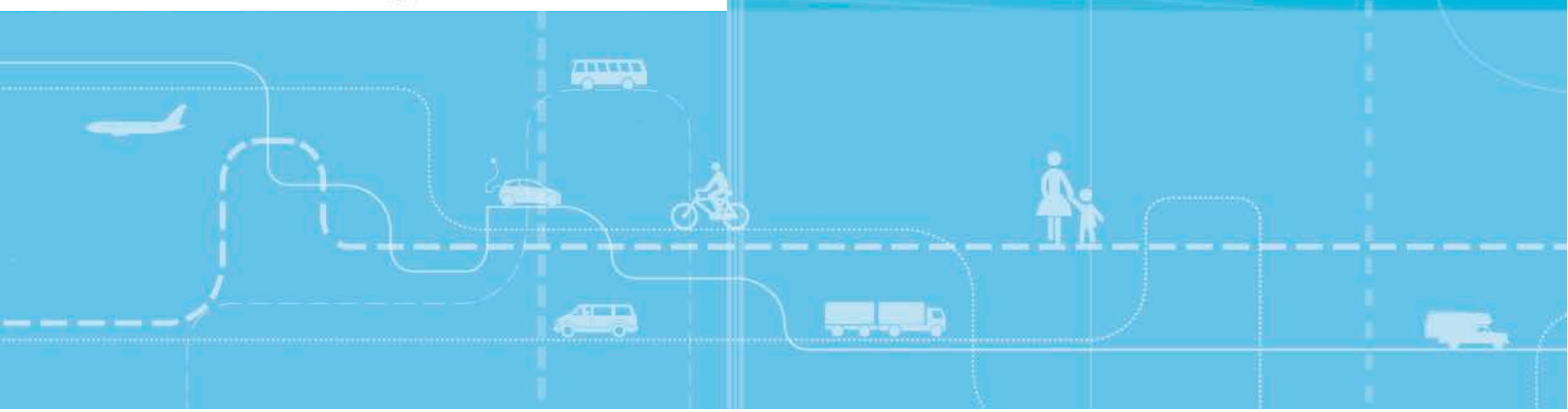
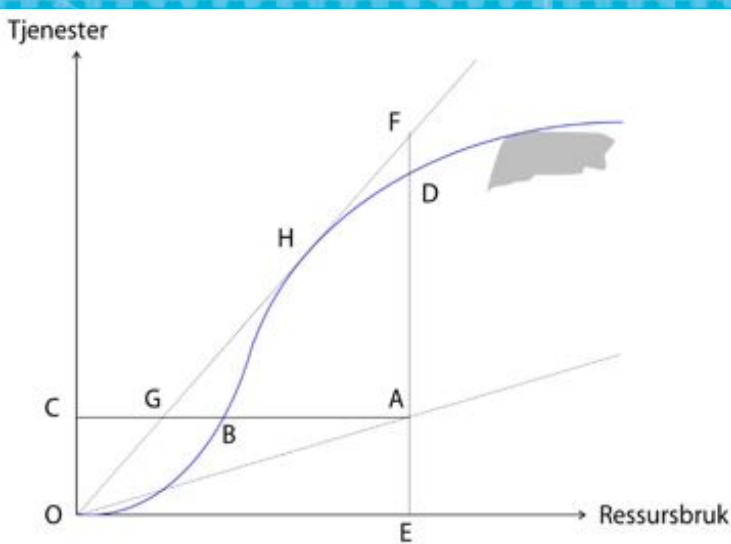


Forbedringspotensial ved måling av effektivitet i kommunal sektor



Forbedringspotensial ved måling av effektivitet i kommunal sektor

Kenneth Løvold Rødseth
Rasmus Bøgh Holmen
Sverre Kittelsen
Finn Før Sund

Transportøkonomisk institutt (TØI) har opphavsrett til hele rapporten og dens enkelte deler. Innholdet kan brukes som underlagsmateriale. Når rapporten siteres eller omtales, skal TØI oppgis som kilde med navn og rapportnummer. Rapporten kan ikke endres. Ved eventuell annen bruk må forhåndssamtykke fra TØI innhentes. For øvrig gjelder [åndsverklovens](#) bestemmelser.

Tittel: Forbedringspotensial ved måling av effektivitet i kommunal sektor

Forfattere: Kenneth Løvold Rødseth
Rasmus Bøgh Holmen
Sverre Kittelsen
Finn Førsumund

Dato: 04.2022

TØI-rapport: 1879/2022

Sider: 89

ISSN elektronisk: 2535-5104

ISBN elektronisk: 978-82-480-1924-4

Finansieringskilde: Kommunal- og distriktsdepartementet (KDD)

Prosjekt: 5077 – Prosjektnavn

Prosjektleder: Kenneth Løvold Rødseth

Kvalitetsansvarlig: Rune Sørensen

Fagfelt: Næringsøkonomi og godstransport

Emneord: Effektivitet; Kommuner; Data Envelopment Analysis; Malmquist indeks

Sammendrag:

Teknisk beregningsutvalg for kommunal og fylkeskommunal økonomi (TBU) ønsket et prosjekt som vurderer mulighetene for og den praktiske anvendbarheten av effektivitetsanalyser av offentlig tjenesteyting. Denne rapporten gir en gjennomgang av beste praksis innen metodikk for effektivitetsanalyse, aktuelle data til måling av kommuneeffektivitet og empiriske sammenlikninger av metoder og modellspesifikasjoner. Rapporten gir anbefalinger om hvordan TBUs etablerte metodikk for måling av kommuneeffektivitet kan forbedres samt aktuelle temaer til videre utredning i TBUs arbeid med å etablere et solid metodisk rammeverk.

Title: Potential for improving efficiency analyses of Norwegian municipalities

Authors: Kenneth Rødseth
Rasmus Bøgh Holmen
Sverre Kittelsen
Finn Førsumund

Date: 04.2022

TØI Report: 1879/2022

Pages: 89

ISSN Electronic: 2535-5104

ISBN Electronic: 978-82-480-1924-4

Financed by: Ministry of Local Government and Regional Development (KDD)

Project: 5077 – Project Name

Project Manager: Kenneth Løvold Rødseth

Quality Manager: Rune Sørensen

Research Area: Industry and Freight

Keywords: Efficiency; Municipalities; Data Envelopment Analysis; Malmquist index

Summary:

Teknisk beregningsutvalg for kommunal og fylkeskommunal økonomi (TBU) requests an assessment of the possibilities for and practical applicability of efficiency analyzes of municipalities. This report provides a review of best practices in efficiency analysis, data for municipal efficiency measurement and empirical analysis. The report provides recommendations on how TBU's established methodology for measuring municipal efficiency can be improved, as well as areas for further research related to TBU's work to establish a solid methodological framework for public sector efficiency measurement.

Language of report: Norwegian

Transportøkonomisk Institutt
Gaustadalléen 21, 0349 Oslo
Telefon 22 57 38 00 - www.toi.no

Institute of Transport Economics
Gaustadalléen 21, N-0349 Oslo, Norway
Telephone +47 22 57 38 00 - www.toi.no

Forord

Rundt 1/5 av de sysselsatte i Norge jobber i kommunesektoren og kommunesektoren yter en betydelig del av offentlige tjenester. Det er derfor viktig at kommunene utnytter ressursene sine best mulig. I lyset av dette har Teknisk beregningsutvalg for kommunal og fylkeskommunal økonomi (TBU) ønsket et prosjekt som vurderer mulighetene for og den praktiske anvendbarheten av effektivitetsanalyser av offentlig tjenesteyting.

TBU gir årlig ut to rapporter med informasjon om utviklingen i kommuneøkonomien de siste årene, med hovedvekt på siste rapporteringsår. I høstrapporten fra Det tekniske beregningsutvalget for kommunal og fylkeskommunal økonomi beskriver TBU den årlige utviklingen i effektivitet i offentlig sektor. Utvalget initierte vinteren 2010 et prosjekt om effektivitet i kommunale tjenester. Formålet var å få mer kunnskap om effektiviteten og effektivitetsutviklingen for de kommunale tjenestene, samt å utvikle et beregningsopplegg for sektorovergrepene effektivitet. Formålet til denne rapporten er å gi en vurdering av det etablerte metodiske opplegget til måling av kommuneeffektivitet, samt å gi innspill til muligheter for videre utvikling.

Prosjektet er bestilt av TBU og er utført av Transportøkonomisk institutt (TØI) og Frischsenteret. Kenneth Løvold Rødseth har ledet prosjektet, som er gjennomført i samarbeid med Rasmus Bøgh Holmen, Sverre A.C. Kittelsen og Finn Førund. Prosjektet har mottatt bistand fra TBU, Senter for Økonomisk Forskning (SØF), Utdanningsdirektoratet (UDir) og Statistisk Sentralbyrå (SSB). Vi takker spesielt Jonathan Smith-Isaksen og Sebastian Brandsrud i Udir, Arne Jensen, Elisabeth Haraldsrud, Else Helena Bredeli, Kjartan Steffensen, Line Gjermshusenge og Peder Næs i SSB og Lars Erik Borge og Ole Henning Nyhus ved SØF. Det ble avholdt to møter med TBU i prosjektperioden. Forfatterne er takknemlig for all støtten som prosjektet har mottatt.

Professor Rune Sørensen ved BI har kvalitetssikret denne rapporten.

Oslo, april 2022

Transportøkonomisk institutt

Bjørne Grimsrud
Aministrerende direktør

Kjell W. Johansen
Avdelingsleder

Innhold

Sammendrag

Summary

1	Innledning	1
1.1	Formålet med rapporten	1
1.2	Framgangsmåte.....	2
1.3	Avgrensning	2
1.4	Rapportstruktur	2
2	Metodikk	3
2.1	DEA-metoden	3
2.2	Måling av effektivitetsendring over tid.....	13
2.3	Effektivitetsanalyse basert på produksjonsindeksen	18
2.4	Benchmarking	26
2.5	Alternative metoder	26
3	Data	29
3.1	Overordnede betraktninger.....	29
3.2	Valg av variabler	36
4	Empiriske analyser	48
4.1	Effektivitetsanalyse.....	48
4.2	Produktivitets- og effektivitetsendring.....	75
4.3	Produksjonsindeksen	80
5	Konklusjon og diskusjon	85
	Referanser	87

Sammendrag

Forbedringspotensial ved måling av effektivitet i kommunal sektor

TOI rapport 1879/2022

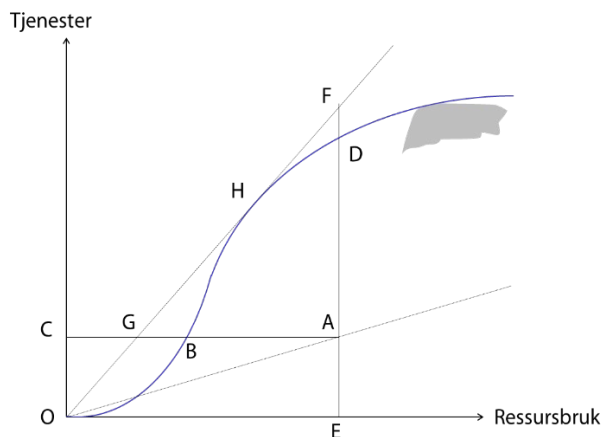
Forfattere: Kenneth Løvold Rødseth, Rasmus Bøgh Holmen, Sverre Kittelsen og Finn Førsund

Oslo 2022 89 sider

Teknisk beregningsutvalg for kommunal og fylkeskommunal økonomi (TBU) ønsket et prosjekt som vurderer mulighetene for og den praktiske anvendbarheten til effektivitetsanalyser av offentlig tjenesteyting. Transportøkonomisk institutt og Frisch-senteret fikk oppdraget. TBU ba om at det skal gjøres en vurdering av dagens metoder for å analysere effektiviteten i kommunene og effektivitetsutviklingen i kommunene som Senter for økonomisk forskning (SØF) gjør årlig for TBU. De ulike metodenes evne til å analysere effektivitetsutviklingen for de største tjenesteområdene samlet over tid og for den enkelte kommune over tid skal vurderes, samt metodenes evne til å sammenlikne effektiviteten mellom kommuner. Det skal gis anbefalinger om hvordan fremtidige effektivitetsanalyser bør gjennomføres enten ved bruk av de metodene som benyttes i dag eller ved å ta i bruk andre metoder. I sammendraget har vi fremhevet konklusjoner og anbefalinger i fet skrift.

Effektivitetsmål

Mens SØFs analyser er fokusert på *effektivitet*, er det vanlig i litteraturen også å være opptatt av *produktivitet*. Slik begrepene brukes i denne rapporten kan forskjellen mellom disse illustreres i en figur:



Figur S.1: Effektivitet og produktivitet.

Figur S.1 ser på tilfellet hvor en tjeneste produseres med en innsatsfaktor eller ressurs. Vi ser på en hypotetisk observert kommune A. Denne kommunen produserer et tjenestevolum lik avstanden OC og bruker OE innsatsfaktor eller ressurs.

Den blå kurven beskriver beste praksis ved gitt teknologi, dvs. effektiv produksjon eller *fronten*. Denne kan estimeres ved ulike metoder; her er den vist som en glatt kurve, mens den i f.eks. data omhyllingsanalyse (Data Envelopment Analysis på Engelsk, forkortet til DEA) vil være en stykkevis lineær front. Kommunens produktivitet (dvs. produkt/innsatsfaktor) er gitt av helningen på den stiplede linjen OA gjennom datapunktet A. Produktiviteten kan angis som selve forholdstallet (f.eks. elever per lærer), men kan også normaliseres

ved å se produktiviteten i forhold til en fast referanse, f.eks. størst mulig produktivitet. I figuren er den størst mulige produktiviteten tegnet som den rette linjen OGHF, siden produktiviteten i punktet H er den høyest mulige med denne teknologien. Når vi her baserer oss på teknologien kalles den normaliserte produktiviteten gjerne *teknisk produktivitet* (TP) og er forholdet mellom helningene på de to linjene. Dette kan forstås som den faktiske tjenesteproduksjonen per enhet ressursbruk i forhold til den høyst mulige tjenesteproduksjonen per enhet ressursbruk. Det kan vises at dette geometrisk kan beregnes som $TP=CG/CA$.

Ressursbesparende *teknisk effektivitet* er derimot et mål på avstanden fra datapunktet til fronten, som her kan beregnes i innsatsfaktorbesparende retning som $TE=CB/CA$. Dette er nødvendig ressursbruk i forhold til faktisk ressursbruk, gitt størrelsen på tjenesteproduksjonen. Forskjellen på de to målene skyldes at kommune A ikke er av optimal størrelse, dvs. at selv om kommunen hadde vært effektiv og redusert sin ressursbruk til avstanden CB ville den fortsatt hatt lavere produktivitet enn om den hadde vært større. Vi definerer derfor ressursbesparende skalaeffektivitet som forholdstallet mellom teknisk produktivitet og teknisk effektivitet $SE=TP/TE=CG/CB$. Alle tre mål vil ha verdier fra 1 og nedover, der f.eks. effektivitet lik 1,0 (eller 100%) tilsier at en kommune er fullt effektiv mens verdier lavere enn 1,0 tilsier at kommunen anslås til å være mer eller mindre ineffektiv.

Dersom det faktisk er konstant skalutbytte i sektoren vil teknologifronten være en rett linje som sammenfaller med linjen maksimal produktivitet. Da vil teknisk effektivitet være lik teknisk produktivitet, og dermed vil også skalaeffektiviteten være 1 over alt.

Metoden som SØF bruker til å estimere fronten er ikke-parametrisk og kalles DEA.

Metoden baseres på å sammenlikne observerte data for produkter og innsatsfaktorer for kommunale tjenestesektorer mot en referansefront som bygger på effektive kommuner.

I avsnitt 2.1 vises de matematiske formuleringene av DEA-metoden som lineære programmeringsproblemer (LP-problemer), og det drøftes spesielt hvordan en kan håndtere ulike variabeltyper i en DEA-analyse. I øvrige metodavsnitt beskrives metoder for å analysere utviklingen over tid, bruk av produksjonsindeksen til å måle effektivitet og produktivitet, samt enkelte alternative metoder. Deretter gir kapittel 3 en gjennomgang av datagrunnlaget og kapittel 4 sammenligner empiriske resultater ved bruk av SØFs metodikk med ulike alternativer. I sammendraget drøfter vi datagrunnlaget og metodenes styrker og svakheter sammen med de empiriske analysene.

De tre sektorer som det beregnes effektivitet for

SØF tar for seg tre sektorer; grunnskole, barnehage og pleie- og omsorgstjenester. Disse sektorene står for brorparten av kostnadene. I skolesektoren benyttes data for de enkelte skoler, og det innføres det en skolebidragsindikator som bygger på karakterer. Kontekstuelle variabler¹ er foreldres utdanning, inntekt og innvandrerbakgrunn.

Det innføres også en miljøindikator som bygger på bl.a. tilfeller av mobbing.

Datagrunnlaget

Rapporten er basert på de data SØF har benyttet i sine analyser og hentes i hovedsak fra KOSTRA og data fra SSB. Ideelt skal data i en effektivitetsanalyse være komplette, dvs. dekke all ressursbruk og all tjenesteproduksjon for en klart avgrenset sektor. Det må heller ikke være overlapp eller dobbelttelling som gjør at samme informasjon opptrer flere ganger. Variable bør dessuten måle de viktige aspektene uten målefeil. Hvilke data som praktisk

¹ Variabler som en kommune ikke kan påvirke.

sett bør inngå i effektivitetsanalyser av kommunal sektor er likevel gjestand for en rekke kompromisser knyttet til tilgjengelighet, datakvalitet, sammenlignbarhet over tid og på tvers av kommuner, samt viktigheten av en eventuell utelatt variabel. Ofte vil også antall variable i en analyse begrenses av antallet observasjoner (frihetsgrader) hvis den skal kunne gi meningsfylt informasjon.

