfaktorer, och avståndet mellan frontfunktion och en enskild observation utgör ett mått på enhetens relativa effektivitet. Någon förutsättning om effektivitetsfördelningens utseende görs ej i grundmodellen. Antalet fullt effektiva enheter, frontenheterna med effektivitetstalet 1, bestäms av problemets matematiska karaktär (och uppgår högst till det antal parametrar som ska bestämmas).

Tvärsnittsmodeller och tvärsnittsproblem

Efter en trevande inledning med parametriska deterministiska fronter (Deterministic Frontier Analysis, DFA) med start i Aigner & Chu (1968) skedde det stora genombrottet på den parametriska stokastiska sidan 1977, i tre separata artiklar, Aigner, Lovell and Schmidt (1977), Meeusen and van den Broeck (1977) samt Battese and Corra (1977). I dessa artiklar introducerades den sk ”composed-error-” (CE) modellen.

Den grundläggande filosofin här är att det är rimligt att räkna med att den spridning i effektivitet mellan produktionsenheter som vi observerar i en sektor är ett uttryck såväl för äkta ineffektivitet som för slumpfaktorer. I CE-modellen delas avståndet mellan front och produktionsenhet upp i en symmetrisk (normalfördelad) slumpterm och en icke-symmetrisk (snedfördelad) effektivitetsterm. Vid en front som skattats med CE-ansatsen kommer vissa produktionsenheter att ligga ovanför fronten beroende på slumpmässiga faktorer, medan detta är omöjligt vid en deterministisk front.

I CE-modellen påverkas effektivitetsmåten dels av den specificerade produktionsfunktionen (t ex Cobb-Douglas eller Translog) dels av den antagna sneda effektivitetsfördelningen. Så t ex specificerade, Aigner, Lovell and Schmidt (1977) en trunkerad normalfördelning, medan Meeusen and van den Broeck (1977) antog att effektiviteten var exponentialfördelad. Empiriskt visar det sig att slumptermen ”suger upp” en stor del av variationen mellan produktionsenheter. Rimligheten i detta kan ifrågasättas speciellt om data är av god kvalitet. Empiriskt visar det sig också att effektivitetsfördelningen är relativt robust med avseende på olika antaganden om effektivitetsfördelningen (halv-normal, trunkerad normal, exponential, gamma); se t ex Greene (1990) eller Kumbhakar & Lovell (2000).

Med OLS är det möjligt att skatta struktureffektiviteten för en sektor. Däremot är det inte möjligt att erhålla konsistenta individuella effektivitetsmått. En del av forskningen har därför varit inriktad mot att finna konsistenta estimat av effektivitetsmåten och signifikanstest. Ett viktigt bidrag är Horrace & Schmidt (1996) som visade hur man kan beräkna konfidensintervall för punktestimaten på teknisk effektivitet. Denna metod utnyttjades bl a av Bera & Sharma (1996) samt Hjalmarsson, Kumbhakar & Heshmati (1996) för två olika punktestimatorer.

Paneldatamodeller: Förändringar i produktivitet och effektivitet

Den i många avseenden mest betydelsefulla utvecklingen inom SFA har gällt paneldatamodeller. Med tillgång till paneldata, dvs. data för T år och N produktionsenheter, mildras vissa av nackdelarna med CE-modellen, samtidigt som nya möjligheter öppnar sig på modellsidan. Paneldatamodellen har ett antal attraktiva egenskaper;

speciellt erbjuder den möjlighet att analysera förändringar i såväl produktivitet som effektivitet över tiden och särskilja den tekniska utvecklingen från förändringar i effektivitet. Det är heller inte nödvändigt att anta en specifik effektivitetsfördelning, utan denna kan ersättas med antagande om konstant effektivitet eller en viss skevhet.

I den grundläggande paneldatamodellen är residualen uppdelad i fyra komponenter: enhetsspecifik komponent + tidsspecifik komponent + effektivitet + slumpterm. Medan den tidsspecifika komponenten är ett uttryck för den tekniska utvecklingen reflekterar den enhetsspecifika komponenten konstanta, dvs. tidsberoende, skillnader mellan enheterna, typ jordmån i jordbruk eller organisation i företag. Den enhetsspecifika komponenten har därför i varje fall delvis karaktär av effektivitetskomponent som i vissa fall borde adderas till den tidsvarierande effektivitetskomponenten (som har både enhets- och tidsindex). Ett typiskt exempel på en renodlad paneldatamodell är Kumbhakar & Hjalmarsson (1995). Observera att en panel inte behöver vara fullständig utan kan mycket väl vara roterande (då samplet gradvis förnyas så att en viss enhet ingår under tex 3 år) eller obalanserad (då vissa enheter ingår under kortare tidsperioder än hela sampleperioden); se t ex Heshmati (1994) och Hjalmarsson, Kumbhakar & Heshmati (1996) för obalanserade paneler och Kumbhakar & Heshmati (1995) för roterande paneler.

Paneldatamodellen är inte entydig utan öppnar upp för ett stort antal olika modeller och skattningsmetoder. En central fråga i paneldatamodellering är specifikationen av effektivitet och teknisk utveckling. Istället för en tidsspecifik komponent kan den tekniska utvecklingen beaktas genom att specificera parametrarna i produktionsfunktionen som funktioner av tiden.

Paneldatamodellen i effektivitetslitteraturen introducerades av Pitt & Lee (1981). Medan 1980-talets paneldatamodeller antog tidsberoende effektivitet fick vi under 1990-talet ett antal modeller med tidsvarierande effektivitet; se Cornwell, Schmidt & Sickles (1990), Battese & Coelli (1992) och Lee & Schmidt (1993). Dessa modeller

avslöjar utvecklingen över tiden för effektiviteten hos mikroenheterna och svarar på frågor om stabiliteten i effektiviteten och huruvida ineffektivitet är ett permanent eller övergående tillstånd.

Cornwell, Schmidt & Sickles (1990) specificerar den enhetsspecifika effektiviteten som en kvadratisk funktion av tiden, olika för varje mikroenhet, medan Lee & Schmidt (1993) inte specificerar någon bestämd funktion men däremot antar att tidsmönstret för effektiviteten är detsamma för alla mikroenheter. Battese & Coelli (1992) inför en mera restriktiv funktion också med tidsmönstret detsamma för alla mikroenheter. Kumbhakar, Heshmati & Hjalmarsson (1997) presenterar resultaten av en jämförelse mellan de olika modellerna skattade på samma datamaterial. Resultaten visar sig i hög grad vara modellberoende. Ytterligare en specifikation av tidsvarierande effektivitet föreslogs i Kumbhakar & Hjalmarsson (1993). Här uppdelas effektiviteten i en enhetsspecifik komponent som fångar heterogeniteten på producentsidan och en tidsspecifik komponent som fångar den tekniska ineffektiviteten.

Förklaring av ineffektivitet

En kontroversiell fråga i effektivitetsforskningen har varit hur bestämningsfaktorerna bakom ineffektivitet ska modelleras. Längre dominerade den sk tvåstegsmetoden, dvs. i ett första steg beräknades eller estimerades effektiviteten för enheterna och i ett senare steg utnyttjades effektivitetsmåten i en regression med ett antal exogena faktorer som förklaringsvariabler. Alternativet är en enstegsmetod, introducerad av Reifschneider & Stevenson (1991), som inkluderade de exogena faktorerna direkt i skattningen av den stokastiska fronten. Denna ansats utnyttjades i Hjalmarsson, Kumbhakar & Heshmati (1996) för beräkning av konfidensintervall för effektivitetsmåten. Ett par av de mest välkända enstegsmodellerna är Kumbhakar, Ghosh & McGuckin (1991) och Battese & Coelli (1995). I dessa modeller parametriseras *genomsnittet* hos den pre-trunkerade fördelningen som ett sätt att fånga exogena faktorerers inflytande på effektiviteten. Med hjälp av Monte Carlo-analys visade Schmidt & Wang (2002) att enstegsmetoden är överlägsen tvåstegsmetoden ur statistisk synvinkel.