Vi gir i avsnitt 3 en gjennomgang av datagrunnlaget i de ulike sektorene. En rekke variabler er allerede tatt i bruk eller vurdert for SØFs analyser, men vi viser også eksempler på mulige nye variabler i sektorer som allerede analyseres samt andre sektorer hvor det vil være mulig å gjennomføre effektivitetsanalyser. **Vi anbefaler at det gjennomføres årlige effektivitetsanalyser også i flere delsektorer enn skole, pleie og omsorg og barnehager. I sektoravgrensningen er det viktig å ta hensyn til eventuelle samdriftsfordeler og vanskeligheter med å skille data mellom sektorer, f.eks. mellom SFO og skole. Vi anbefaler også at det gjennomføres videre forskning på hvordan en best kan ta hensyn til forskjeller i kostnader mellom by og land og betydningen av interkommunalt samarbeid.**

Innsatsfaktorer

På innsatsfaktorsiden kan det være særlige problemer med måling av kapital på en konsistent måte. Bruk av verdimål heller enn kvantum skjuler problemene med ulik prissetting over tid og på tvers av kommuner; tomtepriser er f.eks. svært ulike i ulike distrikter. Også bruk av lønnskostnader kan skjule hetereogenitet i arbeidskraften; er det høye lønnskostnader fordi en har høyt kvalifiserte og derfor høyt lønnede medarbeidere, eller fordi en bruker mye arbeidskraft?

Hvis en skal aggregere ulike innsatsfaktorer er kostnader det beste målet fordi det veier sammen med priser som skal reflektere den samfunnsøkonomiske verdien av hver enkelt faktortype. Over tid er det likevel riktig å deflatere med gode prisindekser for hver komponent separat. I avsnitt 4.1 har vi sammenliknet resultatene i modeller med ulike kapitalmål og arbeidskraftsmål, og finner rimelig samsvar mellom de ulike modellene.

Vi finner ingen begrunnelse i materialet vi har mottatt for valget å justere av driftskostnader med delkostnadsnøklene i inntektssystemet for kommunene utover at det tar hensyn til bosettingsmønsteret. Det er vanskelig å vurdere kvaliteten i disse korreksjonsfaktorene. Det er rimelig å tro at bosettingsmønsteret vil ha betydning for eller samvariere med kostnadene, men vi drøfter i avsnitt 2.1 om informasjonen om bosettingsmønsteret er utnyttet på den rette måten i effektivitetsanalysene (jf. også neste avsnitt om kontekstuelle variable i tjenesteproduksjonen).

Etter vår vurdering har SØF fremskaffet gode data for de viktigste innsatsfaktorene og i hovedsak behandlet disse på en rimelig måte. Likevel er vi kritiske til å korrigere driftsutgifter med mål for bosettingsmønsteret og anbefaler at en i stedet benytter sentralitetsmål i en samvariasjonsanalyse.

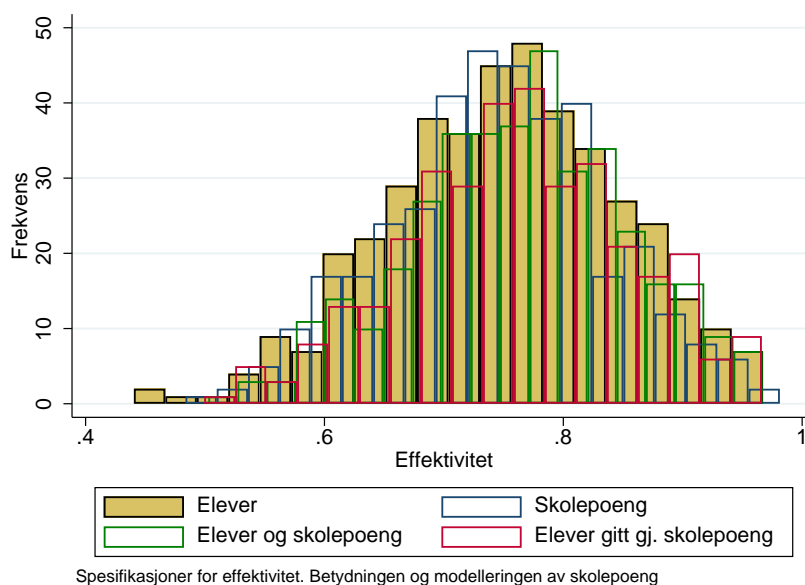
Tjenesteproduksjonen

Måling av tjenesteproduksjonen byr på store problemer. Det finnes gode data på viktige volummål som f.eks. antall elever på hvert trinn i skolen, antall sykehjemspasienter osv. Derimot kan det være store forskjeller i sammensetningen av disse. Det gjelder f.eks. hvor krevende de er å yte tjenester til (vanskelighet) og kvaliteten på tjenestene som ytes. Bruk av mikrodata til å beskrive og kvantifisere slike forskjeller er en lovende utvikling, slik f.eks. skolebidragsindikatoren fanger opp kommunenes bidrag til karakterer for elevene på ulike trinn. Danske myndigheter er her et godt forbilde. **Utvikling av slike indikatorer også på**

andre kommunale sektorer er en forskingsoppgave som TBU og KDD bør oppmuntre til, og ta i bruk i effektivitetsanalysene etterhvert som de foreligger.

Mens det er viktig å kartlegge kvaliteten til tjenestene har bruken av slike indikatorer i selve produksjonsmodellen i DEA-analysen en rekke fallgruver og vi foreslår derfor i stede å ta hensyn til kvalitetsaspektene i analyser av samvariasjonen med volumbaserte effektivitetsmål. Først og fremst er det metodisk ikke rimelig å behandle slike indekser som produkter/outputs i en effektivitetsanalyse; det er f.eks. ikke det samme å øke antall elever og øke det gjennomsnittlige karakternivået. Bruken av delkonstnadsindekser for å korrigere for geografiske og topografiske forskjeller mellom kommunene er av samme type. SØF korrigerer driftskostnadene for slike forskjeller, men i prinsippet gjenspeiler disse forskjeller i vanskeligheten av å yte tjenester heller enn feilkilder i den målte faktorbruken. I tillegg dekker ikke slike kvalitetsindikatorer alle aspekter av sammensetning og kvalitet. Det er f.eks. mange andre aspekter av skolehverdagen enn karakterer og skolemiljø som er viktige for elevenes utvikling. Skal en ta hensyn til mange sammensetnings- og kvalitetsaspekter blir det etter hvert en komplisert og lite gjennomsluktig modell.

I avsnitt 2.1.3 gjennomgås bruken av skolebidragsindikatoren i detalj. I figur S.2 hentet fra den empiriske analysen i avsnitt 4.1 har vi illustrert effektivitetsscorene for grunnskolen estimert ved DEA ved tre ulike håndteringer av kvalitetsaspektet sett opp mot en spesifisering der kun elever per hovedtrinn er benyttet som produkter.



Figur S.2: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalautbytte for 2018 og 2019.

Faktorinnsats: Bruttodriftsutgifter mot bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger

Produkter: a) og d) Elever på hvert hovedtrinn, b) samlede skolepoeng på hvert hovedtrinn og

c) Elever og samlede skolepoeng på hvert hovedtrinn

Kontekstuelle variabler: a) til c) Ingen, d) gjennomsnittlig skolepoeng på hvert hovedtrinn

Figuren gir ikke indikasjon på store avvik i resultatene for ulike måter å behandle skolebidrag/skolepoeng. Generelt vil mindre avvik i kvalitet også gi små avvik i effektivitetsmål i de ulike metodene. Derimot vil en analyse basert på rene volumtall i produksjonsmodellen gi mere gjennomsluktige og metodisk konsistente resultater, samtidig som koeffisientene i den estimerte ettermodellen vil gi informasjon om betydningen av de ulike kontekstuelle variablene for ressursbruk og effektivitet. **Vi anbefaler at variabler som beskriver sammensetning og kvalitet i stedet blir behandlet som kontekstuelle variabler, som**

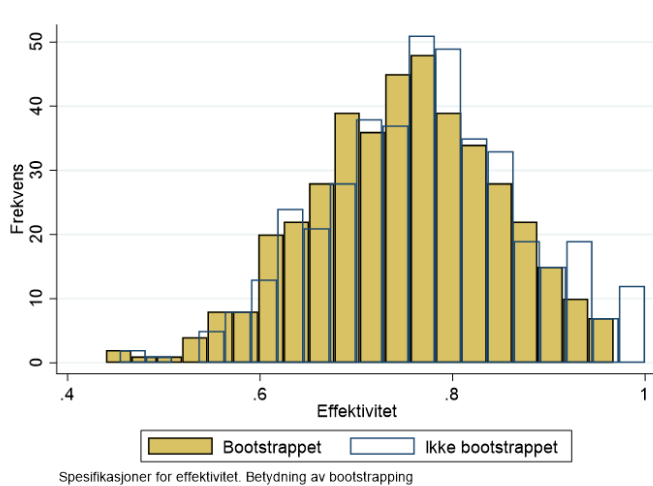
i en DEA-analyse tilsier at det legges i en trinn-to analyse som viser samvariasjonen mellom disse kjennetegnene og effektiviteten basert på en ren volum-modell.

I kapittel 2.1.2 gir vi også en oversikt over metodeutvidelser av DEA-metoden som tillater å inkludere variabler som ikke er på volummål. TBU kan vurdere uttesting av slike modeller som del av det videre arbeidet med effektivitetsanalysene.

Bootstrapping

Effektivitet og produktivitet målt ved DEA-metoden vil være lett å beregne uansett antall variabler og antall observasjoner, men estimatene vil være mer usikre desto færre observasjoner som foreligger. Dette problemet kan avhjelpest ved å bruke *bootstrapping* til å tallfeste usikkerheten i form av konfidensintervall eller standardfeil (Simar og Wilson, 1998). I stolpediagrammet figur S.3 ser vi betydningen av effektivitetsfordelingen med og uten bootstrapping. Bootstrapping vil generelt gi en lavere effektivitet. Vi ser at dette gjelder spesielt for de høyeste effektivitetstall.

Boostrapping hindrer at resultater overfortolkes gjennom å tro at estimatene er mere presise enn de er. Også rangeringen av kommuner kan være usikker og bruk av bootstrap gir mulighet for å vise i hvilken grad en kommune er statistisk signifikant mer effektiv enn en annen. **Vi anbefaler at en fortsatt bruker bootstrapping til å beskrive usikkerheten til effektivitetsestimaterne i DEA-metoden.**



Figur S.3: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet og ikke bootstrappet DEA med variabelt skalautbytte for 2018 og 2019. Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger. Produkter: Elever på hvert hovedtrinn. Kontekstuelle variabler: Ingen.

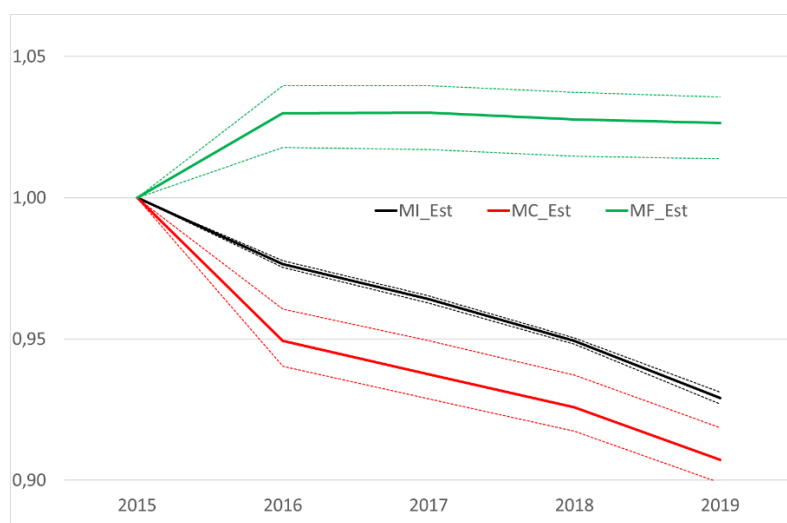
Malmquist-indeksen for produktivitetsendringer

SØF bruker det som kan kalles en vindu-analyse for å beskrive effektivitetsutviklingen over tid. Denne er basert på utvikling av effektivitetstall fra et år til det neste i hvert par av år. Referansteknologien endres med overlapp av årene hver annen periode. I SØFs opplegg vet en ikke om en endring over et lengre tidsrom skyldes endring i variabelspesifikasjonen, i dimensjonaliteten til modellen, i bruk av bootstrap for å korrigere effektivitetsestimaterne, eller i faktisk endring av kommunenes effektivitet mellom tidspunktene. Dermed er ikke disse effektivitetsmål sammenliknbare over tid og metoden er ikke standard i litteraturen.

Innen effektivitetsanalysen er det vanlig å benytte den såkalte Malmquistindeksen (Malmquist, 1953; Caves mfl. 1982) til å måle produktivitet utviklingen og så dekomponere denne til ett mål for *frontendring*, ett for *effektivitetsutviklingen* og ett mål for *endring i skalaeffektivitet*. Fordelene med Malmquistanalyse drøftes nærmere i avsnitt 2.2 mens avsnitt 4.2 viser hvordan dette kan slå ut empirisk for skolesektoren.

Som en ser av figur S.1 må referansen for produktivitet ha konstant skalautbytte (være en rett linje). Over tid vil både fronten og referansen kunne endre seg. Ved beregning av produktivitet *endring* er det derfor også nødvendig å bruke en fast referanse, slik f.eks. prisindekser bruker faste vekter. Den faste referansen kan f.eks. være fronten i første periode, fronten i siste periode, eller en referansefront som utnytter observasjonene i alle år.

En illustrasjon av resultater for Malmquist-indeksen og dets hovedkomponenter vises i figur S.4 hvor vi har beregnet indeksen med en referansefront basert på observasjoner i alle år:



Figur S.4: Estimert Malmquistindeks MI og dets komponenter frontskift MF og opphenting MC; $MI=MC \times MF$. Estimater og 95% konfidensintervall for snittet over kommuner ved bruk av bootstrapping.

«Bootstrapping»² brukes for å beregne konfidensintervall for tallene for indeksen. Vi ser at Malmquist-indeksen MI synker med omtrent 7%, men at konfidensintervallet er ganske smalt. Multiplikativ dekomponering av indeksen i to deler viser at frontindeks MF og opphentingsindeksen MC begge har vesentlig bredere konfidensintervall enn MI indeksen. Dette viser usikkerheten ved estimatene

Malmquistindeksen gir resultater som er lett tolkbare og basert på samme definisjoner over tid. Indeksen og dekomponeringen gir også informasjon om produktivitet utviklingen, i tillegg til effektivitetsutviklingen. Produktivitet er et mål på hvor mye samfunnet får ut av ressursene. Effektivitet er kun et mål på tjenesteproduksjon i forhold til ressursbruk ut fra mulighetene på hvert tidspunkt og ved sin størrelse.

Utfordringen ved Malmquistindeksen er kun at en må ha en enkel modellspesifikasjon hvor det er mulig å skaffe data tilbake i tid og må gjennomføre en større modellkjøring over alle relevante år på nytt. En enkel modell er likevel enklere å forstå. En enkel modell tilbake i

² Bootstrapping brukes av SØF for første gang i TBU rapporten Borge et al. (2020). Meget kort sagt er metoden å analysere sensitiviteten til estimerte effektivitetstall ved å trekke nye sampel for enhetene mange ganger, f.eks. 2000, og så utføre egne effektivitetsberegninger på hvert enkelt datasett. Konfidensintervall på avviket mellom det originale effektivitetstallet til en enhet og det estimerte ved observasjonene som er trukket kan da lages.

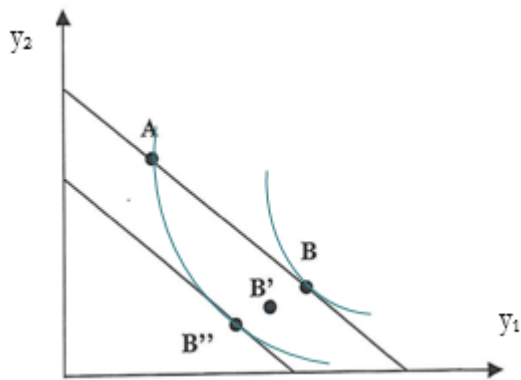
tid forhindrer dessuten ikke at rikere modellspesifikasjoner kan brukes i analysen for det siste året eller de siste årene. **Vi anbefaler at TBU tar i bruk Malmquistindeksen for analyse av effektivitetssutviklingen over tid basert på en enkel modellspesifikasjon.**

Produksjonsindeksen og uavhengighet av prioritering

Produksjonsindeksen (Borge m.fl., 2001) er ment som et mål på kommunenes samlede tjenestetilbud. Den er hovedsakelig utviklet med sikte på å måle produksjon og ikke effektivitet, men TBU har bedt om en vurdering av bruk av indeksen til effektivitetsmåling.

Avsnitt 2.3 beskriver indeksens oppbygging og egenskaper.

Overordnet kan indeksen beskrives som bestående av et sett av kriterier med tilhørende vekt, hvor vektene i hovedsak er budsjettandeler. Det beregnes delindekser for barnehage, grunnskole, primærhelsetjeneste, pleie og omsorg, barnevern, sosialkontortjenester og kultur, samt den overordnede produksjonsindeksen som omfatter alle disse tjenestene.

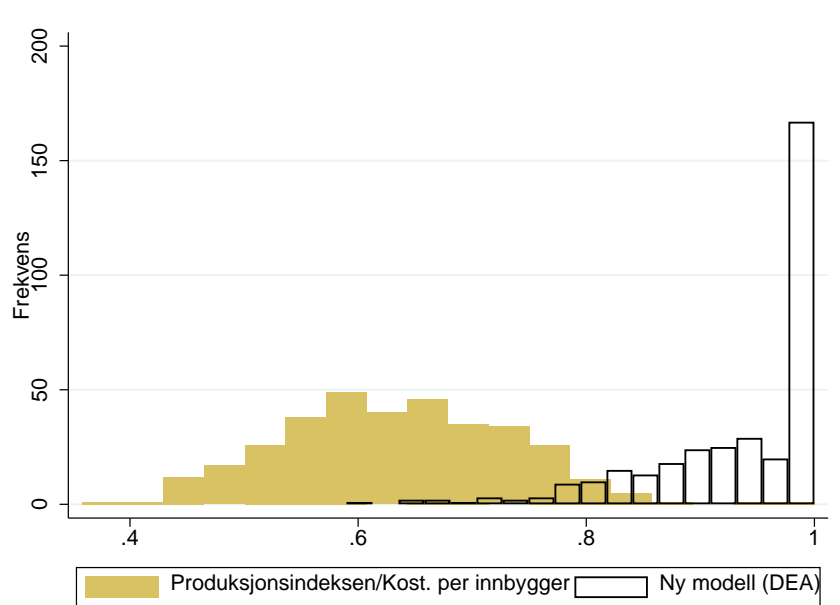


Figur S.5: Nytteoptimering under budsjettbetingelse.

Figur S.5 viser budsjettlinjer for kostnader for kommunene A, B, B' og B''. SØF innfører en forutsetning om at punkter på den samme kostnadslinje har den samme produksjonen. Kommune A har altså den samme produksjonen som kommune B. Men hvis de to kommunene prioriterer tjenestene forskjellig, f.eks. hvis aldersfordelingen av innbyggerne er forskjellig, så gir det ikke god mening at produksjonen er lik. De to blå kurvene («indifferenskurver») representerer nivåer på preferansene kommunene har. Med de samme preferanser med nytte som øker utover i diagrammet, så har A en klart lavere nytte enn B for samme kostnad. En optimering av nyten når preferansene tas i betraktning er ikke forenlig med samme kostnad. Vi innser at utledning av alternative vektorer basert på en sosial nyttefunksjon kan vise seg både vanskelig og svært ressurskrevende. Vi vil i stedet anbefale metodikk til effektivitetsmåling som tillater at vi ikke må benytte priser eller kostnader i beregningen av effektivitet i produksjonsindekssammenheng.

Den enkleste bruken av produksjonsindeksen til å måle produktivitet er å dele produksjonsindeksen på korrigerede brutto driftskostnader pr innbygger. For å normalisere dette som et effektivitetsmål må en velge enheten med høyest produktivitet som benchmark for andre enheter, og beregne effektivitet som forholdet mellom hver kommunes produktivitet og den høyeste produktiviteten målt i utvalget. Dette kalles Corrected ordinary Least Squares (COLS) i effektivitetslitteraturen. I avsnitt 2.3.2 beskrives et alternativt opplegg til måling av effektivitet ved bruk av DEA-metoden på produksjonsindeksens delindekser. Forskjellen mellom SØFs produksjonsindeks og en DEA modell når det gjelder effektivitet vises i figur S.6. DEA modellen har konstant skalautbytte og vektorer for produkter som

bestemmes av modellen og ikke som produksjonsindeksen som bruker gitte budsjett-andeler. Begge modeller har kostnadene per innbygger som eneste innsatsfaktor:



Figur 5.6: En sammenlikning mellom kommuneeffektivitet målt ved DEA og SØFs produksjonsindeks som er estimert med Corrected Least Square-estimatoren (COLS).

Vi ser at COLS-metoden gir effektivitetsscore som i større grad er normalfordelte – noe som kan sees i sammenheng med at produksjonsindeksen er innrettet mot å måle avvik fra det gjennomsnittlige produksjonsnivået.

En DEA-analyse basert på produksjonsindeksen har også klare svakheter ved at den i stor grad er basert på variabler med vidt forskjellig målestokk. Det er ordinale data (f.eks. skolemiljø), korrigerede negative data (f.eks. skolebidrag), og forholdstall. Vi har tidligere drøftet at standard DEA ikke er tilpasset denne typen data, men at det kreves justeringer av DEA-modellen. **En anbefaling, i alle fall i forbindelse med DEA, vil være å i stor grad benytte tradisjonelle volummål i definisjonen av frontmodellen, og heller å korrigere for mer komplekse variabler i en ettermodell. Vi viser til kapittel 2.1.2 for ytterligere detaljer rundt dette.**

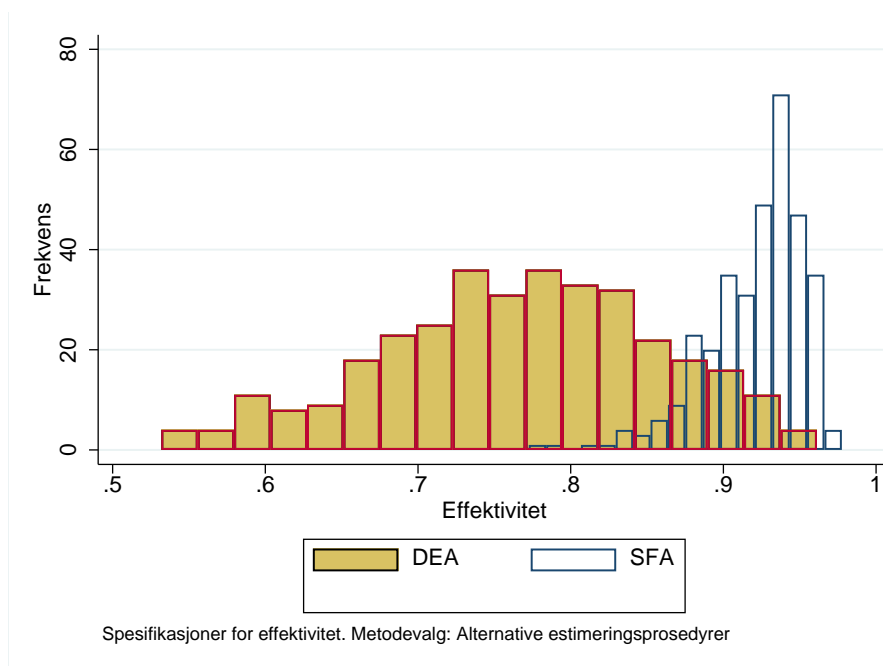
Benchmarking

Benchmarking brukes gjerne som betegnelse på nøkkelindikatorer for tjenester. Slike mål er partielle størrelser. For en kommune kan det være driftsutgifter til barnehager per innbygger, elever per lærer, pleietrengende per pleier og liknende. Men slike partielle mål kan være misvisende når det gjelder å måle samlet effektivitet for en kommune som har flere tjenester og innsatsfaktorer. Metoden tar heller ikke inn over seg avveininger i ressursbruk (f.eks. avveininger mellom ulike produkter i tjenesteproduksjonen), noe SØFs metodikk er i stand til. Benchmarking basert på partielle indikatorer kan heller ikke ta hensyn til utnyttelse av skalaøkonomi og samdriftsfordeler eller ulemper, noe DEA-metoden er velegnet til. Veies de ulike partielle forholdstallene sammen vil det ha samme egenskaper som en effektivitetsindeks basert på produksjonsindeksen som beskrevet i forrige avsnitt. **Mens partielle forholdstall kan gi viktig sektorinformasjon anbefaler vi ikke bruk av disse i en effektivitetsanalyse.**

Sammenlikning av frontmetoder

I tillegg til standard DEA modell er to andre modeller i bruk som er basert på stokastikk, SFA (Stochastic Frontier Analysis) og StoNED modellen (Stochastic Nonparametric Envelopment of Data). StoNED er en videreutvikling av DEA som gjør det mulig å ta simultant hensyn til kontekstuelle variabler. En standard DEA modell som SØF bruker har ikke stokastiske variabler, dvs. variabler med målefeil eller tilfeldige utslag. Dette er en svakhet da stokastisk variasjon i data som oftest vil være relevant for estimering av effektivitet. En modell som er mye brukt er SFA. Men en slik modell krever at man må formulere en parametrisert (matematisk) frontfunksjon. En funksjon som er brukbar kan være vanskelig å konstruere spesielt for tjenesteytende sektorer som kommuner.

Metodene er nærmere forklart i i avsnitt 2.5. En sammenlikning av effektivitet når DEA og SFA brukes til estimering av grunnskole-effektivitet med SØFs variabelspesifikasjon vises i figur S.7:



Figur S.7: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og a) bootstrappet DEA og b) SFA og 2018 og 2019 med variabelt skalautbytte. Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger. Produkter: Elever og samlede skolepoeng på hvert hovedtrinn og skolemiljø. Kontekstuelle variabler: Ingen

SFA gir her en klart sterkere konsentrasjon av kommuner med høyere effektivitet enn DEA. SFA tolker derfor en betydelig del av variasjonen mellom kommunene som tilfeldigheter og målefeil heller enn som forskjeller i effektivitet. Et problem med SFA er imidlertid at det på grunn av de sterke forutsetningene ikke alltid er mulig å finne en løsning. DEA-modellen er velkjent i litteraturen, lett å beregne, og konseptuelt enkel å forholde seg til. DEA-metoden vil også ha mere presis informasjon om effektiviteten og potensialet for forbedring for den enkelte kommunen. **Vi anbefaler fortsatt bruk av DEA (eller dets nære utvidelse StoNED) som hovedmodell i analysene, men ser gjerne at en også bruker SFA til robusthetsanalyser, spesielt fokusert på om rangeringen av kommuner er ulik mellom metodene.**

Paneldataanalyse

Det er fokusert på metodikk for estimering av effektivitet basert på tverrsnittsdata. SFA og StoNED (og varianter av DEA) kan derimot utvides til å ta hensyn til at kommunedata har en paneldatastruktur, dvs. at den samme kommunen observeres gjentatte ganger over tid.

En viktig fordel med paneldata over tverrsnittsdata er at det er mulig å identifisere uobserverbar heterogenitet; dvs. kjennetegn ved kommunene som påvirker deres produksjon men som ikke enkelt kan kvantifiseres ved hjelp av relevante variabler i modellspesifikasjonen. Dette er blitt utnyttet i litteraturen om effektivitetsmåling, hvor det er blitt utviklet modeller som skiller mellom kommuneheterogenitet (som ikke er en del av effektivitetsleddet, men som heller tilskrives variasjon i uforklarte kontekstuelle variabler) og effektivitet. Dette betinger at en er i stand til å skaffe data for de samme variablene for noen år tilbake i tid, noe som igjen antakelig innebærer å velge en enkel variabelspesifikasjon. **Vi anbefaler at paneldataanalyser vurderes nærmere framover.**

Konklusjoner

SØF har opp gjennom årene levert et solid arbeid i sine analyser av effektivitet i kommunesektoren, i all hovedsak basert på det som har vært standard innenfor effektivitetslitteraturen. Likevel har det i analysene vært foretatt tilpasninger som fremstår som noe mer ad hoc og derfor mindre gjennomskubare og etterprøvbare. SØFs analyser avviker fra standard metodikk innen effektivitetsanalyser på en del punkter.

Dette gjelder bl.a. behandling av kontekstuelle variabler og endring over tid. Vi anbefaler mer bruk av standardmetoder. DEA er primært egnet for volummål, men SØF blander data med ulike målestokker. Dette fører til at effektivitetsberegninger ikke blir korrekte og vanskelige å tolke. Behandling av kvalitet og sammensetning er et vanskelig punkt. Mange aktuelle variabler har spesielle måleenheter som ikke egner seg for DEA som er basert på volumdata og ikke data basert på forholdstall eller diskrete kategorier. Også korreksjoner av driftskostnadene kan spille inn på effektivitetsmålingen. Vi anbefaler bruk av disse som kontekstuelle variabler i en tilleggs- eller ettermodell som kortsiktig løsning, men anbefaler videre forskning på temaet.

Vi finner empiriske forskjeller mellom teknisk effektivitet og produktivitet. Det viser at SØF kan ha nytte av å utvide modellopplegget til å dekomponere teknisk effektivitet og skalaeffektivitet. Det sistnevnte kan gi nyttig informasjon om hvordan kommunestørrelse påvirker produktiviteten.

Spesielt for analyser av effektivitet over tid ser vi ikke SØFs vinduanalyse som hensiktsmessig da effektivitetstallene ikke er sammenliknbare når modellen skifter over tid. Våre beregninger viser at dette har til dels betydelige innvirkninger på resultatene. Kjeding av effektivitetsscore viser da en sammensatt effekt av modell- og effektivitetsendringer. Malmquist-indeksen for produktivitetsendring over diskret tid er standarden i litteraturen og vi anbefaler at denne tas i bruk, da med en enkel variabelspesifikasjon som lar seg bruke for noen år tilbake. Malmquist-indeksen gir også informasjon om produktivitetsutviklingen i tillegg til effektivitetsutviklingen.

SØFs opplegg for å måle effektivitet basert på produksjonsindeksen gir ikke et korrekt mål på effektivitet fordi opplegget gir en indeks for produktivitet, ikke effektivitet, slik begrepene vanligvis brukes. Man trenger en referanseproduktivitet for å beregne effektivitet. Produksjonsindeksen har også flere uønskede egenskaper, som bl.a. konstant bytteforhold mellom produkter. Vi foreslår en mulig alternativ produksjonsindeks som tilfredsstillende vanlige krav til effektivitetsmåling, men vil ikke anbefale dette som hovedtilnærming nå. Det virker hensiktsmessig å ta en grundig gjennomgang av variablene i

produksjonsindeksen og i hvilken grad dagens opplegg er vellykket i å konvertere dem til en felles målestokk. En anbefaling i forbindelse med DEA-metoden vil være å i stor grad benytte tradisjonelle volummål i definisjonen av modellen for produktfunksjonen, og heller korrigere for mer komplekse variabler i en ettermodell.

DEA-metoden er enkel å forstå og solid etablert som en hovedmetode i effektivitetsmåling internasjonalt. Vi anbefaler at denne fortsatt benyttes som hovedmetode, og at usikkerheten fortsatt beskrives ved bruk av bootstrapping. Benchmarking i betydningen partielle forholdstall er ikke spesielt egnet til effektivitetsanalyse. En utvidelse av metodeopplegget er å anvende stokastiske metoder (SFA; StoNED) og/eller paneldatametodikk til å beriket analysene og belyse robustheten til resultatene. Våre beregninger viser at resultatene fra stokastiske modeller avviker sterkt fra den deterministiske DEA-modellen. Paneldata kan gi ny innsikt, bl.a. ved å gjøre det mulig å skille mellom effektivitet og andre forskjeller mellom kommunene som har betydning for ressursbruken.

Løpende statistikk og analyser av effektivitets- og produktivitetsutviklingen i kommunesektoren har stor samfunnsmessig verdi. Vi anbefaler at arbeidet fortsetter, men med større vekt på standard metoder.

Vi har gjennomgått en del kilder til data for ressursbruk og tjenesteproduksjon i kommunesektoren, og dette bør kunne gi grunnlag for rikere modeller i flere delsektorer. Likevel, enkle modeller er oftest mest gjennomslagsfulle og forståelige.

Etterprøvnbarhet og reproducerbarhet er viktig for resultatenes legitimitet, i tillegg til å gjøre det enklere med evalueringer som denne. Det bør derfor etableres ordninger for deponering eller offentliggjøring (f.eks. på web) av datagrunnlaget både i rå form og i bearbejdede form slik at det kan for fremtiden er enklere å gjøre reproduksjonsstudier eller alternative antakelser og metoder.

I tillegg er det ønskelig å stimulere til videre forskning, f.eks. på hvordan kvalitetsaspekter kan bli tatt hensyn til, særlig i ikke-parametriske eller semi-parametriske metoder. Det er dessuten ønskelig å stimulere forskning om hvordan en skal behandle interkommunalt samarbeid, betydningen av bosettingsmønster og ikke minst utvikling av flere kvalitetsmål basert på mikrodata.

1 Innledning

Rundt 1/5 av de sysselsatte i Norge jobber i kommunesektoren og kommunesektoren yter en betydelig del av offentlige tjenester som er svært viktig for innbyggerne. Det er derfor viktig at kommunene utnytter ressursene sine best mulig. Begrepet *produktivitet* beskriver forholdet mellom samlet produksjon og ressursbruk, mens *effektivitet* beskriver forholdet mellom en organisasjons faktiske produktivitet og norm for beste praksis produktivitet. Produktivitets- og effektivitetsanalyse er følgelig et sentralt virkemiddel til å analysere produktivitetsutviklingen og kartlegge potensialet for effektivitetsforbedringer i kommunene.

Det tekniske beregningsutvalget for kommunal og fylkeskommunal økonomi (TBU) innledet i 2010 et prosjekt for å beskrive den årlige utviklingen i effektivitet i offentlig sektor. Senter for økonomisk forskning (SØF) har utviklet metodikken som benyttes til dette formålet, og som er tuftet på Data Envelopment Analysis (DEA). Høstrapporten fra TBU beskriver den årlige utviklingen i effektivitet i offentlig sektor.

Kommunal- og Moderniseringsdepartementet peker også på andre aktuelle metoder for effektivitetsanalyse av kommunene. En enklere tilnærming er en systematisk sammenlikning av kommuner basert på utvalgte nøkkeltall fra KOSTRA eller andre datakilder. Dette omtales her som *benchmarking*. En tredje tilnærming er å sammenlikne kommunenes kostnader/ressursbruk mot *produksjonsindeksen*. Denne er tilgjengelig på kommunedata.no og i TBUs høstrapport, og er ment å si noe om omfang og kvalitet på tjenestetilbudet til kommunene.

Disse tre tilnærmingene omtales som «dagens metodikk» i det følgende.