Ett annat bidrag på den ekonometriska sidan är hanteringen av heteroskedasticitet i slump- och effektivitetstermen (dvs. när slumptermerna inte har samma varians i hela stickprovet); se t ex Caudill, Ford & Gropper (1995), som angriper problemet med heteroskedasticitet genom parametrisera *variansen* hos den trunkerade fördelningen. Wang (2002) knyter samman denna ansats med ansatsen att analysera exogena effekters inflytande på ineffektiviteten.

Funktionsformer

Även om Translog-funktionen kommit att dominera som specifikation i SFA har en lång rad andra funktionsformer också utnyttjats i empiriska analyser. Kumbhakar, Heshmati & Hjalmarsson (2002) utnyttjar en sk ”labour requirement frontier” med arbetskraft som input på vänstersidan och en kvadratisk flexibel form med kapital, output och tidstrend på högersidan. En liknande ansats utnyttjas av Bjurek & Hjalmarsson (1995) med en deterministisk kvadratisk labour requirement frontier. En sådan bakvänd produktionsfunktion kan vara en lämplig ansats när arbetskraften dominerar som input i produktion av flera output; typiskt i tjänsteproduktion.

Ett annat exempel är utnyttjande av en Fourier flexibel funktion (som innehåller trigonometriska komponenter) som är mera flexibel än Translog. En jämförelse mellan denna och Translog ges i Huang & Wang (2004).

Gagné och Oulette (2002) jämför tre flexibla funktioner, Translog, Symmetrisk McFadden och Symmetrisk Generaliserad Barnett, när teknisk utveckling specificeras som en tidstrend i en stokastisk frontmodell. Resultaten illustrerar hur känsliga resultaten är för val av funktionsform och hur ineffektivitet speciellt påverkar mätning av stordriftsfördelar.

En annan typ av jämförelse görs av Kumbhakar, Heshmati & Hjalmarsson (1999) som jämför sex olika specifikationer av teknisk utveckling och produktivitsutveckling i en traditionell neoklassisk produktionsfunktion av translogtyp som grundläggande funktionsform. Även här är resultaten relativt känsliga för val av specifikation, dvs. exakt hur den tekniska utvecklingen specificeras matematiskt.

Stordriftsfördelar och specialiserings-/diversifieringsfördelar

Det existerar ingen klar distinktion mellan varor och tjänster, men i stora delar av tjänsteproduktionen har arbetskraften en central roll, medan denna roll innehas av kapitalstocken i varuproduktionen. I varuproduktionen är teknologin ofta bunden till kapitalstocken, varför det krävs nyinvesteringar för ökad produktivitet (kapitalbunden eller embodied teknisk utveckling), medan effektivitet i tjänsteproduktionen i högre grad bestäms av organisation och institutionella faktorer. Informationsteknologin är givetvis mycket betydelsefull inom stora delar av tjänstesektorn men det är den också inom stora delar av varuproduktionen.

Medan stordriftsfördelar karakteriserar nästan alla sektorer av varuproduktion är resultaten av empiriska undersökningar av tjänsteproduktion mera tvetydiga och svårtolkade. Eftersom tjänsteproduktion ofta har karaktär av flervaruproduktion är det viktigt att skilja mellan stordriftsfördelar (economies of scale) och specialiseringsfördelar (economies of scope) samt i vissa fall täthetsfördelar (economies of density). (För definition av stordriftsfördelar vid multiple output, se Starrett (1977)). Teori- och metodproblemen är här betydande.

A priori kan man förvänta sig att stordriftsfördelarna inom tjänsteproduktionen i allmänhet är små medan specialiseringsfördelarna kan vara betydande. Det som genererar stordriftsfördelar i varuproduktionen, kapitalstockarna, saknas i stor utsträckning i tjänsteproduktionen. Däremot kan transaktionskostnaderna vara av stor betydelse i vissa tjänstesektorer (t ex bank och försäkring) varför informationsteknologins framväxt kan komma att öka stordriftsfördelarna genom minskade transaktionskostnader. Eftersom det – såvitt jag vet – saknas en översikt över de empiriska resultaten citeras några få enskilda studier här. Battese, Heshmati & Hjalmarsson (2000) liksom Kumbhakar Heshmati & Hjalmarsson (2002) finner betydande stordriftsfördelar i svensk bankverksamhet. Kumbhakar & Hjalmarsson (1995) finner också betydande stordriftsfördelar i svensk socialförsäkringsverksamhet, och liknande resultat erhålls av Kumbhakar & Hjalmarsson (1997) för svensk eldistribution. Det senare resultatet kan jämföras med Salvanes & Tjøtta (1994) som

med en annan ansats finner mycket små stordriftsfördelar men betydande economies of density i norsk eldistribution.

En sektor med betydande stordriftsfördelar synes också vara detaljhandeln med Wal-Mart som paradexempel. Det kan dock också vara så att Wal-Marts framgång beror på ett snabbt utnyttjande av ny teknologi utan att stordriftsfördelar har någon betydelse. Det senare visar sig vara fallet i en studie av svensk detaljhandel där hypotesen om konstant skalavkastning inte kunde förkastas; Se Maican & Orth (2008).

5. Data Envelopment Analysis

Medan Koopmans (1951) definierade teknisk effektivitet var det Farrell (1957) som bidrog med den första empiriska analysen av effektivitet inom en sektor. Han använde sig av linjärprogrammeringsteknik och i en komplicerad modell (som varit mödosam att replikera) beräknades avståndet från individuella produktionsenheter till teknikfronten för samtliga enheter i ett enda steg. Charnes, Cooper & Rhodes (1978), ofta benämnd CCR-modellen, erbjöd en mycket mera attraktiv beräkningsmetod, en relativt enkel metod där effektiviteten för varje produktionsenhet beräknades separat från övriga produktionsenheter i ett (till synes) enkelt linjärprogrammeringsproblem.

Flervaruproduktion utgör i denna modell inget problem. CCR-modellen var baserad på konstant skalavkastning, CRS. I ett senare arbete, Banker, Charnes & Cooper (1984), ofta benämnd BCC-modellen, introducerades variabel skalavkastning, VRS.

DEA blev snabbt en extremt populär metod för effektivitetsmätning. Till skillnad från regressionsmetoder har den blivit populär såväl inom nationalekonomin som inom Operationsanalys (OR) och Management Science (MS); se Forsund & Sarafoglou (2005) för en analys av skillnader i inriktning mellan bidragen från OR/MS och nationalekonomi. Beräkningsmässigt är DEA en (på ytan) mycket enkel (men i grund och botten mycket komplicerad) modell och kräver, i sin grundversion, inga avancerade statistiska eller ekonometriska kunskaper. Den hanterar enkelt flervaruproduktion. Dessutom har det utvecklats användarvänlig software.