1.1 Formålet med rapporten

Kommunal- og moderniseringsdepartementet (heretter Oppdragsgiver) ber om en ekstern vurdering av dagens metodikk og dens evne til å analysere effektivitetsutviklingen i) per tjeneste, ii) per kommune, iii) over tid og iv) mellom kommuner. Prosjektet skal innrettes mot:

- Å gjøre en vurdering av dagens metoder
- Å vurdere dagens metoders evne til å analysere effektivitetsutviklingen
 - for de største tjenestene levert av kommunene
 - for den enkelte kommune
 - mellom kommuner.
- Forhold som strukturelle/kontekstuelle forhold, aktivitetssammensetning og organisering i kommunene skal tas hensyn til
- Å gi anbefalinger til framtidig metodebruk, både når det gjelder metodevalg, datagrunnlag og presentasjon/visualisering av analysene

1.2 Framgangsmåte

Prosjektet er gjennomført med støtte av en referansegruppe med utspring i TBU. For å sikre at vi får med den nødvendige data- og områdekompetanse har vi også vært i dialog med eksperter på relevante statistikkområder.

Data er samlet inn fra SØF, men er også supplert med innhentede data fra offentlig statistikk. Beregninger i rapporten er gjennomført i all hovedsak i Stata SE v.16. Beregningene av Malmquistindeksen er gjort i FrischNonParam v.3.9.

1.3 Avgrensning

Oppdragsgiver påpeker at de tre sektorene som inngår i dagens metodikk er de aller viktigste for prosjektet. Utover det er det også av interesse med vurderinger knyttet til muligheter for analyse også av andre sektorer, og eventuelle sammenhenger mellom sektorer som kan være relevant. Av prioriteringshensyn vil vi derfor i begrenset grad vurdere mulighetene for å utvide analysene til flere sektorer enn de som inngår i gjeldende rapport.

1.4 Rapportstruktur

Kapittel 2 beskriver og kommenterer metodikk. Kapittel 3 drøfter alternative og supplerende metoder og data til effektivitetsanalyse av kommunene. Kapittel 4 gjengir resultater fra empirisk analyse basert på utvalgte metoder og modellspesifikasjoner. Kapittel 5 oppsummerer funnene i studien og gir anbefalinger om data- og metodebruk for effektivitetsanalyse av kommunene.

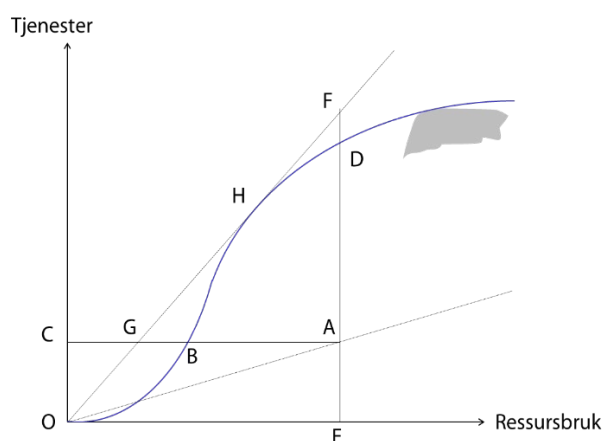
2 Metodikk

Dette kapitlet gjennomgår dagens metodikk. Vi deler inn kapitlet i diskusjon om metoder og data/variabelvalg. Innen hvert underkapittel gir vi innledningsvis en overordnet beskrivelse av metodene og datasettene, før vi kommenterer og evaluerer dem.

2.1 DEA-metoden

Vår oversikt over SØFs metodikk baserer seg på Kommunal- og Moderniseringsdepartementet (2020) og Borge mfl. (2020). Vi fokuserer her på metoden, mens SØFs bruk av data blir omtalt i kapittel 3.

Mens SØFs analyser er fokusert på *effektivitet*, er det vanlig i litteraturen også å være opptatt av *produktivitet*. Forskjellen mellom disse to begrepene kan illustreres som følger:



Figur 2.1: Effektivitet og produktivitet.

Figur 2.1 ser på tilfellet hvor en tjeneste produseres med en innsatsfaktor eller ressurs. Den blå kurven beskriver beste praksis ved gitt teknologi, dvs. effektiv produksjon eller *fronten*. Denne kan estimeres ved ulike metoder; her er den vist som en glatt kurve, mens den i f.eks. DEA-metoden vil være en stykkevis lineær front. Illustrasjonen viser også datapunktet til en kommune (angitt A). Kommunens produktivitet (dvs. produkt/innsatsfaktor) er gitt av den stiplede linjen OA gjennom datapunktet A. Produktiviteten kan angis som selve forholdstallet (f.eks. elever per lærer), men kan også normaliseres ved å se produktiviteten i forhold til en fast referanse. I figuren er den størst mulige produktiviteten tegnet som den rette linjen OGHF, siden produktiviteten i punktet H er den høyest mulige med denne teknologien. Når vi her baserer oss på teknologien kalles den normaliserte produktiviteten gjerne *teknisk produktivitet (TP)* og er forholdet mellom helningene på de to linjene. Det kan vises at dette geometrisk kan beregnes som $TP=CG/CA$.

Ressursbesparende *teknisk effektivitet* er derimot et mål på avstanden fra datapunktet til fronten, som her kan beregnes i innsatsfaktorbesparende retning som $TE=CB/CA$. Forskjellen på de to målene skyldes at kommune A ikke er av optimal størrelse, dvs. at selv

om kommunen hadde vært effektiv og redusert sin ressursbruk til avstanden CB ville den fortsatt hatt lavere produktivitet enn om den hadde vært større. Vi definerer derfor ressursbesparende skalaeffektivitet som forholdstallet $SE=TP/TE=CG/CB$. Dersom det faktisk er konstant skalausbytte i sektoren vil teknologifronten være en rett linje som sammenfaller med linjen maksimal produktivitet. Da vil teknisk effektivitet være lik teknisk produktivitet, og dermed vil også skalaeffektiviteten være 1 over alt.

DEA er en velkjent metodikk for effektivitetsanalyse. Farrell (1957) innførte en ikke-parametrisk beste praksis front, mens Farrell & Fieldhouse (1962) viste hvordan en kunne formulere metoden som et lineært programmeringsproblem og Charnes mfl. (1978) lanserte begrepet DEA. Metoden bygger på tre sentrale antakelser, nemlig *fravær av målefeil* («feasability», dvs. at det som er observert også er mulig), *monotonisitet* (ofte kalt «free disposal», dvs. at man alltid kan bruke mer av ressursene uten å endre produksjonen, evt. produsere mindre uten å endre innsatsfaktorbruken) og *konveksitet* (dvs. at det er mulig å lineært kombinere eksisterende produksjonsplaner). DEA vil være den metoden som gir det minste mulighetsrommet («minimum extrapolation») gitt disse tre antakelsene (Banker 1996).

Et av metodens viktigste fortrinn er at den er *ikke-parametrisk*, noe som betyr at forskeren slipper å velge formen på funksjonen som skal tilpasses (og at man dermed unngår valg av funksjoner som ikke passer dataene). En ulempe med metodikken er at den er *deterministisk*. Det vil si at den ikke skiller mellom effektivitet og stokastisk variasjon i dataene, men at alle avvik fra den estimerte fronten tolkes av modellen som ineffektivitet. Metoden blir derfor sårbar for målefeil i data og spesielt innflytelsesrike datapunkter, såkalte «uteliggere». Metoden er også sårbar for «dimensjonalitetens forbannelse». Dette betyr at om forholdet mellom antall observasjoner i datasettet og antall variabler i modellen blir lavt vil modellen i liten grad være i stand til å skille mellom effektive og ineffektive enheter. Det er også kjent at metoden underestimerer effektivitet i små utvalg.

La $k=1,\dots,K$ være en indeks for kommunene, $n=1,\dots,N$ være en indeks for innsatsfaktorene og $m=1,\dots,M$ være en indeks for produktene. Da kan innsatsfaktorbesparende DEA under variabel skalaavkastning for kommunen k' generelt defineres:

$$TE(\mathbf{x}^{k'}, \mathbf{y}^{k'}) = \min_{\theta, \lambda} \left\{ \theta : \begin{aligned} & \sum_{k=1}^K \lambda^k \mathbf{y}_m^k \geq \mathbf{y}_m^{k'}, \forall m \\ & \sum_{k=1}^K \lambda^k \mathbf{x}_n^k \leq \theta \mathbf{x}_n^{k'}, \forall n \\ & \sum_{k=1}^K \lambda^k = 1 \\ & \boldsymbol{\lambda} \geq \mathbf{0} \end{aligned} \right. , \quad (1)$$

hvor TE er et effektivitetsmål (*innsatsfaktorbesparende teknisk effektivitet*) som tar verdien 1 dersom kommune k' befinner seg på produksjonsfronten, men er mindre enn 1 dersom kommunen har et effektiviseringspotensial. Modellen i (1) kan forstås på følgende måte: Venstre side av ulikhetene for innsatsfaktorer og produkter bestemmer fronten som kommune k' skal sammenliknes mot. Fronten lages ved å kombinere dataen fra de ulike kommunene. Dette gjør modellen ved å velge verdier for intensitetsvariablene $\boldsymbol{\lambda}$, som i optimeringen i likning (1), vil ta en positiv verdi for enheter som er referent for kommune k' og null for andre enheter. Begrensningen at summen av intensitetsvariablene skal være lik 1 pålegger variabel skalaavkastning. Den sikrer at fronten gir en «tett» omhylling av eksisterende datapunkter, og at det ikke er mulig å f.eks. doble et datapunkt (noe som i så tilfelle måtte innebære at intensitetsvariablene summeres til 2).

Dersom denne begrensningen fjernes fra programmet estimeres modellen under *konstant skalaavkastning*. Ofte brukes betegnelsen CRS om en slik DEA-modell, som gir et mål på *teknisk produktivitet*. Forholdet mellom effektivitetene målt med og uten begrensningen vil følgelig gi et mål på *skalaeffektivitet*.

Effektivitet og produktivitet målt ved DEA-metoden vil være lett å beregne uansett antall variabler og antall observasjoner, men estimatene vil være mer usikre desto færre observasjoner som foreligger. Dette problemet kan avhjelpest ved å bruke *bootstrapping* til å tallfeste usikkerheten i form av konfidensintervall eller standardfeil (Simar og Wilson, 1998). Denne metoden innebærer at det lages alternative datasett f.eks. ved å gjøre flere trekninger av delutvalg av alle kommuner, og at det gjennomføres egne effektivitetsberegninger på hvert enkelt datasett. Antall replikasjoner er gjerne 2000.

Bootstrapping hindrer at resultater overfortolkes gjennom å tro at estimatene er mere presise enn de er. Også rangeringen av kommuner kan være usikker og bruk av bootstrap gir mulighet for å vise i hvilken grad én kommune er statistisk signifikant mer effektiv enn en annen.

Borge mfl. (2020) anvender DEA til å beregne effektivitet sektorvis per kommune per år mellom 2017-2018 basert på en front konstruert ved å slå sammen alle datapunktene for 2017-2018. Etter estimeringen aggregeres de sektorvise effektivitetsscorene opp til et effektivitetstall per kommune ved å anvende sektorenes budsjettandeler som vekt. Borge mfl. (2020) anvender bootstrapping til beregning av effektivitetsscore, mens tidligere års effektivitetsstudier av kommunene ikke benytter denne tilnærmingen.

DEA-metoden er standard metodikk innen produktivitets- og effektivitetsanalyse. Som vi har vært inne på har denne både fordeler og ulemper vurdert opp mot konkurrerende metoder som f.eks. Stochastic Frontier Analysis (SFA). Vi vil derfor også se på andre modeller som blant annet tillater å ta hensyn til stokastisk variasjon i neste kapittel. I dette kapitlet fokuserer vi på mulige forbedringspunkter knyttet til SØFs anvendelse av DEA til måling av kommuneeffektivitet.

2.1.1 Håndtering av uteliggere

Fjerning av uteliggere er spesielt viktig i deterministiske metoder som DEA, som ikke har noe god måte å håndtere stokastisk støy på, men isteden tolker den som en del av effektivitetsforskjellene. Kommunene som ikke er med i SØF analyser inkluderer kommuner med manglende data eller som SØF skjønnsmessig har definert som uteliggere. Det kan være fornuftig å fjerne de mest markante uteliggerne ved visuell inspeksjon, men det beste er om man klarer å begrunne denne fjerningen av uteliggere med noen objektive kriterier.

Et mulig grep er å fjerne enheter som ligger mer enn for eksempel 25 prosent foran en front definert av alle andre observasjoner (fjerning av enheter basert på såkalt supereffektive enheter, jamfør Andersen and Petersen 1993). I et annet mulig grep fjernes enheter som definerer referansefronten for enheter med til sammen mer enn 25 prosent av utfallsvariabelen (fjerning av uteliggere basert på såkalt Torgersens Rho, jamfør Torgersen, Førsumd and Kittelsen 1996). Slike grenser har riktignok heller ikke en solid teoretisk basis, men er likevel basert på erfaringer med spesielle observasjoner i mange studier (se for eksempel Banker 1995).

2.1.2 Kvalitetsjustering av variabler

I SØFs analyser er det i stor grad brukt variabler som er korrigert for ulikheter i rammebetingelser eller kvalitet. Det er viktig å måle og rapportere slike forskjeller og det framstår som en forskningsoppgave å utvikle kvalitetsindikatorer for flere sektorer. Det er likevel

ikke åpenbart *hvordan* kvalitetsindikatorer skal brukes i en analyse av produktivitet og effektivitet.

Særlig gjelder dette ved bruk av DEA. I DEA-metoden er det i utgangspunktet kun to typer variabler: innsatsfaktorer (inputs) og produkter/tjenesteproduksjon (outputs). Variablene må være kardinalmål og ha dimensjon som volum. Det må f.eks. være meningsfylt å snakke om en dobling av en variabel. Dersom to kommuner av omtrent samme størrelse slår seg sammen vil den sammenslåtte naturlig kunne ha omtrent det dobbelte av alle ressursbruksvariablene og tjenesteproduksjonsvariablene. SØFs metodikk tar i liten grad hensyn til denne betingelsen, og benytter gjerne variable med ulike måleenheter i samme analyse.

Negative tall, forholdstall (brøker) og tall på en skala som er naturlig begrenset ovenfra vil derfor ikke kunne brukes direkte i en DEA-analyse. Mens det er lov å endre en variabel multiplikatvnt uten at det får betydning for resultatene (f.eks. gange med en konstant for å uttrykke kostnader i euro heller enn kroner), vil metoden ikke være robust overfor additive endringer (f.eks. legge til en konstant). Det finnes tilpasninger av metodene for å kunne bruke variabler av andre typer, og av og til er det mulig å omforme variablene til å ha en naturlig volumtolkning (Olesen et al., 2017). F.eks. vil et forholdstall som enerom per pasient lett kunne omformes til antall enerom hvis en kjenner antall pasienter.

Faktaboks: Måleenheter i DEA

- **Forholdstall:** Bruken av disse i DEA har vært omdiskutert. Olesen mfl. (2015; 2017) viser at problemet knytter seg til at forholdstall ikke tilfredsstiller standardantakelsene DEA-modellen bygger på. Olesen mfl. (2015; 2017) utvikler nye DEA-modeller som er tilpasset forholdstall.
- **Kvalitative variabler:** Det kan være flere utfordringer med behandling av kvalitative variabler i DEA. For det første er en kvalitativ variabel gjerne definert i et gitt intervall, og skalering ut over intervallet (noe DEA typisk vil tillate) er ikke mulig. For det andre kan det være ulike verdier knyttet til ulike nivå av en kvalitativ variabel: DEA vil for eksempel tolke verdien 40 som en dobling av verdien 20, selv om den kvalitative tolkningen verdsetter utfallet 40 som mange ganger mer betydningsfullt enn utfallet 20. For det tredje angis kvalitative variabler ofte som diskret utfall (eks. ved bruk av en Likert-skala) mens DEA er tilpasset kontinuerlige data. Noen aktuelle metoder for å håndtere kvalitative data i DEA er såkalt Imprecise DEA (IDEA) og Fuzzy DEA (FDEA). Disse er beskrevet bl.a. i Shen mfl. (2011).
- **Negative tall:** DEA-metoden er ikke egnet for variabler som tar både positive og negative utfall siden dette påvirker/endrer ulikhetene i DEA-modellen (jf. likning 1). Det kan være fristende å endre en variabel med negative tall ved å legge til en konstant, for å sikre at variabelen kun tar positive verdier. Dette er normalt ikke en gyldig operasjon siden DEA-modellen ikke er *translasjonsinvariant* (dvs. at effektivitetsscorene og deres rangering kan påvirkes ved denne datamanipulasjonen). En mulighet for å løse problemet med negative utfall (uten å endre variabelen som tar negative tall) er å basere effektivitetsmålingen på den såkalte retningsbestemte (directional) distansefunksjonen (Kersten og van de Woestyne, 2011).
- **Målefeil:** DEA er en deterministisk metode, noe som innebærer at målefeil og støy i variablene ikke tas eksplisitt hensyn til. Dersom dette er en viktig utfordring kan det være relevant å benytte stokastiske metoder (jf. kapittel 2.5).

Standard metode innen DEA er likevel å inkludere slike variabler i en trinn 2 analyse, dvs. beregne effektivitetstall i en modell med de rene kvantitative målene, men så kjøre en regresjon med effektivitetstallene som avhengig variabel og kvalitetsindikatorer og andre kontekstuelle variabler som uavhengige variabler. Vi vil i det følgende bruke begrepet *ettermodell* om regresjonsanalysen. I den rene deterministiske DEA-metoden vil det da være problemer med uavhengigheten i de to trinnene, men ved bruk av bootstrap finnes det nå enkle metoder for å iterativt estimere de to trinnene samtidig (Simar & Wilson 2007). Det er uansett viktig å tolke resultatene som statistisk samvariasjon og ikke nødvendigvis som årsakssammenhenger.

I andre metoder som SFA og StoNED er det mulig å legge inn andre typer variabler direkte i samme trinn som effektivitetsestimeringen, enten ved at slike variabler skifter selve fronten av mulighetsområdet, eller ved at de endrer fordelingen av effektivitetsestimatene. Dette kommer vi tilbake til i senere avsnitt.