Akademiskt var det framförallt inom OR och MS som DEA snabbt fick stor popularitet, medan nationalekonomer var mer kritiska till ansatsen. Den *ursprungliga* huvudkritiken kan sammanfattas i tre punkter:

- DEA är en icke-parametrisk metod. Ingen produktions-, kostnads- eller vinstfunktion skattas. DEA genererar styckvis linjära teknologier. Därför är det omöjligt att utnyttja traditionell marginalanalys eller härleda, skala- och substitutionselasticiteter.
- Eftersom DEA inte beaktar mätfel i data blir resultaten starkt känsliga för extremobservationer. LP-formuleringen styr dessutom antalet fullt effektiva enheter (linjärprogrammeringens fundamentalteorem).
- DEA är en deterministisk ansats. LP-lösningen genererar inga standardavvikelser och ger inget utrymme för hypotestestning. Varje avvikelse från fronten klassificeras som ineffektivitet.

Den första och sista punkten är betydligt mindre relevanta idag. Låt mig börja med den första punkten.

Från OR till neoklassisk produktionsteori

Även om DEA är en modell för en styckvis linjär (piecewise linear) teknologi så är det en produktionsfunktionsmodell. Som produktionsfunktion är den givetvis öppen för samma typ av analys som den neoklassiska produktionsfunktionen. Då existerar marginalproduktiviteter, skal- och substitutionselasticiteter, även om dessa funktioner nu inte beter sig på samma ”smidiga” sätt som i standard neoklassisk produktionsteori. Intresset för dessa begrepp, med undantag för optimal skala, har däremot inte varit så stort, speciellt inte från den ena delen av forskningsfältet, OR/MS, som mera fokuserat på de matematiska egenskaperna hos DEA och betydligt mindre på de ekonomisk-teoretiska egenskaperna; se dock Cooper, Park & Ciurana (2000) som analyserar substitutionselasticiteten i DEA samt Forsund & Hjalmarsson (2004a) och (2004b) som presenterar en översättning av DEA till de neoklassiska begreppen och speciellt skala-

elasticitet och optimal skala och visar hur dessa begrepp kan beräknas i DEA; se också Forsund et al (2007).

Stordriftsfördelar och optimal skala

Eftersom graden av stordriftsfördelar är en central aspekt vid en teknologi, har beräkning av stordriftsfördelar i DEA kommit att generera en mycket omfattande, delvis kontroversiell och enligt min uppfattning delvis mycket förvirrande, litteratur. Den främsta orsaken till denna förvirring är, enligt min uppfattning, en bristande insikt i grundläggande neoklassisk produktionsteori. En översikt över denna litteratur ges dels i Banker et al (2004) dels i Forsund & Hjalmarsson (2004b). Med undantag för Forsund & Hjalmarsson (2004a) och (2004b) samt Forsund et al (2007) har DEA-litteraturen i stort sett varit inriktad mot kvalitativa beräkningar av stordriftsfördelar, dvs. produktionsenheterna har klassificerats i grupper med skal-elasticitet större eller mindre än 1. Detta kan enkelt ske genom att inspektera summan av DEA-vikterna för varje enhet. Om denna summa är mindre än 1 har vi avtagande skalavkastning, om den är större än 1 har vi tilltagande skalavkastning.

Eftersom DEA är en icke-parametrisk metod är de empiriska resultaten extremt känsliga för mikroenheternas lokalisering i input-output-rummet, dvs. hur enheterna fördelar sig med avseende på såväl resursinsatsens som produktionens storlek och sammansättning. Ett exempel på detta är begreppet optimal skala, dvs. den produktionsskala som ger högsta möjliga produktivitet och lägsta genomsnittliga produktionskostnader. Forsund & Hjalmarsson (2004a) visar dels hur optimal skala och värdet på skal-elasticiteten kan beräknas för olika enheter, dels att nivån på optimal skala är extremt känslig för valet av faktorproportioner; se också Hjalmarsson, Kumbhakar & Heshmati (1996) där samma slutsats kan dras. För en given sektor är variationen i optimal skala nästan lika stor som variationen i produktion mellan mikroenheterna. Den något deprimerande slutsatsen som dras är att optimal skala inte är ett ekonomiskt meningsfullt begrepp i DEA, dvs. det saknar fullständigt policyrelevans. Skattning av optimal skala kräver helt enkelt en produktionsfunktion med mera struktur än den DEA erbjuder.

Extrema observationer och känslighetsanalys

DEA är som metod baserad på extremobservationer. Känsligheten för extremobservationer medför att känslighetsanalys är ett naturligt inslag i denna typ av effektivitetsanalys. Huvudproblemet består i att skilja relevanta extremobservationer från irrelevanta.

Eftersom antalet fullt effektiva enheter bestäms av linjärprogrammeringens fundamentalteorem genererar DEA ofta ett stort antal fullt effektiva enheter. Medan vissa av de effektiva enheterna ”spänner upp” fronten, är andra fullt effektiva enheter fullt effektiva enbart på grund av metoden. Om DEA appliceras på en sektor med relativt få mikroenheter blir antalet fullt effektiva enheter oproportionerligt stort. En relevant fråga är därför om vissa effektiva enheter kan betraktas som mera effektiva än andra, dvs. hur ser en rangordning av de effektiva enheterna ut? Denna fråga har undersökts i flera bidrag. Ett förslag, Andersen & Petersen (1993), är att rangordna effektiva enheter genom att beräkna ett mått på sk supereffektivitet för att identifiera de mest inflytelserika effektiva enheterna. Wilson (1995) tog ett steg vidare och föreslog, som ett sätt att identifiera enheter med ”alltför stort” inflytande, en vägning av supereffektivitetsmåttan med genomsnittsförändringen i effektivitetsmåttan för de resterande enheterna när en viss effektiv frontenhet tagits bort.

Torgersen, Førsund & Kittelsen (1996) föreslår en metod som går ut på att rangordna effektiva enheter efter vilket inflytande dessa har som benchmark för ineffektiva enheter. Mikroenheter som tjänar som föredöme (sk peers) för ineffektiva enheter rangordnas som mera effektiva än enheter som inte har denna roll (sk self evaluators). Rangordningsmåttet är den totala viktade produktionsökningen, eller resursbesparingen, som kan erhållas från de ineffektiva enheterna med de effektiva som benchmark.

Edvardsen, Førsund & Kittelsen (2003) och (2008) går ett steg vidare, men nu är det inte inflytande utan klassificering av peers utifrån den information statusen ger om deras egen effektivitet som är i fokus. Här presenteras en metod för att klassificera fullt effektiva enheter som interna (interior) respektive externa (exterior) och som ”isolerade” (self evaluators) respektive inflytelserika (peers). De isolerade enheterna

återfinns ofta i "utkanten" av teknologin¹ men kan också ligga i det inre. Att skilje de externa från de interna är viktigt eftersom de externa bidrar med mycket mindre kunskap än de interna. Vid en sk "second-stage regression", dvs. en efterföljande regression för att förklara skillnaderna i effektivitet mellan enheter är det en stor fördel att inte ha med de externa effektiva enheterna som bara bidrar till att snedvrider resultaten. De inflytelserika (peers) kan också karakteriseras på liknande sätt vilket kan vara värdefullt vid tolkningen av DEA-resultat; speciellt kan de ge information om vilka delar av fronten som bäst stöds av data.