Den etablerte metodikken for måling av kommuneeffektivitet inneholder en rekke variabler som ikke er volummål. Som vi har forsøkt å synliggjøre med faktaboksen bryter slike alternative variabler med grunnforutsetningen i DEA-analysen, og krever i en strengt teknisk forstand en tilpassing av DEA-metoden for å gi korrekte analyser. I hvilken grad resultater fra standard DEA og tilpassede DEA-modeller avviker vil være et empirisk spørsmål. Det har vært utenfor prosjektets ressursrammer å gå videre inn på en empirisk studie av dette.

Selv om vår hovedanbefaling er å benytte variabler med volummål i DEA-modellen og variabler med andre målestokker som kontekstuelle variable i en ettermodell bør det presiseres at heller ikke dette nødvendigvis er en ideell løsning. Dette skyldes spesielt at en ettermodell kun er gyldig dersom den tilfredsstillende *separabilitetsantakelsen*, dvs. at kontekstuelle variabler kun påvirker kommunenes effektivitet og ikke deres produksjonsmuligheter (Simar og Wilson, 2007). I de senere år er det blitt utviklet en test for dette, og en anvendelse av Hampf og Rødseth (2019) viser at mange empiriske modellspesifikasjoner i praksis ikke tilfredsstillende separabilitetsantakelsen. Det er derfor grunn til å stille spørsmål ved den utbredte bruken av ettermodeller innen effektivitetsmåling. Det finnes andre metoder som ikke er avhengig av separabilitetsantakelsen som f.eks. conditional DEA (Daraio og Simar 2005; 2007). Men det finnes i liten grad litteratur som sier noe om hvordan variable med spesielle målestokker skal behandles innenfor dette rammeverket.

Et alternativ til bruk av en ettermodell er som nevnt å benytte egne DEA-modeller som er tilpasset slike variabler (jf. faktaboks). Vi tror at dette vil gjøre metodeopplegget for komplisert og lite transparent for anvendt måling av kommuneeffektivitet. SØFs metodiske opplegg benytter gjerne variabler med svært ulike målestokk i *en enkelt* analyse (eks. i produksjonsindeksen), og det er heller ikke klart hvordan man skal kombinere ulike spesialmetoder i en helhetlig modell. Generelt vil vi anbefale videre utredning av dette kompliserte området, men en foreløpig anbefaling er å benytte variabler med spesielle måleenheter i en ettermodell/regresjonsanalyse og kun benytte volumtall i DEA-modellen. Behandling av variabler med ulike måleenheter innenfor samme effektivitetsanalyse er et tema som er av interesse for den internasjonale litteraturen om produktivitets- og effektivitetsanalyse. Publisering av fagfelleverderte artikler om temaet bør derfor være en naturlig del av kvalitetssikringen av et revidert opplegg for måling av kommuneeffektivitet.

2.1.3 Skolebidragsindikatoren som eksempel

I SSBs dokumentasjon av [skolebidragsindikatoren](#) er denne forklart som (Steffensen et al, 2017, s. 8):

Skolebidraget kan dermed tolkes som det karakter- eller resultatgjennomsnittet vi forventer at en skole ville hatt, om elevene ved skolen var gjennomsnittlige med tanke på de elevkjenningene som er inkludert i beregningen.

I SSBs publiserte tall er derfor alle karakterene presentert i karakterskalaen 1-6, med et gjennomsnitt på 3,4 som er gjennomsnittskarakteren på landsbasis³. Riktignok har enkelte av trinnene i grunnskolen rapportert skalapoeng som er resultatene på nasjonale prøver. Disse har et snitt på rundt 50, men SSBs rapport har av pedagogiske grunner omgjort skalaene til en felles karakterskala for alle trinn. En slik multiplikativ endring er uten betydning for DEA-modellens effektivitetsestimater.

SSBs estimater på skolebidrag tar utgangspunkt i en regresjon der den enkelte elevs karakter er avhengig variabel og dennes bakgrunnsvariabler er uavhengige variabler. Slik får en et estimat på hvor mye bakgrunnsvariablene har å si for karakterene. La A_i^k være karakteren for hver elev i i kommune k . Blant de b bakgrunnsvariablene i vektoren Z_{hi}^k er foreldrenes utdanning, innvandringsbakgrunn og husholdsinntekt. For mellomtrinnet og ungdomsskolen kontrolleres også for elevens resultat på lavere trinn. Indikatoren får da tolkning som «value-added», dvs. hvor mye tilleggsferdigheter eleven har fått i forhold til sitt inngangsnivå. I sin enkleste form kan en da uttrykke karakteren for hver elev:

$$A_i^k = \beta_0 + \sum_h \beta_h Z_{hi}^k + \varepsilon_i^k \quad (2)$$

Skolens bidrag til karakterene finnes enklest ved å endre (2) og inkludere en fast faktor for hver skole lik for alle elevene på trinnet ved denne skolen. I vår sammenheng er vi interessert i en felles skolefaktor for alle skolene i kommunen og forenkler derved ved å se på en «fast effekt» b^k for hver kommune (kommunebidrag). Dette kan forstås som et konstantledd for hver kommune til erstatning for det felles konstantleddet β_0 i (2).

I tillegg er det ønskelig å forklare karakterene ved avvik fra gjennomsnittseleven i landet. For hvert kjennetegn Z_h vil en kunne skrive landsgjennomsnittet som $\bar{Z}_h = \sum_k \sum_i Z_{hi}^k / N$, hvor N er antall elever i landet. Regresjonen får da formen

$$A_i^k = b^k + \sum_h \beta_h (Z_{hi}^k - \bar{Z}_h) + \varepsilon_i^k \quad (3)$$

Denne omformingen vil ikke endre estimatene på koeffisientene β_h , kun på konstantleddene, dvs. de kommunefaste effektene b^k . Slik kan den enkelte elevs karakter skrives som summen av kommunebidraget b^k , bidraget fra foreldrebakgrunn og tidligere prestasjoner, og et tilfeldig feilledd. For landet som helhet vil forventningen til det andre leddet være null, i likhet med feilleddet. For den enkelte kommunen vil derimot gjennomsnittlig

³ Noe forvirrende brukes begrepet *skolebidrag* av Utdanningsdirektoratet (og SSB i sin dokumentasjon av leveransen til Udir) som avviker fra landsgjennomsnittet, og vil derfor ha et gjennomsnitt på 0. Et slikt tall er en additiv endring (trukket fra en konstant) og vil dessuten ha både negative og positive verdier. Disse tallene kan derfor ikke brukes direkte i en DEA-analyse. Også semantisk er det nok uheldig å si at en skole eller kommune yter et negativt bidrag til elevenes kunnskaper. SØF har riktignok hentet skolebidrag fra Udir men har lagt til igjen gjennomsnittlig skalapoeng for å rekonstruere SSBs opprinnelige skolebidragsindikator.

karakter avhenge av forskjellen gjennomsnittlig bakgrunnskjenntegn i kommunen og i landet.

Dette ser vi enklest ved å finne total karaktersum for kommunen:

$$\begin{aligned} \sum_i A_i^k &= N^k b^k + \sum_i \sum_h \beta_h (Z_{hi}^k - \bar{Z}_h) + \sum_i \varepsilon_i^k \\ &= N^k b^k + N^k \sum_h \beta_h (\bar{Z}_h^k - \bar{Z}_h) + \sum_i \varepsilon_i^k \end{aligned} \quad (4)$$

\bar{Z}_h^k er bakgrunnskjenntegnene til kommunens gjennomsnittselev og N^k er antall elever i kommune k . Siden forventningen til feilleddet er null, også for den enkelte kommunen, og vi lar gjennomsnittskaracteren i kommune k være a^k , vil vi kunne skrive kommunebidraget som:

$$\begin{aligned} b^k &= \frac{\sum_i A_i^k}{N^k} - \sum_h \beta_h (\bar{Z}_h^k - \bar{Z}_h) \\ &= a^k - \sum_h \beta_h (\bar{Z}_h^k - \bar{Z}_h) \end{aligned} \quad (5)$$

Her er andre ledd null dersom bakgrunnen til kommunens elever er lik landsgjennomsnittet. Kommunebidraget er derfor gjennomsnittskaracteren i kommunen, minus et korreksjonsledd for om kommunens elever har et bedre utgangspunkt ved å ha bedre verdi på bakgrunnsvariablene enn landsgjennomsnittet. Har f.eks. kommunen en høy gjennomsnittsinntekt vil en korrigerer en høy gjennomsnittskaracter nedover for å finne skolebidraget. En annen måte å uttrykke dette på er at kommunebidraget er den karakteren elevene i kommunen ville ha forventet å få dersom deres bakgrunn var lik landsgjennomsnittet, jfr. sitatet ovenfor.

Det er av betydning at karakterene er presentert som gjennomsnittskaracter per elev. Som nevnt i forrige avsnitt må variablene i en DEA-analyse ha samme dimensjonalitet som andre volummål, og ikke gjennomsnitt eller forholdstall. Det er totalkaracteren gitt ved (4) som da er det enkleste tjenestemålet, men dette er ikke korrigerert for bakgrunnsvariablene. I måling av effektivitet for undervisningssektoren benytter SØF derfor kommune-bidragsindikatoren ganget med antall elever som mål på produksjon i skolen, dvs.:

$$B^k = N^k b^k = N^k a^k - N^k \sum_h \beta_h (\bar{Z}_h^k - \bar{Z}_h) \quad (6)$$

Det totale kommunebidraget B^k er derved en kvalitetsjustert indeks som sammenfatter tre aspekter ved kommunens tjenesteproduksjon: Antall elever, karakternivå og elevbakgrunn. Elevbakgrunnen kunne igjen vært oppdelt i hvert enkelt kjenntegn Z_b .

Det er det totale kommunebidraget B^k som inngår i kostnadsfunksjonen (eller teknologien) som estimeres av SØF. For DEA-metoden er det totale kommunebidraget av riktig dimensjonalitet, men det er problematisk at det låser forholdet mellom de tre aspektene og deres relative betydning for ressursbruk eller kostnader.

Problemene er i korthet

- at substitusjonsforholdet mellom kvantitet (dvs. antall elever) og kvalitet (dvs. karakterer) er forhåndsdefinert og ikke bestemmes av DEA-modellen. Det er gitt at å øke elevtallet med en fast faktor er nøyaktig like ressurskrevende som å øke karakternivået (kommunebidraget) med samme faktor. Å øke fra f.eks. 1000 til 1250 elever vil være like ressurskrevende som å øke karakternivået fra 4 til 5.
- at betydningen av bakgrunnsvariablene er fastlåst på bakgrunn av en annen analyse (SSBs skolebidragsanalyse). Statistisk gir dette opphav til et endogenitetsproblem fordi sammenhengene ikke bestemmes simultant. Det kan også gi skjeve resultater fordi betydningen av elevenes bakgrunn kun går gjennom karakternivå. Det utelukkes f.eks. at ressursbruken også avhenger direkte av foreldrekjennetegn som inntekt, utdanning og innvandringsstatus.

2.1.4 Forslag til alternativ metodikk for undervisningssektoren

I dette avsnittet foreslår vi en alternativ modell som løser utfordringene vi har tatt opp i de foregående delkapitlene. Vår løsning tar utgangspunkt i at

- Elever og resultatmål (dvs. karakterer) bør vurderes som mål på henholdsvis kvantitet og kvalitet ved produksjonen. Vi bør a) tillate at det er substitusjonsmuligheter mellom disse i modellen vår og b) ikke gjøre noen forutsetninger om substitusjonsmulighetene *a priori*.
- Kontekstuelle forhold (familiebakgrunn ol.) tas hensyn til direkte i DEA-analysen, heller enn å benytte residualer fra en annen regresjonsanalyse (dvs. fra SSBs skolebidragsanalyse).

Ved å følge Johnson og Kuosmanen (2012) kan vi definere en ett-steps DEA-modell som følger:

Her vil substitusjonsforholdet mellom elever og karakterer bestemmes av δ -parameterne som estimeres av modellen, samt at effekten av kontekstuelle faktorer (Z) korrigeres for direkte i modellen. Det er dermed ingen utfordringer knyttet til at Z -variablene korrelerer med produktene i denne modellspefisikasjonen.

$$\min \sum_{k=1}^K (\varepsilon^k)^2$$

s.t.

$$\ln x^k = \ln \phi^k + \sum_h \beta_h \bar{Z}_h^k + \varepsilon^k, \forall k \quad (7)$$

$$\phi^k = \alpha^k + \delta_N^k N^k + \delta_A^k A^k, \forall k$$

$$\phi^k \geq \alpha^{k'} + \delta_N^{k'} N^k + \delta_A^{k'} A^k, \forall k, k'$$

$$\delta^k \geq 0, \forall k$$

$$\varepsilon^k \geq 0, \forall k$$

Her vil substitusjonsforholdet mellom elever og karakterer bestemmes av δ -parameterne som estimeres av modellen, samt at effekten av kontekstuelle faktorer (Z) korrigeres for direkte i modellen. Det er dermed ingen utfordringer knyttet til at Z -variablene korrelerer med produktene i denne modellspefisikasjonen.

En ulempe ved den skisserte løsningen i modellen (7) er at en bruker kommunegjennomsnittene for bakgrunnskjenetegnene slik at en ikke får utnyttet individvariasjonen slik SSB gjør når de estimerer (2). En mulig modifikasjon er da å bruke SSBs analyse til å fastlegge sammenhengen mellom elevbakgrunn og karakterer ved å benytte deres totale kommunebidrag B^k i stedet for totale karaktersum A^k . Alternativt kan en benytte det individestimerte avviket i forventet karaktersum ($A-B$) på kommunenivå som forklaringsvariabel, med andre ord la de faktiske skalapoeng inngå i kostnadsfunksjonen, mens forventede skalapoeng behandles som kontekstuelle variabler. En bør uansett inkludere elevtallet N^k og kommunegjennomsnittet for bakgrunnsvariablene \bar{Z}_n^k i modellen for å fange opp den direkte effekten av de ulike variablene på ressursbruken.

2.1.5 Behandling av variabler med spesielle måleenheter: Læringsmiljø

I SØF sin beskrivelse av variabelen læringsmiljø heter det «For læringsmiljø inngår snitt av samlescore for «Elevdemokrati og medvirkning», «Mobbing på skolen» og «Trivsel» for både 7. og 10. trinn. Mobbing er omdefinert ettersom lav score i rådata indikerer positiv verdi for mobbing. Multiplisert med antall elever.»

Mange av de samme innvendingene som i kapitlene 2.1.1 og 2.1.3 gjelder også for denne variabelen. Men i motsetningen til den mer kontinuerlige karakterskalaen innebærer læringsmiljøindikatoren en ordinal heller enn en numerisk (kardinal) rangering. DEA er i utgangspunktet ikke tilpasset denne type data. Vi viser til egen faktaboks over for detaljer.

2.1.6 Behandling av andre kontekstuelle faktorer

I kapitlene 2.1.1 og 2.1.3 drøftet vi standard behandling av kontekstuelle faktorer, nærmere bestemt elevers familiebakgrunn, i effektivitetsanalysen. Gjengs praksis i effektivitetslitteraturen er at slike faktorer behandles i en regresjonsanalyse enten simultant med estimering av DEA-modellen (som vist av (7)) eller som trinn 2 i et metodeopplegg hvor a) DEA brukes til å estimere effektivitet i første steg og b) regresjonsanalyse brukes til å forklare effekten av kontekstuelle variabler på teknisk effektivitet i andre steg (se den innflytelsesrike publikasjonen til Simar og Wilson (2007) for ytterligere detaljer). Uavhengig av hvilken modell som benyttes vil denne typen metodikk tilkjenne effekten av kontekstuelle variabler som passer best for det enkelte datasettet som benyttes i en analyse.

I stedet for å benytte konvensjonell metodikk for justering av kontekstuelle forhold benytter SØF justeringer *ex ante*, slik vi viste for den justerte skolebidragsindikatoren. Spesielt gjelder også dette for innsatsfaktoren driftskostnader i sektoranalysene for grunnskole og for pleie og omsorg (men ikke for barnehager), hvor materialet vi har mottatt beskriver følgende:

Kostnadene nedjusteres (oppjusteres) hos kommuner med spredt (konsentrert) bosettingsmønster målt ved reisetid (sone og nabo). Korrigeringen tar utgangspunkt i vektingen av sone- og nabokriteriene i delkostnadsnøkkelene gjennom følgende formel:

$$\frac{\text{rapportert innsatsfaktor}_{ij}}{(1 - \text{vekt}_{\text{sone}} - \text{vekt}_{\text{nabo}}) + \text{vekt}_{\text{sone}} \times \text{sonekrit}_i + \text{vekt}_{\text{nabo}} \times \text{nabokrit}_i},$$

hvor i angir kommune og j angir innsatsfaktor. Vektene er lik kriterienes vekt i delkostnadsnøkkelene. Dersom en kommune har en indeksverdi på både sone- og nabokriteriet over 1, det vil si at reisetid per innbygger i kommunen er høyere enn landsgjennomsnittet (spredt bosettingsmønster), vil nevneren i brøken få en verdi over 1. Dette medfører at korrigert kostnad blir lavere enn den faktiske/rapporterte innsatsfaktorbruken.

Vi finner ingen begrunnelse for valget av denne sektorspesifikke justeringsfaktoren (som vi senere vil referere til som sonevariabelen) i materialet vi har mottatt, utover at den er basert på delkostnadsnøklene i inntektssystemet for kommunene. Av kommuneproposisjonen for 2021 framgår det at disse er basert på ulike estimeringer i SSB og SØF med justeringer foretatt av departementet (Prop. 105 S. 2019-20, s. 22). Det er derfor vanskelig å vurdere kvaliteten i disse korreksjonsfaktorene. Det synes likevel som korreksjonen som brukes i inntektssystemet er vel innarbeidet i kommunesektoren, og det er rimelig å tro at bosetningsmønsteret vil ha betydning for eller samvariere med kostnadene.