Antal enheter och struktureffektivitet

Att DEA är en datadriven metod innebär också att effektivitetsmått är känsliga för såväl antal enheter som dessas lokalisering i input-outputrummet. Det finns en tendens att effektivitetsmålet för en viss enhet tenderar att sjunka när antalet enheter ökar i en sektor. Orsaken till detta är att när antalet enheter ökar så närmar sig den front som konstrueras av DEA asymptotiskt mot den "sanna" fronten (dvs. gradvis allt närmare för

att vid oändligt många enheter sammanfalla med den sanna fronten); se Diewert (1993) och Banker (1989). Detta innebär också att om vi jämför struktureffektiviteten för två sektorer med olika antal mikroenheter så kommer den med färre enheter att i allmänhet uppvisa en högre effektivitet än den med många enheter. A priori skulle man förvänta sig det omvända eftersom flera företag i en sektor borde leda till ökad konkurrens och därmed ökad effektivitet. Vi ser här att en "teknisk" egenskap hos DEA leder till resultat som lätt kan leda tanken i fel riktning. För att få en rättvisande jämförelse av effektiviteten i två sektorer med olika antal produktionsenheter krävs en normering med avseende på antalet enheter. Hur en sådan justering ska gå till visas av Zhang & Bartels (1998), som med Monte Carlo-simulering analyserar sambandet mellan antalet mikroenheter i en sektor och dess struktureffektivitet. I en kommentar visar Staat (2001) att samma problem vidlåder ett antal olika DEA-modeller med mera implicit jämförelse av subsamples av olika storlek (t ex modeller för icke-diskretionära variabler).

¹ "Efficient by default"; se Tulkens (1993).

Utvecklingen av metoder, speciellt sk bootstrapping, för beräkning av konfidensintervall har på sätt och vis löst problemet med känsligheten för antalet enheter (vilket också föreslogs av Zhang & Bartels). I en sektor med få enheter kan vi förvänta oss mycket stora konfidensintervall för effektivitetsmåten.

Statistiska egenskaper och statistisk signifikans

En orsak till den, i varje fall inledningsvis, bristande entusiasmen från nationalekonomer för DEA är avsaknaden av statistiska mått för osäkerheten i punktestimaten för effektivitetsmåten. En relativt färsk rapport Bird et al (2005), från "British Working Party of Performance Monitoring in the Public Sector" har som en av sina rekommendationer att prestationsmått (performance measures) alltid ska inkludera mått på osäkerhet.

Även deterministiska modeller har emellertid statistiska egenskaper. Så är också fallet med DEA. En programmeringsmetod är ju ett sätt att skatta en okänd frontproduktionsfunktion. Inom detta område, dvs. kartläggning av de statistiska egenskaperna hos DEA, skedde en betydande utveckling under 1990-talet med start i Banker (1993), följd av Simar (1996), Kittelsen (1998) och Simar & Wilson (1998, 1999, 2000).

Ända sedan Farrell (1957) är det väl känt att en styckvis linjär front utgör en *pessimistisk* bias av den "sanna" fronten eftersom det som observeras endast är ett fåtal dragningar från den "äkta" men okända teknologin. Detta innebär också att effektivitetsmåten får en *optimistisk* bias. Ju färre observationer vi har tillgång till desto större bias kan förväntas. Banker (1993) visade i fallet med single input – single output att när antalet observationer går mot oändligheten så går avståndet mellan DEA-effektivitetsmåten och de sanna effektivitetsmåten mot noll, dvs. DEA-estimatoren är konsistent. Simar & Wilson (2000) analyserar en generalisering till multipel input – multipel output.

Steget från analys av statistiska egenskaper till beräkning av konfidensintervall är inte långt. Simar och Wilson (1998) introducerade bootstrapping för beräkning av

konfidensintervall för effektivitetsmått. Exakt hur detta ska gå till är kontroversiellt. Ett tidigt bidrag av Simar (1992) fick begränsad uppmärksamhet, men följdes upp i en artikel av Ferrier & Hirschberg (1997), som verkligen tände bootstrappingbrasan med närmast mördande kritik i två artiklar av Simar & Wilson (1999a, 1999b) med mellanliggande svar av Ferrier & Hirschberg (1999). Huvudkritiken gick ut på att den metod som Ferrier & Hirschberg (1997) föreslagit inte var konsistent.

I ett antal artiklar med start i i Simar & Wilson (1998, 1999 och 2000) utvecklades bootstrappingtekniken.

En länk mellan parametriska och icke-parametriska frontansatser är estimering av parametriska sk distansfunktioner; se t. ex. Coelli & Perleman (2000) med referenser. En distansfunktion innehåller all information om teknologin och klarar dessutom av att hantera multiple-output multiple-input problem. Det råder dock fortfarande viss oklarhet exakt hur modellen ska formuleras.

Paneldata i DEA

Medan paneldata är mycket värdefulla i SFA är det inte självklart hur sådana bäst kan utnyttjas i DEA, speciellt inte vid jämförelser med parametriska modeller; se Tulkens & van den Eckaut (1995) för en diskussion av detta. Tulkens & van den Eckaut skiljer mellan tre olika metoder att utnyttja paneldata:

- Temporal
- Intertemporal
- Sekventiell

Temporalmetoden innebär att varje år behandlas separat från övriga år, dvs. effektiviteten år 1 beräknas med data från år 1, effektiviteten år 2 beräknas med data från år 2, etc.

Den intertemporala metoden innebär att data från samtliga år poolas till ett enda dataset

och att effektiviteten beräknas för alla enheter för alla år med detta dataset som referens. Den front som då erhålles kallar Tulkens & van den Eckaut *intertemporal front*. Här utnyttjas information om teknologin från alla tidsperioder.

Den sekventiella metoden innebär att data ackumuleras så att effektiviteten år 1 beräknas med data från år 1, effektiviteten år 2 beräknas med data från år 1 plus data från år 2, etc. Detta torde vara det mest naturliga valet med data för många år och vid jämförelser med resultat från parametriska modeller; se Kumbhakar & Hjalmarsson (1998).

Förändringar i produktivitet och effektivitet

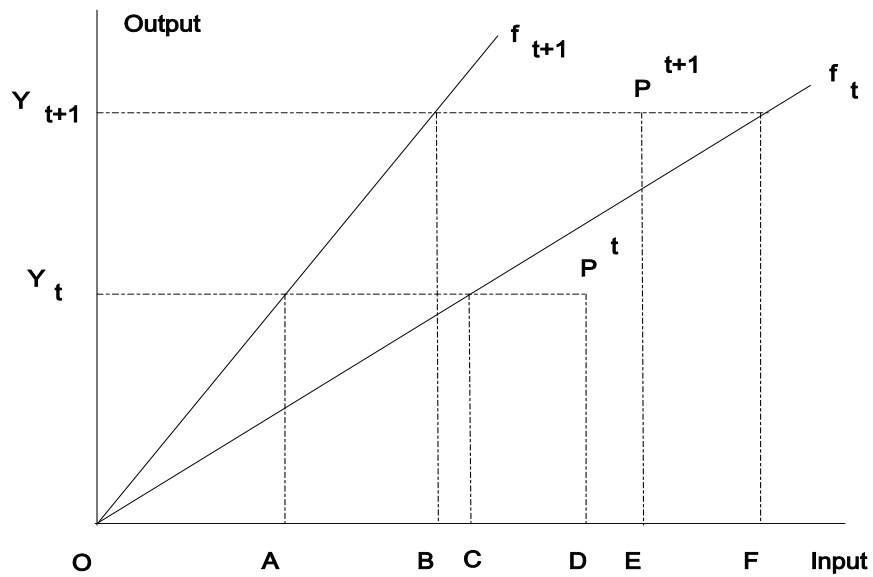
Totalfaktorproduktivitet, TFP, definieras normalt som kvoten mellan produktion och insatsfaktorer. Vid flera produkter och flera produktionsfaktorer måste dessa vägas samman till enskilda siffror, dvs. till index för output respektive input. Vikterna kan antingen vara exogena, dvs. givna utifrån, eller endogena, dvs. bestämda av teknologin. Exogena vikter utgörs normalt av priser som i de välkända produktivitetsindexen av Törnqvist och Fischer. Standardexemplet på endogena vikter är de som erhålls vid appliceringen av en, single output - multiple input neoklassisk produktionsfunktion, då vikterna utgörs av marginalelasticiteterna för produktionsfaktorerna. I vissa fall, som i offentlig tjänsteproduktion, existerar inte priser på alla output och input och då återstår egentligen bara teknologiansatsen. Eftersom multiple output inte utgör något problem i DEA har denna produktionsfunktionsansats kommit att utnyttjas flitigt för analys av produktivitetsanalyser.