Som for skolebidrag er det et spørsmål om informasjon om bosetningsmønsteret er utnyttet på rette måten i effektivitetsanalysene. Det er bl.a. lite sannsynlig at bosetningsmønsteret har noen betydning i barnehagesektoren. Mens skolebidragsindikatoren var estimert basert på individinformasjon om de enkelte elvene er det for bosetningsmønster antakelig kun brukt informasjon på kommunenivå. Da er det i utgangspunktet mer nærliggende å estimere effekten simultant som en del av effektivitetsanalysen. Dette kan f.eks. gjøres ved å inkludere sonekriteriet og nabokriteriet hver for seg eller den samlede indeksen som kontekstuelle variabler på samme måte som elevbakgrunn foreslås modellert i avsnitt 2.1.4 over.

2.1.7 Aggregering av effektivitetsindekser

Aggregering av effektiviteten til enkeltenheter (f.eks. firmaer) opp til et effektivitetsmål for en større enhet (f.eks. industrien) har fått mye oppmerksomhet i den teoretiske litteraturen om effektivitetsanalyse. Overordnet finnes det to ulike tilnærminger til aggregering: Med *strukturell effektivitet* menes samlet effektivitet i tilfellet hvor ressurser ligger fast per enhet, mens med *industrieffektivitet* menes samlet effektivitet i tilfellet hvor ressurser kan allokere mellom enheter i aggregeringen (se Färe og Karagiannis (2017) for diskusjon av disse konseptene og assosiert litteratur)⁴. Vi fokuserer på strukturell effektivitet i dette kapitlet.

SØF benytter budsjettvekter til å beregne strukturell effektivitet, dvs. til å utlede et samlet effektivitetsmål for kommunene ut fra effektivitetsmåling per delsektor. Vi finner ingen begrunnelse for denne vektningen i SØFs publikasjoner. Allikevel mener vi at dette gir en hensiktsmessig økonomisk tolkning av strukturell kommuneeffektivitet, i tråd med Färe og Zelenyuk (2003) sitt opplegg for aggregering av enheters effektivitet. For å illustrere dette lager vi et eksempel med 3 innsatsfaktorer ($x_1; x_2; x_3$), hvorav de to første inngår i produksjonen i sektor 1 (S1) mens x_3 inngår i produksjonen i sektor 2 (S2). Innsatsfaktorene har tilhørende priser w . θ er ressursbesparende effektivitet, som definert i likning (1). I dette tilfellet kan vi skrive strukturell effektivitet som

⁴ Merk at Farrell (1957) bruker disse to begrepene på en annen måte.

$$\begin{aligned}
& \frac{w_1 x_1 + w_2 x_2}{\sum_{n=1}^3 w_n x_n} \theta^{s1} + \frac{w_3 x_3}{\sum_{n=1}^3 w_n x_n} \theta^{s2} \\
&= \sum_{n=1}^2 \frac{w_n x_n}{\sum_{n=1}^3 w_n x_n} \left(\frac{\theta^{s1} x_n}{x_n} \right) + \frac{w_3 x_3}{\sum_{n=1}^3 w_n x_n} \left(\frac{\theta^{s2} x_3}{x_3} \right) \\
&= \frac{\sum_{n=1}^3 w_n x_n^*}{\sum_{n=1}^3 w_n x_n} \leq 1
\end{aligned} \tag{8}$$

hvor x^* er minimal ressursbruk etter at teknisk effektivitet er fjernet. Målet på strukturell effektivitet blir i dette tilfellet altså kommunens totale kostnader gitt teknisk effektiv ressursbruk delt på kommunens faktiske kostnader. I likhet med de sektor-spesifikke effektivitetsmålene vil dette gi et kommunalt effektivitetsmål som tar verdien 1 dersom kommunen har effektiv ressursbruk i alle sektorer og mindre enn 1 dersom det er ineffektiv tjenesteproduksjon i minst 1 sektor.

Merk at utledningen i (8) kun holder dersom det er de samme innsatsfaktorene som benyttes til estimeringen av effektivitet og beregning av aggregeringsvektene. Dersom budsjettandelene inneholder andre innsatsfaktorer/kostnader vil beviset bryte sammen.

Oppsummert finner vi at det å estimere effektivitet per sektor og å aggregere opp til kommunenivået med (riktige) budsjettvekter gir en korrekt økonomisk tolkning av strukturell effektivitet. En fordel med dette opplegget er at det typisk muliggjør å legge inn mange variabler i delsektor-modellene, noe som kan være vanskeligere dersom man skulle inkludert alle tjenester i en enkelt DEA-modell som direkte måler effektivitet per kommune (jf. tidligere diskusjoner om dimensjonalitetens forbannelse). Samtidig kan det være ressurser og/eller interne produkter som oversees av den sektorvise effektivitetsanalysen, eks. administrasjon og overhead, som ville blitt hensyntatt dersom man estimerer modellen for kommunenes samlede ressurser. Det finnes en rekke internasjonale studier som estimerer effektivitet per kommune heller enn per sektor (se eks. Balaguer-Coll mfl., 2010). Vi vil ikke gå videre inn i denne diskusjonen her, men et punkt for videre arbeid kan være å sammenlikne sektorvise og aggregerte effektivitetsmåling for kommuner.

2.2 Måling av effektivitetsendring over tid

Måling av produktivitets- og effektivitetsendring over tid er også et tema som er omfattende diskutert i litteraturen. Også på dette punktet har SØF utviklet et metodeopplegg som avviker fra normen i litteraturen om effektivitetsanalyse. Vi ønsker derfor først å beskrive SØFs opplegg, før vi ser på konvensjonell metodikk for måling av endring over tid.

SØF benytter data for to og to år i sine DEA-analyser, som i formen minner om det som i tidsserieanalytelitteraturen kalles «rolling window analysis». Eksempelvis er det gjort analyser for 2016-2017; 2017-2018; 2018-2019 osv. I hvert sett av analyser (f.eks. for årene 2018-2019) estimeres effektivitet per kommune (og sektor) for hver av de to aktuelle årene basert på en front konstruert på bakgrunn av de samme to år med data (jf. likning (1) for en generell beskrivelse av DEA-modellen). Effektivitetsendringen per kommune måles da ved forholdet mellom effektivitetstallet i de to årene (f. eks., effektivitet 2019/effektivitet 2018).

Videre sammenstilles tallene for en periode (f.eks. 2018-2019) med tidligere perioder (f.eks. 2016-2017; 2017-2018) for å vise effektivitetsendringen per kommune over tid.

Det er to egenskaper ved SØFs metodikk som kan gjøre det vanskelig å vurdere den reelle effektivitetsutviklingen en kommune har hatt:

1. Modellspesifikasjonen endres mellom år. Dette gjelder spesielt variablene (innsatsfaktorene og produktene) brukt i DEA-analysene, men også endring av metode. Eksempelvis ble årsverk tidligere benyttet som innsatsfaktorer i undervisningssektoren, men har i senere tid blitt erstattet med brutto driftsutgifter som eneste ressursmål. Analysen for 2018-19 bruker DEA med bootstrapping, noe som kan forventes å gi noe lavere effektivitetsestimater enn den deterministiske DEA-metoden brukt i tidligere år. Et spørsmål er derfor om en endring fra en periode til den neste er et resultat av at modellen er endret eller om effektiviteten er endret.
2. Selv om fronten er den samme innad i en periode (f.eks. 2018-2019) trenger ikke fronten være den samme i neste eller forrige periode (f.eks. 2017-2018) siden den er tilpasset på et annet datasett. Dersom fronten endrer seg kan en kommunes effektivitet endre seg uten at kommunen reelt sett har fått endret produktivitet.

For å kunne utdype disse poengene trenger vi litt mer notasjon knyttet til estimeringen av effektivitet. La t og $t+1$ definere to tidsperioder, og anta at definisjonen av innsatsfaktorer i modellen endres mellom periodene. La $\mathbf{x}=(x_0, x_1)$ og $\mathbf{y}=(y_0, y_1)$ dele inn innsatsfaktorene og produktene før og etter endringen av modelloppsettet.

$TE_{0,t}(\mathbf{x}_{0,t}^{k'}, \mathbf{y}_{0,t}^{k'}) =$ Teknisk effektivitet for enhet k' målt for ressursinnsats og produksjon observert i periode t i hht. modellspesifikasjon 0, mot en front laget på data fra periode t

$TE_{1,t+1}(\mathbf{x}_{1,t+1}^{k'}, \mathbf{y}_{1,t+1}^{k'}) =$ Teknisk effektivitet for enhet k' målt for ressursinnsats og produksjon observert i periode $t+1$ i hht. modellspesifikasjon 1, mot en front laget på data fra periode $t+1$

$TE_{0,t+1}(\mathbf{x}_{0,t+1}^{k'}, \mathbf{y}_{0,t+1}^{k'}) =$ Teknisk effektivitet for enhet k' målt for ressursinnsats og produksjon observert i periode $t+1$ i hht. modellspesifikasjon 0, mot en front laget på data fra periode $t+1$

$TE_{1,t}(\mathbf{x}_{1,t}^{k'}, \mathbf{y}_{1,t}^{k'}) =$ Teknisk effektivitet for enhet k' målt for ressursinnsats og produksjon observert i periode t i hht. modellspesifikasjon 1, mot en front laget på data fra periode t

2.2.1 Dekomponering av bidraget til endring i modelloppsettet til effektivitetsendringen

Hensikten med dette delkapitlet er å vise at effektivitetsendringen presentert i SØF sine rapporter ikke viser en «ren» effektivitetsendring, men også påvirkes av endring i modellspesifikasjonen over tid. Formlene som utledes kan brukes til å beregne den faktiske eller «rene» effektivitetsendringen.

La oss først definere effektivitetsendringen fra periode t til $t+1$ som:

$$TEC_{t,t+1} = TE_{1,t+1}(\mathbf{x}_{1,t+1}^{k'}, \mathbf{y}_{1,t+1}^{k'}) / TE_{0,t}(\mathbf{x}_{0,t}^{k'}, \mathbf{y}_{0,t}^{k'}) \quad (9)$$

Anta nå at modellspesifikasjonen (f.eks. variablene) endrer seg mellom periode t og $t+1$. Vi kan da dekomponere (9) som følger:

$$\begin{aligned}
 TEC_{t,t+1} &= \frac{TE_{1,t+1}(\mathbf{x}_{1,t+1}^{k'}, \mathbf{y}_{1,t+1}^{k'})}{TE_{0,t}(\mathbf{x}_{0,t}^{k'}, \mathbf{y}_{0,t}^{k'})} \\
 &= \frac{TE_{1,t+1}(\mathbf{x}_{1,t+1}^{k'}, \mathbf{y}_{1,t+1}^{k'})}{\underbrace{TE_{0,t+1}(\mathbf{x}_{0,t+1}^{k'}, \mathbf{y}_{0,t+1}^{k'})}_{\text{Modellendring}}} \frac{TE_{0,t+1}(\mathbf{x}_{0,t+1}^{k'}, \mathbf{y}_{0,t+1}^{k'})}{\underbrace{TE_{0,t}(\mathbf{x}_{0,t}^{k'}, \mathbf{y}_{0,t}^{k'})}_{\text{Effektivitetsendring}}} \\
 &= \frac{TE_{1,t}(\mathbf{x}_{1,t}^{k'}, \mathbf{y}_{1,t}^{k'})}{\underbrace{TE_{0,t}(\mathbf{x}_{0,t}^{k'}, \mathbf{y}_{0,t}^{k'})}_{\text{Modellendring}}} \frac{TE_{1,t+1}(\mathbf{x}_{1,t+1}^{k'}, \mathbf{y}_{1,t+1}^{k'})}{\underbrace{TE_{1,t}(\mathbf{x}_{1,t}^{k'}, \mathbf{y}_{1,t}^{k'})}_{\text{Effektivitetsendring}}}
 \end{aligned} \tag{10}$$

Likning (10) viser at SØFs effektivitetsendring kan dekomponeres i to effekter, nemlig «ren» effektivitetsendring og konsekvenser av at modellspesifikasjonen har endret seg. For å unngå å velge basisår kan man eventuelt definere dekomponeringen som et geometrisk gjennomsnitt av endringene beregnet ved de to basene:

$$\begin{aligned}
 TEC_{t,t+1} &= \left[\frac{TE_{1,t}(\mathbf{x}_{1,t}^{k'}, \mathbf{y}_{1,t}^{k'})}{TE_{0,t}(\mathbf{x}_{0,t}^{k'}, \mathbf{y}_{0,t}^{k'})} \frac{TE_{1,t+1}(\mathbf{x}_{1,t+1}^{k'}, \mathbf{y}_{1,t+1}^{k'})}{TE_{0,t+1}(\mathbf{x}_{0,t+1}^{k'}, \mathbf{y}_{0,t+1}^{k'})} \right]^{0.5} \\
 &\quad \underbrace{\hspace{10em}}_{\text{Modellendring}} \\
 &\times \left[\frac{TE_{0,t+1}(\mathbf{x}_{0,t+1}^{k'}, \mathbf{y}_{0,t+1}^{k'})}{TE_{0,t}(\mathbf{x}_{0,t}^{k'}, \mathbf{y}_{0,t}^{k'})} \frac{TE_{1,t+1}(\mathbf{x}_{1,t+1}^{k'}, \mathbf{y}_{1,t+1}^{k'})}{TE_{1,t}(\mathbf{x}_{1,t}^{k'}, \mathbf{y}_{1,t}^{k'})} \right]^{0.5} \\
 &\quad \underbrace{\hspace{10em}}_{\text{Effektivitetsendring}}
 \end{aligned} \tag{11}$$

Merk at problemet med at effektivitetsendringen blir en sammenflettet effekt av modell- og effektivitetsendringer kan enkelt omgås ved å benytte identisk modellspesifikasjon for alle år som analyseres i effektivitetsstudien. I dette tilfellet faller helt enkelt leddet «modellendring» bort, og man står igjen med den «rene» effektivitetsendringen.

Oppdragsgiver beskriver at årsaken til at modellen er i endring er at man til enhver tid ønsker å benytte de beste variablene som er tilgjengelige. Dette er et viktig hensyn, men det kan også tilsløre effektivitetsutviklingen. Det er i praksis to mulige veier rundt dette:

1. Det kan være mulig å etablere to modeller: En basismodell som er mulig å estimere konsistent over tid og en mer kompleks modell som benytter de beste dataene i hver tidsperiode. Resultatene fra basismodellen kan benyttes til å beregne effektivitetsendringen som vist i avsnitt 2.3 nedenfor, mens resultatene fra de to modellene sammenliknes for å identifisere eventuelle avvik som skyldes modellspesifikasjon.
2. Å tillate at modellspesifikasjonen endres over tid men – om mulig – dekomponerer endringen i modell- og effektivitetsendring ved å anvende formelapparatet som er utviklet i dette delkapitlet.

Generelt kan strukturelle skift i data være et problem for vurdering av effektivitetsutviklingen over tid. Vi er ikke kjent med at dette er godt behandlet i litteraturen, og vi anser det derfor som et relevant område for videre forskning.

2.2.2 Dekomponering av produktivetsendring

I dette delkapitlet ser vi bort fra at metodene, innsatsfaktorene og produktene endres over tid, men fokuserer på at dataene som brukes og derfor også fronten som estimeres kan endre seg over tid.

Et viktig poeng er at SØF måler effektivitetsendring fra f.eks. 2018 til 2019 basert på samme front (konstruert basert på data fra hhv. 2018 og 2019). Men i sammenstillingen av effektivitetsendring over flere perioder er det ingen garanti for at fronten som ligger bak analysen ikke har endret seg.

Det er viktig å forstå at a) dersom fronten endrer seg mens b) *produktiviteten* til en kommune er konstant vil kommunens *effektivitet* endre seg. I denne sammenhengen betyr det at SØFs beregnede endring for hver kommune vil være et resultat både av effektivitetsendring og av frontendring, selv ved gitt modellspekifisasjon. For å vise dette går vi tilbake til notasjonen fra forrige delkapittel, men dropper nå indeksen for modellspekifisasjonene 0 og 1 (m.a.o., vi antar at modellspekifisasjonen er uendret i perioden vi ser på). Vi ser på 3 perioder; t , $t+1$ og $t+2$, og bruker toppskrift til å betegne hvilke perioder fronten er estimert på og fotskrift til å betegne hvilket år observasjonen er fra. I henhold til SØFs opplegg definerer vi da for hvert par av år:

$$TEC_{t,t+1}^{k'} = \frac{TE_{t,t+1}(\mathbf{x}_{t+1}^{k'}, \mathbf{y}_{t+1}^{k'})}{TE_{t,t+1}(\mathbf{x}_t^{k'}, \mathbf{y}_t^{k'})} \quad (12)$$

$$TEC_{t+1,t+2}^{k'} = \frac{TE_{t+1,t+2}(\mathbf{x}_{t+2}^{k'}, \mathbf{y}_{t+2}^{k'})}{TE_{t+1,t+2}(\mathbf{x}_{t+1}^{k'}, \mathbf{y}_{t+1}^{k'})}$$

Endringen i effektivitet fra år t til $t+1$ er målt som effektiviteten til observasjon $t+1$ dividert med effektiviteten til observasjon t , begge målt opp mot fronten i de to periodene $t, t+1$ samlet. Når det i SØFs analyser beregnes endring over flere år ved kjeding multipliseres de årlige endringene sammen. Vi kan da vise at en beregnet endring fra t til $t+2$ er en kombinert effekt av at fronten og effektiviteten har endret seg:

$$TEC_{t,t+2} = TEC_{t,t+1} * TEC_{t+1,t+2}$$

$$= \frac{TE_{t,t+1}(\mathbf{x}_{t+1}^{k'}, \mathbf{y}_{t+1}^{k'})}{TE_{t,t+1}(\mathbf{x}_t^{k'}, \mathbf{y}_t^{k'})} * \frac{TE_{t+1,t+2}(\mathbf{x}_{t+2}^{k'}, \mathbf{y}_{t+2}^{k'})}{TE_{t+1,t+2}(\mathbf{x}_{t+1}^{k'}, \mathbf{y}_{t+1}^{k'})} \quad (13)$$

$$= \underbrace{\frac{TE_{t,t+1}(\mathbf{x}_{t+1}^{k'}, \mathbf{y}_{t+1}^{k'})}{TE_{t+1,t+2}(\mathbf{x}_{t+1}^{k'}, \mathbf{y}_{t+1}^{k'})}}_{\text{Frontendring}} * \underbrace{\frac{TE_{t+1,t+2}(\mathbf{x}_{t+2}^{k'}, \mathbf{y}_{t+2}^{k'})}{TE_{t,t+1}(\mathbf{x}_t^{k'}, \mathbf{y}_t^{k'})}}_{\text{Effektivitetsendring}}$$

Likning (13) gir formler som kan benyttes direkte til å dekomponere hvorvidt endringen i målt kommuneeffektivitet skyldes at fronten har endret seg, eller at kommunen er blitt mer/mindre effektiv. Dersom komponenten frontendring er større enn 1 betyr det at frontendring har bidratt til forbedret effektivitet, mens hvis den er lavere enn 1 betyr det at frontendringen har bidratt til at kommunen framstår som mindre effektiv.