Det mest originella bidraget i utnyttjandet av icke-parametriska fronter för produktivitetsanalyser kom med Färe et al (1989), som i sin tur bygger på Caves et al (1982), som baserade sin definition av Malquistindexet på en (neoklassiskt) translogteknologi. Färe et al (1989) visar hur effektivitetsmått från DEA kan utnyttjas för att beräkna Malmquist produktivitetsindex. Detta index är baserat enbart på data för produktion och produktionsfaktorer. Detta index kräver inga förutsättningar om kostnadsminimering eller vinstmaximering vilka ligger underförstådda i Törnqvist och Fisher indexen. Det kräver däremot en skattning av sk distansfunktioner eller

Farrellska effektivitetsmått. Indexet kan närmast karakteriseras som en kvot mellan olika Farrellmått på teknisk effektivitet, där en enskild produktionsenhets effektivitet vid två olika tidpunkter mäts i förhållande till fronten vid båda tidpunkterna. Detta gör det möjligt att dekomponera Malmquistindexet i två delar, så att förändringen i totalproduktivitet kan uppdelas i förskjutning av fronten och förskjutning i förhållande till fronten.

Liksom ifråga om effektivitetsmått har det ursprungliga Malmquistindexet dels en input-orientering dels en output-orientering. I det förra fallet mäts produktivitetsskillnader eller produktivitetsutveckling som skillnader i minimal inputmängd givet en viss outputnivå, i det senare fallet som skillnader i output givet en viss inputkvantitet. Vid konstant skalavkastning sammanfaller de båda måtten. Låt oss börja med att anta konstant skalavkastning illustrerad i Figur 1.

Antag att P är en produktionsenhet som observeras vid två tidpunkter, t och $t+1$. Mellan dessa båda tidpunkter har frontfunktionen förskjutits från f_t till f_{t+1} . Samtidigt har P ändrat läge från P^t till P^{t+1} . P kan ha närmat sig frontfunktionen (catching up) eller halkat tillbaka (lagging behind) som i Figur 1. År t är den tekniska effektiviteten för P mätt mot f_t , $E^{t,t}$, där första index anger frontår och andra index observation, och mätt mot f_{t+1} , $E^{t+1,t}$. År $t+1$ är på motsvarande sätt den tekniska effektiviteten för P relativt f_{t+1} , $E^{t+1,t+1}$, och mätt mot f_t , $E^{t,t+1}$.



Figur 1. Konstruktion av Malmquistindex.

All indexberäkning kräver ett basår eller en basregion. Som vanligt utgörs ytterpunkterna av Laspeyre- och Paasche-indexen med första respektive sista år (region) som bas. Det vanliga torde vara att utnyttja basårets teknologi som utgångspunkt och mäta den tekniska utvecklingen i förhållande till denna. Malmquist produktivitetsindex, M_t , med frontteknologi f_t som bas definieras då som

$$M_t = \frac{E_{t,t+1}}{E_{t,t}}$$

och med frontteknologi f_{t+1} som bas

$$M_{t+1} = \frac{E_{t+1,t+1}}{E_{t+1,t}}$$

Om M är större än 1 har produktivitetsutvecklingen varit positiv. Malmquistindexet kan dekomponeras i två delar:

MC: Förändring i effektivitet i förhållande till fronten (Catching up productivity index)

$$MC = \frac{\frac{OE}{OF}}{\frac{OC}{OD}} = \frac{E^{t+1,t+1}}{E^{tt}}$$

MF_i: Förskjutning i avståndet till fronten (Frontier productivity index) , i = t,t+1

$$MF_i = \frac{E^{t,j}}{E^{t+1,j}}, \quad i, j = t, t+1 \quad i \neq j$$

och

M_i = MC x MF_i Totalproduktivitetsindex, i = t,t+1 ,

För en observation som är tekniskt effektiv båda tidpunkterna är MC = 1, varför indexet då är ett rent frontavståndsmått.

Färe och Grosskopf (1990a) och (1990b) definierar Malmquists produktivitetsindex som det geometriska medelvärde av M_t och M_{t+1}.

$$MG = [M_t M_{t+1}]^{1/2}$$

Catching-up-indexet förblir detsamma medan frontindexet utgörs av det geometriska medelvärde av avståndet mellan fronterna för båda observationerna. Med denna definition av Malmquistindexet krävs kunskap om frontteknologin båda åren medan de ovägsda indexen endast förutsätter kunskap om ett basårs teknologi.

De teoretiska bidragen under 1990- och 2000-talet har framförallt gällt en kartläggning av olika egenskaper hos Malmquistindexet och omformuleringar av indexet så att det uppfyller vissa önskvärda egenskaper. (Som tidigare refererats till visar Simar & Wilson (1999a) bootstrapping för Malmquistindex.)

Bjurek (1996) tar sin utgångspunkt dels i den grundläggande definitionen av totalfaktorproduktivitet som kvoten mellan ett output-index och ett input-index dels det faktum att det "ursprungliga" Malmquistindexet inte alltid existerar vid variabel skalavkastning. Han omformulerar därför indexet till ett totalfaktorproduktivitetsindex, dvs. som en kvot mellan ett outputindex och ett inputindex. I Bjurek et al (1998) jämförs de olika Malmquistindexen.

I ett antal artiklar har det visats hur Malmquistindexet kan dekomponeras i olika komponenter. Till bidragen hör Färe et al (1994), Grifell & Lovell (1999), Balk (2001) och Orea (2002) som på olika sätt visar hur Malmquistindexet kan dekomponeras i teknisk förändring, förändring i effektivitet och förändring i skala. Forsund (1997) bidrar med ytterligare aspekter på optimal skala. Diewert och Nakamura (2003) analyserar mera utförligt egenskaper och dekomponeringar inte bara av Malmquistindexet utan också Laspeyre-, Paasche-, Fisher- och Törnqvistindexen.

6. Reglering, mechanism design och effektivitet

Effektivitetsanalyser har också kommit att utnyttjas vid regleringen av naturliga monopol inom avreglerade marknader, typ elnät, telenät, gas, vatten etc. I samband med den brittiska avregleringen utvecklades en s.k. performance-based regulation (som ett alternativ till den klassiska avkastningsregleringen) i form av en pristaksreglering med nyckelformeln $RPI-X+Y$; se Littlechild (2000) och Jamasb & Politt (2001).

Tolkningen är att prisnivån på en viss tjänst, t ex elenergi, tillåts öka med konsumentprisindex (RPI = retail price index) minus ett produktivitetsökningsmått (X) plus råkraftpriset på el (Y, som är exogent för den enskilde el-distributören). För att fastställa ett mått på rationaliseringspotentialen (X) krävs ett mått på den relativa

effektiviteten. Ett nätbolag som redan befinner sig på fronten för bästa organisering och teknologi bör avkrävas ett lägre rationaliseringskrav än ett nätbolag som befinner sig på långt avstånd från best-practice.