2.2.3 Malmquistindeksen

Det er viktig å presisere av SØFs opplegg for å vurdere effektivitetsendring over tid ikke er konvensjonell metodikk innen effektivitetsmåling: Innen effektivitetsanalysen er det vanlig å benytte den såkalte Malmquistindeksen (Malmquist, 1953; Caves mfl. 1982) til å måle produktivitetens utviklingen.

Malmquistindeksen kan brukes enten frontene estimeres ved DEA, SFA eller andre metoder. Skal en måle endringer over tid er det viktig å ta utgangspunkt i målt produktivitet i forhold til en fast referanse og ikke effektivitet i forhold til teknologifronten⁵. (Fotnote inn her) Deretter kan en dekomponere produktivitetsendring.

Ved beregning av produktivitetsendring er det også nødvendig å bruke en fast referanse, slik f.eks. prisindekser bruker faste vektorer. Den faste referansen kan f.eks. være fronten i første periode, fronten i siste periode, eller en referansefront som utnytter observasjonene i alle år. I DEA-metoden vil dette svare til at indeksen k i (1) løper over observasjonene i første år, siste år eller alle år, uavhengig av hvilket år vi har kommuneobservasjonen k' vi skal måle produktiviteten til. Som ovenfor bruker vi toppskrift til å betegne hvilke perioder fronten eller referansen er estimert på og fotskrift til å betegne hvilke(t) år observasjon(e) er i. Toppskrift f er for den faste referansen vi har valgt.

Vi kan da definere Malmquistindeksen for produktivitetsendringen til en kommune mellom periodene t og s målt mot den felles referansefronten f som

$$MI_{ts}^f = \frac{TP_s^f}{TP_t^f} \quad (14)$$

Merk at vi i dette avsnittet har utelatt indeksen for kommunen vi ser på for å forenkle uttrykkene. Siden teknologifronten kan ha endret seg mellom t og s kan vi dekomponere indeksen i frontendring og faktisk endring av hvor langt kommunen ligger bak fronten (oppheving) på hvert tidspunkt på følgende måte:

$$MI_{ts}^f = \underbrace{\frac{TP_s^s}{TP_t^t}}_{\text{Oppheving}} \times \underbrace{\left[\frac{TP_s^f}{TP_s^s} \times \frac{TP_t^t}{TP_t^f} \right]}_{\text{Frontendring}} \quad (15)$$

som en kan kort uttrykke som $MI=MC \times MF$. Opphevingen måler om kommunen har kommet nærmere den maksimale produktiviteten i egen periode. Frontendringskomponenten måler produktiviteten dobbelt relativt ved at frontene i de to årene sammenlignes med den felles referansefronten, både ved første og siste observasjon. Komponentene tar verdier over 1 dersom hhv oppheving eller frontendring har bidratt til økt produktivitet over tid. Frontendringen tar kun verdier større eller lik 1 hvis en ikke tillater teknologisk tilbakegang.

Siden opphevingen er endring i teknisk produktivitet målt i forhold til maksimal produktivitet i samme periode som observasjonen, kan vi ved å bruke $TP=TE \times SE$ dekomponere dette leddet videre i endring i teknisk effektivitet og i skalaeffektivitet:

⁵ Når referansefunksjonen er fast og både den og beste praksis funksjon har konstant skalaavkastning, så kan endringen over tid også gjøres ved det relative forhold mellom effektivitetstall for to perioder.

$$MI_{ts}^f = \underbrace{\frac{TP_s^s}{TP_s^s}}_{\text{Opphenting}} \times \underbrace{\left[\frac{TP_s^f}{TP_s^s} \times \frac{TP_t^t}{TP_t^f} \right]}_{\text{Frontending}} = \underbrace{\frac{TE_s^s}{TE_t^t}}_{\text{Effektivitetsendring}} \times \underbrace{\frac{SE_s^s}{SE_t^t}}_{\text{Skalaeffektivitetsendring}} \times \underbrace{\left[\frac{TP_s^f}{TP_s^s} \times \frac{TP_t^t}{TP_t^f} \right]}_{\text{Frontending}} \quad (16)$$

Strengt tatt trenger vi ikke toppskrift s og t på teknisk effektivitet TE og skalaeffektivitet SE siden disse alltid måles mot fronten i samme periode. Enkelt kan vi skrive dette som $MI=MC \times MF=ME \times MS \times MF$.

Vi anbefaler at Malmquist-indeksen tas i bruk til å vurdere endring i produktivitet over tid, og at indeksen dekomponeres i front-, skala- og effektivitetsendring. Dette for å sikre at det metodiske opplegget er i henhold til internasjonale standarder, noe som også kan gjøre internasjonale sammenlikninger enklere. Merk allikevel at en viktig forskjell mellom Malmquist-indeksen (16) og SØFs opplegg (13) er at SØF definerer effektivitetsendring målt mot parvis samme front, mens den konvensjonelle Malmquist-metodikken legger til grunn at effektivitetsendring vurderes basert på effektivitet målt mot periode-spesifikke fronter og produktivitetsendring vurderes mot en felles referansefront for alle perioder.

2.2.4 Oppsummering av utviklingen over tid

Bruk av Malmquistindeksen og dens komponenter basert på en fast modell gir viktige fordeler framfor SØFs nåværende innfallsvinkel til å analysere utviklingen over tid.

1. Den gir resultater som er lett tolkbare og basert på samme definisjoner over tid. I SØFs opplegg vet en ikke om en endring over et lengre tidsrom skyldes endring i variabelspesifikasjonen, i dimensjonaliteten til modellen, i bruk av bootstrap for å korrigere effektivitetsestimaterne, eller i faktisk endring av kommunenes effektivitet mellom tidspunktene.
2. Den gir informasjon om mer enn effektivitetsutviklingen. Det er f.eks. ikke slik at en bedring i effektiviteten over tid er det samme som at kommunene får mere tjenesteproduksjon ut av ressursene. Det kan bare måles ved produktiviteten slik Malmquistindeksen måler fordi denne relateres til en fast referanse uanvveghig av skala og tid. Effektivitetsendring sier kun noe om hvor godt kommunene gjør det relatert til mulighetene på ethvert tidspunkt og gitt sin egen størrelse. Malmquist-analysen gir også som delkomponent en effektivitetsindeks, samt komponenter som måler endring i mulighetsområdet (teknologien) og i optimal skala.

Utfordringen ved Malmquistindeksen er kun at en må ha en enkel modellspesifikasjon hvor det er mulig å skaffe data tilbake i tid og må gjennomføre en større modellkjøring over alle relevante år på nytt. En enkel modell er likevel enklere å forstå. En enkel modell tilbake i tid forhindrer dessuten ikke at rikere modellspesifikasjoner kan brukes i analysen for det siste året eller de siste årene.

2.3 Effektivitetsanalyse basert på produksjonsindeksen

Produksjonsindeksen ble først utviklet av Borge mfl. (2001), og en revidert indeks er presentert i Borge og Tovmo (2009). Indeksen er ment som et mål på kommunenes samlede tjenestetilbud. Den er hovedsakelig utviklet med sikte på å måle produksjon og ikke effektivitet, men i denne rapporten ønsker vi å sette søkelys på bruken av indeksen til måling av produktivitet og effektivitet.

Overordnet kan indeksen beskrives som bestående av et sett av kriterier med tilhørende vektorer, hvor vektene i hovedsak er budsjettandeler. Det beregnes delindekser for barnehage, grunnskole, primærhelsetjeneste, pleie og omsorg, barnevern, sosialkontortjenester og kultur, samt den overordnede produksjonsindeksen som omfatter alle disse tjenestene.

Borge mfl. (2001) definerte ulike krav som de mente en ideell produksjonsindeks for kommunene bør tilfredsstill:

- Dersom kommune A produserer mer enn kommune B av minst en tjeneste og like mye som kommune B av alle andre tjenester, så har kommune A en høyere indeksverdi enn kommune B
- To kommuner A og B har like økonomiske rammebetingelser. Dersom begge kommunene utnytter ressursene effektivt men prioriterer forskjellig, så har de to kommunene samme indeksverdi.
- To kommuner A og B har like økonomiske rammebetingelser. Dersom kommune A utnytter sine ressurser effektivt mens kommune B er ineffektiv, så har kommune A en høyere indeksverdi enn kommune B
- Kommune A har bedre økonomiske rammebetingelser enn kommune B. Dersom begge kommuner utnytter sine ressurser effektivt, så har kommune A en høyere indeksverdi enn kommune B.

De finner at det i praksis er vanskelig å definere en ideell indeks, og de baserer seg på en additiv indeks med nasjonale budsjettandeler som vektorer. Borge og Tovmo (2009) peker på flere problemer med dette, bl.a. at kommuner med høye enhetskostnader tillegges større vekt enn kommuner med lave enhetskostnader. De peker også på at det i mange tilfeller kan være riktig å la kvantitet og kvalitet i tjenesteproduksjonen inngå multipliktivt i indikatoren, men at det i praksis legges opp til at kvantitets- og kvalitetsindikatorer inngår additivt. Begrunnelsen er at det er vanskelig å finne gode indikatorer for kvalitet, og at en multipliktiv tilnærming har en tendens til å «blåse opp» eller overdrive måleproblemene.

Som i DEA-analysen lar vi $\mathbf{y} \in \mathfrak{R}_+^M$ være en vektor av kriterier (eller produkter) som samlet beskriver kommunal tjenesteproduksjon, mens $\mathbf{p} \in \mathfrak{R}_+^M$ er et sett av tilhørende vektorer. I tråd med Borge mfl. (2001) deler vi disse inn i undergrupper, som hver er tilknyttet en sektor innen kommunal tjenesteproduksjon. Vi henviser til disse med indeksen $s=1,..,S$, mens $m_s=1,..,M_s$ er delsettet av kriterier som tilhører sektor s . I henhold til Borge og Tovmo (2009) kan vi da formelt definere delindeksen for sektor s og kommune k' som følger

$$DPI_s^{k'} = \sum_{m_s=1}^{M_s} p_m \left(\frac{y_m^{k'}}{\frac{1}{K} \sum_{k=1}^K y_m^k} \right) = \sum_{m_s=1}^{M_s} p_m \left(\frac{y_m^{k'}}{\bar{y}_m} \right). \quad (17)$$

Hver delindeks er altså en vektet sum av kriterier tilhørende sektor s , og hvor hvert av kriteriene normaliseres med gjennomsnittet av kriteriet over alle kommuner. Merk at det i SØFs opplegg benyttes vektete gjennomsnitt, hvor antall innbyggere per kommune brukes som vektorer. Men for å unngå kompliserende notasjon overser vi dette her.

I Borge mfl. (2001) forklares normaliseringen (dvs. at hvert kriterium deles på det tilsvarende landsgjennomsnittet) med at indeksen forklarer produksjonen i hver kommune relativt til det nasjonale gjennomsnittet. Det legges til grunn at rangeringen av kommunenes produksjon vil være den samme om man beregner indeksen uten å normalisere den. I neste kapittel viser vi at dette ikke er riktig.

La β_s være vekter per sektor. Den totale produksjonsindeksen for kommune k er da

$$I^{k'} = \sum_{s=1}^S \beta_s \sum_{m_s=1}^{M_s} p_m \left(\frac{y_m^{k'}}{y_m} \right). \quad (18)$$

I Borge mfl. (2011) drøftes en utviding av antall kriterier i produksjonsindeksen for et utvalg av kommuner (såkalte ASSS-kommuner). Her benyttes også indeksen som et mål på produktivitet ved at den deles på kommunenes kostnader. Tanken er å lage et indekstall som viser forholdet mellom produksjon og ressursinnsats. Borge mfl. (2011) refererer til dette som effektivitet, men dette er ikke en riktig definisjon siden det ikke benyttes noen sammenlikning med beste praksis.

I det følgende vil vi peke på mulige forbedringspunkter for produksjonsindekismetoden.

2.3.1 Indeksens egenskaper

Det finnes en stor vitenskapelig litteratur om såkalt indeksnummerteor, hvor de matematiske egenskapene til en indeks vies mye plass. Vår forståelse er at SØFs produksjonsindeks ikke er forankret i denne litteraturen, samt at man i liten grad har viet egenskapene til produksjonsindeksen mye oppmerksomhet, ut over drøftingen som gjøres i Borge mfl. (2001). Vi går ikke i detalj inn på dette her, men en videre undersøkelse kan være nyttig for bedre å forstå styrker og svakheter ved indeksen.

2.3.1.1 Normalisering av indeksen

Borge mfl. (2001) beskriver at rangeringen av kommunene ikke påvirkes av at indeksen normaliseres ved å dele hvert kriterium på det tilsvarende kommunale gjennomsnittet (jf. likning (17)). For å vise at dette ikke er riktig gir vi et regneeksempel med 3 kommuner og 3 kriterier i tabell 2.1. Denne beregner indeksen med og uten normalisering, og viser at kommunene A og C får like indeksnummer i den uvektede indeksen, mens de avviker i den vektete. Dette viser at Borge mfl. (2001) sin påstand om at normaliseringen i likning (17) ikke har betydning for kommunenes rangering ikke nødvendigvis stemmer.

Tabell 2.1: Illustrasjon av produksjonsindeksens transitivitet.

Kommune	p^1	y^1	$y^1/\text{Gj.snitt}$	p^2	y^2	$y^2/\text{Gj.snitt}$	p^3	y^3	$y^3/\text{Gj.snitt}$	Uten normalisering	Med normalisering
A	0,40	5,00	0,50	0,20	10,00	0,86	0,30	15,00	1,80	8,50	0,91
B	0,40	15,00	1,50	0,20	10,00	0,86	0,30	5,00	0,60	9,50	0,95
C	0,40	10,00	1,00	0,20	15,00	1,29	0,30	5,00	0,60	8,50	0,84

En alternativ normalisering av indeksen som oppfyller kravet om at rangeringen skal være upåvirket vil være å normalisere mht. den gjennomsnittlige indeksen:

$$DPI_S^{k'} = \frac{\sum_{m_s=1}^{M_s} p_m y_m^{k'}}{\frac{1}{K} \sum_{k=1}^K \sum_{m_s=1}^{M_s} p_m y_m^k}. \quad (19)$$

som alltid vil tilfredsstillende dersom

$$\sum_{m_s=1}^{M_s} p_m y_m^{k'} > \sum_{m_s=1}^{M_s} p_m y_m^{k''} . \quad (20)$$

så er også den normaliserte produksjonsindeksen i (19) alltid høyere for k' enn k'' . Dette kan enkelt bevises da nevnerne i den første likning under er like og telleren for kommune k' er større enn for kommune k'' :

$$\begin{aligned} & \frac{\sum_{m_s=1}^{M_s} p_m y_m^{k'}}{\frac{1}{K} \sum_{k=1}^K \sum_{m_s=1}^{M_s} p_m y_m^k} > \frac{\sum_{m_s=1}^{M_s} p_m y_m^{k''}}{\frac{1}{K} \sum_{k=1}^K \sum_{m_s=1}^{M_s} p_m y_m^k} \\ & = \sum_{m_s=1}^{M_s} p_m y_m^{k'} > \frac{\sum_{m_s=1}^{M_s} p_m y_m^{k''} \frac{1}{K} \sum_{k=1}^K \sum_{m_s=1}^{M_s} p_m y_m^k}{\frac{1}{K} \sum_{k=1}^K \sum_{m_s=1}^{M_s} p_m y_m^k} . \\ & = \sum_{m_s=1}^{M_s} p_m y_m^{k'} > \sum_{m_s=1}^{M_s} p_m y_m^{k''} \end{aligned} \quad (21)$$

2.3.1.2 Valg av vektor og deres tolkning

I likning (18) spesifiseres produksjonsindeksen som summen av utvalgte kriterier (dvs. produkter) og deres tilhørende vektorer. Dersom det hadde eksistert velfungerende markeder for kommunale tjenester ville bruk av produktenes respektive salgspriser vært hensiktsmessig. På denne måten ville markedets betalingsvillighet for tjenestene kommet til uttrykk gjennom prisene, samtidig som analysen kunne forankres rundt kommunale inntekter og/eller profitt. Samtidig ville bruk av salgspriser gi et korrekt uttrykk for strukturell effektivitet, parallelt med vår utredning om aggregering basert på innsatsfaktorpriser i kapittel 2.1.7.

En utfordring forbundet med måling av effektivitet i offentlig sektor er at priser for varer og tjenester er mangelfulle. Borge mfl. (2001) baserer derfor vektene på *innsatsfaktorpriser*.