Medan Sverige baserade sin nätreglering (fram till 2009) på en ingenjörsbaserad modell som simulerar ett optimalt elnät, den sk nätnytttemodellen, har t.ex. Norge utnyttjat DEA i sin nätreglering (som i Norge utgör en kombination av pristaks- och avkastningsreglering). Generellt är erfarenheterna av pristaksreglering mycket goda och speciellt erfarenheterna från England visar att, i varje full under de första regleringsperioderna efter en avreglering, finns det en tendens att underskatta rationaliseringspotentialen, med vinstexplosion i de reglerade företagen som resultat. (Tony Blair gick till val på införandet av en windfall gain tax för att dra in en del av dessa rationaliseringsvinster till statskassan.) Sverige har, speciellt i jämförelse med England och USA, en historiskt mycket svag regleringstradition. De "gamla" affärsverken var i huvudsak avkastningsreglerade (avskrivningar till nuanskaffningsvärde plus ränta på statskapitalet) och ännu är det egentligen bara elnätsregleringen som är av pristakskaraktär.

Medan produktionseffektiviteten reflekterar avståndet till en effektivitetsfront reflekterar allokeringseffektiviteten hur väl resurserna utnyttjas inom en sektor, givet att företagen befinner sig på best-practice. Med teknisk effektivitet avses att göra saker på rätt sätt medan allokeringseffektivitet avser att göra rätt saker. Traditionellt har analyser av allokeringseffektivitet fokuserat på prisbildningen på olika marknader och speciellt prisernas avvikelser från marginalkostnaderna på en marknad, speciellt i offentlig tjänsteproduktion.

Analyserna har dock vidgats från mera begränsade analyser av hur prissättningen på en marknad fungerar till hela designen av en marknad. Under rubriken mechanism design (MD) har ett nygammalt forskningsfält vuxit fram under senare år, ett fält som dessutom förärades nobelpriset i ekonomi 2007. Det kan närmast karakteriseras som avancerad institutionell ekonomi, dvs. hur utformar vi effektivt fungerande institutioner för fördelning av resurser inom olika områden. Fältet har sin föregångare i litteraturen

på 1950-1970-talen om effektiva planeringsprocesser i ekonomier med planhushållning av den gamla sovjetiska typen. Fältet har upplevt en renässans i samband med avregleringen av tidigare reglerade marknader men också i samband med införandet av nya, tidigare icke-existerande marknader, t ex olika typer av miljömarknader som handel med utsläppsrätter, el-certifikat, fördelning av telekomlicenser etc. MD-forskningen är mycket tvärvetenskaplig i den betydelsen att här förenas relativt skilda områden av ekonomisk teori, Industrial organisation (IO), auktionsteori, experimentell ekonomi och beteende-ekonomi (Behavioral Economics) och finansiell ekonomi.

Ett av de mest berömda exemplen på framgångsrikt utnyttjande av MD utgör fördelningen av telekomlicenser i USA som inbringade \$800 miljoner i ren vinst för den federala statskassan. Bakom den auktionsprocess som utnyttjades för försäljningen låg en omfattande forskning i ett laboratorium för experimentell ekonomi (vid Caltech). Det hör till ödets ironi att kongressen något år tidigare starkt kritiserat federala anslag till denna typ av forskning, som karakteriserades som fullständigt irrelevant, men vars statsfinansiella avkastning visade sig bli gigantisk.

7. Infrastrukturtjänster och produktivitet

Infrastrukturen i ett land producerar tjänster för övriga delen av näringslivet men också i stor utsträckning för hushållen där den tunga posten utgörs av bostadstjänster. Ofta görs en distinktion mellan kärn- (core) infrastruktur som vägar, järnvägar, hamnar, telekom och icke-kärninfrastruktur som forskning. Forskningen har varit relativt starkt inriktad mot analyser av lönsamheten i infrastrukturinvesteringar genom att se på sambandet mellan infrastrukturkapitalstocken och näringslivets produktivitet. Ett antal försök har gjorts att ekonometriskt kvantifiera kausala effekter av en intervention t.ex. investeringar i transportinfrastruktur. Effekterna har av en rad olika skäl visat sig mycket svårfångade. Det är t.ex. svårt att hitta ”naturliga” experiment som kan renodla effekterna; vägbyggen läggs inte ut slumpmässigt utan utgör delar av större nätverk. Som regel har vi gemensamma trender i produktions- och infrastrukturutveckling varför det är svårt att avgöra åt vilket håll kausaliteten går.

Till de mera berömda studierna hör Aschauer (1989) som på basis av nationella data för USA 1949-1985 fann starka positiva effekter. Ökningar i den offentliga investeringstakten indikerade mycket betydande produktivitetsvinster i ekonomin. Studien ledde till en omfattande diskussion om metod och tolkningsproblem. Något senare publicerade Garcia-Milà och McGuire (1992) en studie baserad på delstatsdata för USA 1969-1983. Två typer av investeringar – motorvägar och utbildning – ingick i studien. Resultaten indikerade att transportinvesteringarna har en positiv effekt (på bruttoregionalprodukten) men med mindre effekt än hos Aschauer. Ytterligare en studie som kan nämnas för USA är Hulten och Schwab (1991). De fann att TFP utvecklades snabbare i snöbältet än i solbältet trots snabbare expansion av infrastrukturen i solbältet. Generellt fann man en svag effekt av infrastrukturinvesteringarna. Intrycket av den amerikanska forskningen är att de tidiga mycket positiva resultaten verkar kollapsa när man använder mera sofistikerade ekonometriska metoder. Studierna finner sällan att samhället är underförsörjt med infrastruktur.

En intressant studie är Röller och Waverman (2001) som ser på sambandet mellan investeringar i telekom-infrastruktur och tillväxt med data från 21 OECD länder 1970-1990. Författarna skattar en modell som endogeniserar telekom-investeringar genom att specificera en mikromodell för utbud och efterfrågan på telekom-investeringar som skattas tillsammans med en makroproduktionsfunktion. Hänsyn tas till simultanitet och landspecifika effekter. Man finner ett positivt samband mellan telekom-investeringar och tillväxt. Effekten är tilltagande när man närmar sig full täckningsgrad.

Även om infrastrukturen uppmärksammades av Produktivitetsdelegationen har jag inte funnit exempel på några mera sofistikerade studier baserade på svenska data med undantag för Berndt och Hansson (1992). Här utnyttjas data från svensk tillverkningsindustri 1960-1988. Under perioden 1974-88 var tillväxten i kapitalstocken 0 % om man exkluderar energi. Författarna finner kostnadssänkande effekter men mycket mindre än hos Aschauer. I analysen undersöks om samhället är underförsörjt med infrastruktur (man jämför optimal kapitalstock med faktisk). Resultaten indikerar att motsatsen var fallet, dvs. samhället var överförsörjt under 1980-talet (13 % över

optimum 1983, 9 % 1988). I denna analys räknas nyttan endast som besparing i den privata sektorns kostnader och ingen hänsyn tas till nyttan för slutkonsumenterna.

En nackdel med de aggregerade ansatserna är att resultaten inte säger något om vilka projekt som bör prioriteras utan resultaten säger endast något om huruvida samhället är underförsörjt av infrastruktur eller inte. Projektbedömningarna är därför i allmänhet av cost-benefit-karaktär.

REFERENCES

- Aschauer, D. (1989). Is Public Expenditure Productive? *Journal of Monetary Economics*, 23, 177-200.
- Aigner, D.J., Chu, S.F., 1968. On estimating the industry production function. *American Economic Review* 58, 226–239.
- Aigner, D.J., Lovell, C.A.K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* 6 (1), 21–37.
- Andersen, P., Petersen, N.C., 1993. A procedure for ranking efficient units in data envelopment analysis. *Management Science* 39 (10), 1261–1264.
- Banker, R. D. (1989). “Econometric Estimation and Data Envelopment Analysis.” *Research in Government and Nonprofit Accounting* 5, 231–243.
- Banker, R.D., 1993. Maximum likelihood, consistency and data envelopment analysis: a statistical foundation. *Management Science* 39 (10), 1265–1273.
- Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies. *Management Science* 39, 1261–1264.
- Banker, R.D., Cooper, W.W., Seiford, L.M., Thrall, R.M., Zhu, J., 2004. Returns to scale in different DEA models. *European Journal of Operational Research* 154 (2), 345–362.
- Battese, G.E., Corra, G.S., 1977. Estimation of a frontier production model: With application to the pastoral zone of Eastern Australia. *Australian Journal of Agricultural Economics* 21, 169–179.

Battese, G.E., Heshmati A., and Hjalmarsson L.(2000), Efficiency of Labour Use in the Swedish Banking Industry: A Stochastic Frontier Approach, *Empirical Economics*, 25(4), 623-640.

Battese, G.E. and T.J. Coelli. (1992). "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India." *Journal of Productivity Analysis* 3, pp. 153-169.

Battese, G.E., Coelli, T.J., 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 20, 325–332.

Bera, A.K. and S.C. Sharma (1999) "Estimating Production Uncertainty in Stochastic Frontier Production Models" . *Journal of Productivity Analysis* 12, pp. 187-210.

Berndt, E. and Hansson (1992). Measuring the Contribution of Public Infrastructure Capital in Sweden, *The Scandinavian Journal of Economics*, Vol. 94, Supplement.

Bird, S. M., Sir D. Cox, V. T. Farewell, H. Goldstein, T. Holt and P. C. Smith, (2005): "Performance indicators: good, bad, and ugly," *Journal of the Royal Statistical Society, Series A*, 168 (Part 1), 1-27.

Bjurek H. (1996) "The Malmquist Total Factor Productivity Index, *Scandinavian Journal of Economics*, 98(2), 303-313.

Bjurek H. and L. Hjalmarsson (1995) " Productivity in Multiple Output Public Service: A Quadratic Frontier Function and Malmquist Index Approach". *Journal of Public Economics*, vol 56, 447-460.

Bjurek H., Försund F., and L. Hjalmarsson (1998) Malmquist Productivity Indexes: An Empirical Comparison, in Färe, R., S. Grosskopf and R.R. Russell (eds) *Essays in Honour of Sten Malmquist*, Kluwer Academic Publ., Boston/London/Dordrecht.

Caudill, S. B., Ford, J.M. and D.M. Gropper (1995) "Frontier Estimation and Firm-Specific Inefficiency Measures in the Presence of Heteroscedasticity, *Journal of Business and Economic Statistics*, 13,105-111.

Caves, D.W., Christensen, L.R., Diewert, E., 1982. The economic theory of index numbers and the measurement of input, output, and productivity. *Econometrica* 50 (6), 1393–1414.

Charnes, A., Cooper, W.W., Lewin, A.Y., Seiford, L.M., 1994. Data Envelopment Analysis. Kluwer, Boston.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operations Research* 2, 429–444.

Coelli, T., Perleman, S., 2000. Technical efficiency of European railways: A distance function approach. *Applied Economics* 32 (15), 1967–1976.

Cooper, W.W., Park, K. S. and J.P. Ciurana (2000) Marginal Rates and Elasticities of Substitution with Additive Models in DEA”, *Journal of Productivity Analysis* 13, 105-123.

Cornwell, C., P. Schmidt, and R.C. Sickles. (1990). "Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels." *Journal of Econometrics* 46, pp. 185-200.

Diewert, E. (1993). “Data Envelopment Analysis: A Practical Alternative?” Paper Presented to Swan Consultants (Canberra) Pty Ltd Conference.

Edvardsen DF, Førsund FR, Kittelsen SAC (2003) Far out or alone in the crowd: classification of self-evaluators in DEA. HERO Working Paper 7/2003. Health Economics Research Programme at the University of Oslo

Edvardsen DF, Førsund FR, Kittelsen SAC (2008) Far out or alone in the crowd: a taxonomy of peers in DEA, *Journal of Productivity Analysis* 29:201–210

Färe, R., Grosskopf, S., Lindgren, B., Roos, P., 1989. Productivity developments in Swedish hospitals: A Malmquist output index approach. Discussion Paper 89-3, Department of Economics, Southern Illinois University, Carbondale (mimeographed), published as Chapter 13 in

Førsund, F.R., and L. Hjalmarsson (2004a). Are all scales optimal in DEA? Theory and empirical evidence, *Journal of Productivity Analysis* 21(1), 25-48

Førsund F.R. and L. Hjalmarsson (2004b) “Calculating Scale Elasticity in DEA Models”, *Journal of Operational Research Society*, 55(10), 1023-1039.

Førsund F.R, Hjalmarsson L., V.E. Krivonozhko and O.B. Utkin (2007) “Calculation of Scale Elasticity in DEA Models: Direct and indirect approaches”, *Journal of Productivity Analysis*, 28(1-2), 45-56.

F.R. Førsund, N. Sarafoglou (2005) *International Journal of Production Economics*, 21

Farrell, M.J., 1957. The measurement of productive efficiency. *Journal of the Royal Statistical Society, Series A* 120 (III), 253–281.

Ferrier, G.D., Hirschberg, J.G., 1997. Bootstrapping confidence intervals for linear programming efficiency scores: with an illustration using Italian banking data. *Journal of Productivity Analysis* 8 (1), 19–33.

Ferrier, G.D., Hirschberg, J.G., 1997. Can We Bootstrap DEA Scores? *Journal of Productivity Analysis* 11, 81-92.

Gagné, R. and P. Oulette (2002) "The Effect of Technological Change and Technical Inefficiencies on the Performance of Functional Forms", *Journal of Productivity Analysis* 17, 233-247.

Garcia-Milà, T. and T. McGuire (1992). The Contribution of Publicly Provided Inputs to States' Economies, *Regional Science and Urban Economics*, 22.

Greene, W.H. (1993) "The Econometric Approach to Efficiency Analysis", in Fried, H.O., Lovell, C.A.K. and Schmidt, S.S. (Eds), *The Measurement of Productive Efficiency: Techniques and Applications*, Oxford University Press.

Greene, W.H. (1993). Frontier production functions. In *Handbook of Applied Econometrics*, vol II - Microeconometrics, eds., Pesaran H.C. and P. Schmidt, Basil & Blackwell.

Heshmati, A. (1994). "Estimating Technical Efficiency, Productivity Growth and Selectivity Bias Using Rotating Panel Data: An Application to Swedish Agriculture". PhD Thesis, Department of Economics, University of Gothenburg.

Hjalmarsson, L. (1991). Metoder i forskning om produktivitet och effektivitet med tillämpningar på offentlig sektor. ESO-rapport, *DsFi* 1991:20.

Hjalmarsson, L. (1996). From Club-Regulation to Market Competition in the Scandinavian Electricity Supply Industry. In Richard Gilbert and Edward Kahn (Eds.) *International Comparisons of Electricity Regulation*, Cambridge University Press.

Hjalmarsson, L., S.C. Kumbhakar and A. Heshmati (1996), "DEA, DFA and SFA: A Comparison". *Journal of Productivity Analysis* 7, 303-327.

- Horrace W.C. and P. Schmidt (1996), "Confidence Statements for Efficiency Estimates from Stochastic Frontier Models", *Journal of Productivity Analysis* 7, 257-282.
- Huang, T-H. and M-H Wang (2004) "Comparisons of Economic Inefficiency Between Output and Input Measures of Technical Inefficiency Using the Fourier Flexible Cost Function", *Journal of Productivity Analysis* 22(1-2), 123-142.
- Hulten C. and R.M. Schwab (1991). Public Capital Formation and the Growth of Regional Manufacturing Industries, *National Tax Journal*, XLIV(4), 121-134.
- Jamasb, T. and M. Politt (2001) Benchmarking and Regulation: International Electricity Experience, *Utilities Policy* 9, 107-130.
- Kittelsen, S.A.C., 1998. Using data envelopment analysis to measure production efficiency in the public sector. Økonomiske Doktoravhandlinger, Universitetet i Oslo.
- Koopmans, T. C. (1991) "An Analysis of Production as an Efficient Combination of Activities." In T.C. Koopmans (ed) *Activity Analysis of Production and Allocation*, John-Wiley and Sons.
- Kumbhakar, S.C. (1990). "Production Frontiers, Panel Data, and Time-Varying Technical Efficiency." *Journal of Econometrics* 46, 201-212.
- Kumbhakar, S.C. and A. Heshmati. (1995). "Efficiency Measurement in Swedish Dairy Farms: An application of rotating panel data", *American Journal of Agricultural Economics* 77, 660-674.
- Kumbhakar, S.C. and L. Hjalmarsson (1993). Estimation of Technical Efficiency and Technical Progress Free from Farm-Specific Effects: An Application to Swedish Dairy Farms. In Fried, Lovell and Schmidt (Eds.) *The Measurement of Productive Efficiency: Techniques and Applications*. Oxford University Press.
- Kumbhakar, S.C. and L. Hjalmarsson (1995). "Labour-use Efficiency in Swedish Social Insurance Offices". *Journal of Applied Econometrics*, Vol 10, 33-47.
- Kumbhakar, S.C., and L. Hjalmarsson (1997): "Relative performance of public and private ownership under yardstick regulation", *European Economic Review* 42, 97-122.
- Kumbhakar, S.C. and C. A. K. Lovell (2000) *Stochastic Frontier Analysis*, Cambridge University Press.

Kumbhakar, S.C., A. Heshmati and L. Hjalmarsson (1997), "Temporal Patterns of Technical Efficiency: Results from Competing Models". *International Journal of Industrial Organization*, 15(5), 597-616.

Kumbhakar, S.C., A. Heshmati and L. Hjalmarsson (1999), "Parametric Approaches to Productivity Measurement: A Comparison Among Alternative Models", *Scandinavian Journal of Economics*, 101(3), 405-424.

Kumbhakar, S.C., A. Heshmati and L. Hjalmarsson, (2002) How Fast Do Banks Adjust? A Dynamic Model of Labour-Use with an Application to Swedish Banks, *Journal of Productivity Analysis*, 18(1), 79-102.

Kumbhakar, S.C., Ghosh, S. and J.T. McGuckin (1991) "A Generalised Production Frontier Approach for Estimating Determinants of Inefficiency in U.S. Dairy Farms" *Journal of Business and Economic Statistics* 9, pp. 279-286.

Lee, Y.H. and P. Schmidt. (1993). "A Production Frontier Model with Flexible Temporal Variation in Technical Efficiency." in *The Measurement of Productive Efficiency Techniques and Applications*, eds. , Fried H., C.A.K. Lovell, and S.S. Schmidt, Chapter 8, pp. 237-255, Oxford Academic Press.

Littlechild, S.C. (2000). A Review of UK Electricity regulation 1999-2000. DAE Working Paper 26, University of Cambridge, Department of Applied Economics.

Lovell, C.A.K. (1993). "Production frontiers and productive efficiency, in the Measurement of productive efficiency techniques and applications". Eds. Fried, H., Lovell, C.A.K. and Schmidt, S.. Chapter 1: 3-67, Oxford University Press.

Maican F. and M. Orth (2008). How do "big-box" entrants influence productivity distribution in retailing. Working Paper, Department of Economics, University of Gothenburg.

Meeusen, W., van den Broeck, J., 1977. Efficiency estimation from Cobb–Douglas production functions with composed errors. *International Economic Review* 18, 435–444.

Pitt, M. and L-F. Lee. (1981). "The Measurement and Sources of Technical Inefficiency In the Indonesian Weaving Industry." *Journal of Development Economics* 9, 43-64.

Schmidt, P. and R.C. Sickles (1984). "Production Frontiers and Panel Data." *Journal of Business and Economic Statistics* 2, pp. 367-374.

Reifschneider, D. and R.E. Stevenson (1991), "Systematic Departures from the Frontier: A Framework for Analysis of Farm Inefficiency." *International Economic Review* 32, 715-723.

Röller, L.H. and L. Waverman (2001). Telecommunications Infrastructure and Economic Development: A Simultaneous Approach, *American Economic Review*, 91(4), 909-923.

Salvanes, K. G. and S. Tjøtta (1994). Productivity differences in multiple output industries: An empirical application to electricity, *Journal of Productivity Analysis*, 5(1), 23-43.

Simar, L., 1996. Aspects of statistical analysis in DEA-type frontier models. *Journal of Productivity Analysis* 7 (2/3), 177–185.

Simar, L., Wilson, P.W., 1998. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science* 44 (11), 49–61.

Simar, L. and P.W. Wilson (1999a): "Estimating and bootstrapping Malmquist indices," *European Journal of Operations Research* 115(3), 459-471.

Simar, L. and P.W. Wilson (1999b): "Some Problems with the Ferrier/Hirschberg Bootstrap Idea," *Journal of Productivity Analysis* 11, 67–80.

Simar, L. and P.W. Wilson (1999c): "Of Course We Can Bootstrap DEA Scores! But Does It Mean Anything? Logic Trumps Wishful Thinking," *Journal of Productivity Analysis* 11, 93–97.

Simar, L., Wilson, P.W., 2000. Statistical inference in nonparametric frontier models: The state of the art. *Journal of Productivity Analysis* 13, 49–78.

Simar, L., Wilson, P.W., 2003. Estimation and inference in twostage semi-parametric models of production processes. Technical Report 0310 IAP Statistics Network.

Staat, M. (2001) "The Effect of Sample Size on the Mean Efficiency in DEA: Comment", *Journal of Productivity Analysis*, 15, 129–137.

Starrett, D. A. (1977): "Measuring returns to scale in the aggregate, and the scale effect of public goods", *Econometrica* 45, 1439-1455.

Torgersen AM, Førsund FR, Kittelsen SAC (1996) Slack-adjusted efficiency measures and ranking of efficient units. *Journal of Productivity Analysis*, 7:379–39

Tulkens, H. and P. van den Eeckaut (1995): "Non-parametric efficiency, progress, and regress measures for panel data: methodological aspects," *European Journal of Operational Research* 80, 474-499.

Wang, H-J. (2002) "Heteroscedasticity and Non-Monotonic Efficiency Effects of a Stochastic Frontier Model" *Journal of Productivity Analysis*, 18, 241–253.

Wilson PW (1995) Detecting influential observations in data envelopment analysis. *Journal of Productivity Analysis*, 6, 27–45

Zhang, Y. , and R. Bartels (1998), " The Effect of Sample Size on the Mean Efficiency in DEA with an Application to Electricity Distribution in Australia, Sweden and New Zealand", *Journal of Productivity Analysis*, 9,187-204.