Konkret definerer de gjennomsnittlige enhetskostnader \bar{c}_m for hvert enkelt produkt m , og vekten til produkt m som dets gjennomsnittlig *budsjettandel*:

$$p_m = \frac{\sum_{k=1}^K \bar{c}_m y_m^k}{\sum_{k=1}^K \sum_{m=1}^M \bar{c}_m y_m^k} . \quad (22)$$

Vi gjør tilleggsforutsetningen

$$c_m^k = c_m q^k . \quad (23)$$

hvor q^k er en indikator på kostnadsnivået til kommune k , og c_m er en forventet produksjonskostnad per enhet produsert av m . Dette kan tolkes som om alle kommuner har den samme initiale produksjonskostnaden per enhet, men at denne reelt sett blir ulik mellom kommuner på grunn av ulikheter i kostnadsnivået.

Med dette utgangspunktet kan vi omskrive produksjonsindeksen som (se Borge mfl., 2001):

$$I^{k'} = \frac{\sum_{m=1}^M (c_m^{k'} y_m^{k'} / q^{k'})}{\frac{1}{K} \sum_{k=1}^K \sum_{m=1}^M (c_m^k y_m^k / q^k)}. \quad (24)$$

Det vil si at indeksen er definert ved forholdet mellom totalt budsjett i kommune k' normalisert med kostnadsnivået i denne kommunen, relativt til gjennomsnittlig budsjett for alle kommuner⁶. I lyset av dette er viktige egenskaper ved produksjonsindeksen:

- 1) Dersom alle kommunene har likt kostnadsnivå (q) er produksjonsindeksen gitt ved forholdet mellom en kommunes budsjett og gjennomsnittlig kommunebudsjett for hele landet.
- 2) To kommuner med likt kostnadsnivå og budsjett har lik indeksverdi, uavhengig av hvordan de prioriterer mellom produksjonen av ulike varer og tjenester.
- 3) Hvis budsjettet til kommunene A og B er likt, men A har høyere kostnadsnivå (q) enn B, så er produksjonsindeksen høyere for B.
- 4) Vektingen forutsetter konstante enhetskostnader. Dette innebærer at modellen ser bort fra at marginale kostnader og dermed også substitusjonsmulighetene mellom produkter avhenger av produsert mengde. På tilsvarende måte neglisjeres skalaavkastning i produksjonen.

Vi ønsker å stille spørsmål ved flere av disse egenskapene:

Uavhengighet av prioritering

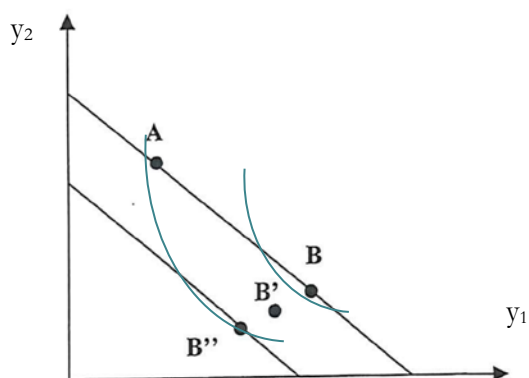
Egenskap 2) sier at kommuner med identisk kostnadsnivå (q) og budsjett får lik indeksverdi, selv om de prioriterer helt forskjellig mht. hvilke varer og tjenester de tilbyr. Dette svarer til ett av de fire ideelle indekseegenskapene som ble postulert av Borge mfl. (2001) og som ble presentert i innledningen til kapittel 2.3. I Figur 2.2, basert på Borge mfl. (2001), illustreres dette prinsippet ved at kommunene A og B er plassert langs den samme budsjettlinjen, noe som innebærer at budsjettet og dermed produksjonsindeksene er identisk.

Forutsetningen om at prioritering ikke er av betydning står i kontrast til den generelle konsumentteorien innen mikroøkonomi. I sitt alternative opplegg til å måle effektivitet i offentlig sektor legger Forsund (2017)⁷ til grunn at befolkningen har *preferanser* for kommunale tjenester, og at disse kan uttrykkes ved en nytte- eller preferansefunksjon. Innenfor

⁶ Formelt er dette identisk med helsesektorens bruk av DRG-poeng for å måle produksjonen i spesialisthelse-tjenesten. For hver diagnoserelaterte gruppe (DRG) beregnes vektor basert på gjennomsnittskostnaden for å behandle en pasient. Der blir vektene normalisert ved å dele på gjennomsnittskostnaden over alle pasienter heller enn som budsjettandeler (Helsedirektoreatet, 2021).

⁷ Borge mfl. (2020) innfører begrepene D- og C-goder fra Bradford mfl. (1969), der D står for direkte konsumgoder som kommuneenheten produserer og C står for tjenestene konsumentene oppfatter. Borge mfl. (2020) gjør ikke noe ut av denne distinksjonen. I Forsund (2017) er ideene til Bradford mfl. (1969) videreutviklet ved å modellere at C-tjenester er en funksjon av D-tjenester. Prioritering bestemmes ved å maksimere en preferansefunksjon over C-tjenestene gitt en budsjettbetingelse for ressurser og funksjonene som gir sammenhengene mellom C-tjenester og D-tjenester, samt produktfunksjoner for D-tjenester.

dette rammeverket blir målsetningen til kommunene å maksimere velferd gitt sine budsjettbetingelser. Dette illustreres i figur 2.2 ved å tegne inn såkalte «indifferenskurver» som viser kombinasjonen av kommunale tjenester som gir innbyggerne den samme nytten. Jo lenger utover (mot høyre) i figuren kurven tegnes, dess høyere er nytten til innbyggerne.



Figur 2.2: Nytteoptimering under budsjettbetingelse.

Når vi legger til grunn velferdsvekter i beregningen av produksjonsindeksen kan rangeringen av kommunene bli en helt annen enn det Borge mfl. (2001) sin indeks beskriver. I figur 2.2 beskrives en situasjon hvor kommune B produserer en miks av tjenester som gir langt høyere nytte for samfunnet enn kommune A, til tross for at begge har identiske utgifter.

I et hypotetisk velfungerende marked for kommunale tjenester ville det vært naturlig å bruke tjenestenes salgspriser som vekter til å beregne produksjonsindeksen, ikke budsjettandeler. De ville da gi et mål på kommunal inntekt direkte. Men siden kommunale tjenester ikke omsettes i markeder må det brukes andre metoder for å finne slike ideelle vekter. Et typisk virkemiddel vil være å anvende såkalt Stated Preference-teknikker. Dette inngår som standard metodikk i nyttekostnadsanalyse i en rekke sektorer i Norge. Et eksempel er transportsektoren, hvor bl.a. verdsetting av tid utgjør blant de viktigste enhetsprisene. Vi innser at utledning av alternative vekter basert på en sosial nyttefunksjon kan vise seg både vanskelig og svært ressurskrevende. I det følgende vil vi derfor også se på metodikk til effektivitetsmåling som tillater at vi ikke må benytte priser eller kostnader i beregningen av effektivitet i produksjonsindekssammenheng.

Påvirkning av kostnadsnivået

Som vist av likning (24) innebærer en økning i kostnadsnivået til kommune k' en reduksjon i produksjonsindeksverdien til kommunen. Et spørsmål er hvordan kostnadsnivåindikatoren q skal tolkes. Dersom den gjenspeiler kostnadsulemper som skyldes ineffektivitet framstår det som hensiktsmessig at indeksverdien faller når ineffektiviteten øker. Kostnadsforskjeller mellom kommuner kan også tilskrives forskjeller i priser (f.eks. lønninger eller eiendomspriser) og kontekstuelle faktorer, slik vi drøftet i kapittel 2.1.6. I dette tilfellet virker det hensiktsmessig å korrigere for kontekstuelle forhold, slik at ikke kommuner med naturgitte ulemper (eks., på grunn av bosetningsmønster) «straffes» for disse i produksjonsmålingen. Dette gjelder spesielt dersom produksjonsindeksen skal benyttes til måling av effektivitet.

Resultatet av en slik maksimering er effektiv utnyttelse av ressurser og en optimal prioritering. En empirisk anvendelse av Førsums (2017) modell finnes i Hanson (2019).

Konstante enhetskostnader

Som tidligere beskrevet benytter Borge mfl. (2001) aggregeringsvekter som baserer seg på at produkter skal verdsettes med sine gjennomsnittlige produksjonskostnader. Metoden ser følgelig bort fra at produsert mengde påvirker marginale kostnader og dermed også skalafordeler/ulempen i produksjonen. På tilsvarende måte antas det et konstant bytteforhold mellom ulike varer og tjenester, noe som er uvanlig i økonomisk produksjonsteori: Typisk antas det at substitusjonskurven mellom to outputs er krummet, slik at det blir mer kostbart å utvide produksjonen av et gode man allerede produserer mye av.

Konstante bytteforhold og skalaavkastning står i kontrast til DEA-analysen (jf. kapittel 2.1), hvor modellen selv definerer variable bytteforhold og skalafordeler/ulempen. I fortsettelsen ønsker vi å vise hvordan en alternativ produksjonsindeks kan utledes basert på DEA.

2.3.2 Effektivitetsmåling basert på produksjonsindeksen

I Borge mfl. (2011) deles produksjonsindeksen på kommunens kostnader for å lage et mål på effektivitet. Det bør bemerkes at dette ikke er et mål på *effektivitet*, men på *produktivitet*. Årsaken er at det ikke gjøres noen sammenlikning mot en referansenorm for best praksis eller teknologi⁸. Det er dessuten kun et forholdstall som ikke tar hensyn til skalafordeler eller ulemper. Borge mfl. (2011) sitt opplegg mangler en klar definisjon av beste praksis, og dermed også av begrepet effektivitet. Vi foreslår i fortsettelsen en alternativ produksjonsindeks basert på konvensjonell metodikk innen produktivitets- og effektivitets-analyse.

2.3.3 Et alternativt opplegg til måling av effektivitet for produksjonsindeksen

Vi har pekt på flere egenskaper som vi vurderer som svakheter ved det etablerte opplegget for måling av effektivitet basert på produksjonsindeksen (*EFFPI*):

- 1) Opplegget mangler en klar definisjon av beste praksis og følgelig også av effektivitet
- 2) Opplegget baserer seg på konstante enhetskostnader og ser dermed bort fra substitusjonsrater mellom produkter og skalaavkastning i produksjonen
- 3) Kommunenes prioriteringer har ingen betydning for indeksverdien. Dette henger også sammen med punkt 2), siden substitusjonsrater antas å være konstante
- 4) Det gjøres ingen korrigerende av effekten av kontekstuelle forhold med hensyn til produksjons- og effektivitetsmåling.

I dette kapitlet foreslår vi alternativ metodikk som beholder fokuset på produksjonsindeksen, samtidig som den lar oss imøtekomme de 4 kritiske merknadene over.

Som i Borge mfl. (2011) definerer vi vårt effektivitetsmål som forholdet mellom vektet produksjon og kommunebudsjettet (kalt C). Men i motsetning til Borge velger vi ikke vekter for produktene selv, men finner disse som del av et optimeringsproblem. Konkret ønsker vi å identifisere de vektene som gjør at kommunen vi vurderer får høyest mulig produktivitet (dvs. forhold mellom vektet produksjon og kostnader), under forutsetningen at produktiviteten skal være teknisk mulig. Som i det vanlige DEA-problemet betyr det at beste praksis bestemmes på bakgrunn av data fra kommunene. Optimaliseringsproblemet som ved sin løsning identifiserer maksimal produktivitet for kommune k er:

⁸ Merk at en slik indeks vil være den inverse av «kostnadsindeksen» som rapporteres for helseforetakene i SAMDATA-rapportene. Der beregnes kostnadsindeksen som kostnader pr DRG-poeng (Helsedirektoratet, 2021).

$$EFFPI = \max \left\{ \frac{\sum_{m=1}^M r_m y_m^{k'}}{C^{k'}} : \frac{\sum_{m=1}^M r_m y_m^k}{C^k} \leq 1, \forall k, r_m \geq 0, \forall m \right\}. \quad (25)$$

Effektivitetsmålet vårt er faktisk en DEA-modell. Det ser vi ved at problemet kan omskrives som følger:

$$\begin{aligned} EFFPI &= \max \left\{ \frac{\sum_{m=1}^M p_m y_m^{k'}}{wC^{k'}} : \frac{\sum_{m=1}^M p_m y_m^k}{wC^k} \leq 1, \forall k, p_m \geq 0, \forall m, w \geq 0 \right\} \\ &= \max \left\{ \begin{array}{l} \sum_{m=1}^M p_m y_m^{k'} : \sum_{m=1}^M p_m y_m^k - wC^k \leq 0, \forall k \\ wC^{k'} = 1 \\ p_m \geq 0, \forall m, w \geq 0 \end{array} \right\}. \end{aligned} \quad (26)$$

DEA er et lineært programmeringsproblem (LP-problem). Det er velkjent at et LP-problem har en såkalt dual – dvs. en alternativ måte å formulere problemet på – som gir den eksakt samme løsningen som det opprinnelige problemet (hvis det er noen entydig løsning). I vårt tilfelle er dualen til likning (26):

$$\begin{aligned} EFFPI &= \min_{\theta, \lambda} \left\{ \theta : \begin{array}{l} \sum_{k=1}^K \lambda^k y_m^k \geq y_m^{k'}, \forall m \\ \sum_{k=1}^K \lambda^k C^k \leq \theta C^{k'} \\ \lambda \geq 0 \end{array} \right\} \end{aligned} \quad (27)$$

Ved å sammenlikne med likning (1) ser vi at effektivitetsmålet for produksjonsindeksen er identisk med ressursbesparende DEA under forutsetning om konstant skalaavkastning.

Når vi har etablert sammenhengen mellom vår effektivitetsindeks og den vanlige DEA-modellen vil vi naturligvis være i stand til å løse alle problemene vi har pekt på for den etablerte produksjonsindeksen. DEA er en velegnet metode til effektivitetsmåling og har en klar definisjon av beste praksis; DEA-modellen tar alltid inn over seg substitusjonsrater og skalaavkastning i produksjonen; en utvidet DEA-analyse vil tillate å korrigere for kontekstuelle faktorer og deres påvirkning på kommuneeffektivitet. Vi kan også estimere DEA-modellen under skalaavkastningene CRS og VRS for å identifisere påvirkningen av teknisk og skalaeffektivitet mht produksjonsindeksen.

I kapittel 2.1.7 beskrev vi at det finnes to ulike tilnærminger til aggregering av effektivitetsindekser: Strukturell effektivitet og industrieffektivitet. I det første tilfellet antas det at ressursene ligger fast per sektor. Vi viste da at budsjettandeler er hensiktsmessige vektorer når sektoreffektivitetene skal aggregeres til kommunenivået. I industrieffektivitet antas det at ressurser kan allokere mellom sektorer for å optimere samlet kommune-produktivitet. Vår alternative produksjonsindeks tilhører denne definisjonen. Hva som er «riktig» mål på kommuneeffektivitet – om det er strukturell eller industrieffektivitet – blir en subjektiv vurdering mer en at det finnes et objektivt svar. I stedet gir vårt opplegg muligheter til å beregne begge målene på kommuneeffektivitet for å kunne sammenlikne dem.

Til slutt er det betimelig å bemerke at produksjonsindeksen er i dag satt sammen av variable med vidt forskjellig målestokk. Det er ordinale data (f.eks. skolemiljø), korrigerede negative data (f.eks. skolebidrag), og forholdstall. Vi har tidligere drøftet at standard DEA ikke er tilpasset denne typen data, men at det kreves justeringer av DEA-modellen. Dette kan være et motargument mot bruk av DEA til måling av produksjonsindeksen, men samtidig framstår det som vanskelig å velge vektorer for variable med så ulike målestokk også i dagens produksjonsindeks. Det virker hensiktsmessig å ta en grundig gjennomgang av variable i produksjonsindeksen og i hvilken grad dagens opplegg er vellykket i å konvertere de til en felles målestokk. En anbefaling, i alle fall i forbindelse med DEA, vil være å i stor grad benytte tradisjonelle volummål i definisjonen av frontmodellen, og heller å korrigere for mer komplekse variable i en ettermodell.

2.4 Benchmarking

KOMMUNE-STAT-RAPPORTERING (KOSTRA) oppgir en rekke nøkkelindikatorer per tjenesteområde (eks. netto driftsutgifter til barnehager, per innbygger) som gjør at man kan sammenlikne kommuner på utvalgte måltall. Ulempen med dette, sammenliknet med SØFs metodikk, er at sammenlikningen vil måtte bestå av flere nøkkeltall eller partielle effektivitetsindikatorer. Det er dermed et spørsmål om hvilken vekt vi skal tillegge den enkelte indikator. Den parvise sammenlikningen gir dessuten ikke en entydig sektorspesifikk effektivitetsscore mellom 0 og 1 (hvor 1 indikerer beste praksis). Metoden tar heller ikke inn over seg avveininger i ressursbruk (f.eks. avveininger mellom ulike produkter i tjenesteproduksjonen), noe SØFs metodikk er i stand til⁹. Benchmarking basert på partielle indikatorer kan heller ikke ta hensyn til utnyttelse av skalaøkonomi og samdriftsfordeler eller ulemper, noe DEA-metoden er velegnet til.

Et sentralt valg i sammenstillingen av slike partielle produktivitetsindikatorer er hvordan de skal vektet sammen. Valg av vektorer drøftes av Kuosmanen og Kortelainen (2005), som foreslår et opplegg som er helt parallelt til produksjonsindeksen vi definerte i likning (25). Dette viser at man kan benytte DEA til vekte sammen utvalgte måletall (f.eks. tjenesteproduksjon per innbygger), og man kan følgelig benytte opplegget til å beregne en effektivitetsscore per kommune. Vi anser dette som et relevant opplegg for bruk av utvalgte nøkkelindikatorer og/eller videreutvikling av produksjonsindeksen til effektivitetsanalyseformål. Ved å bruke DEA-modellen til vektning av partielle indikatorer vil både avveininger i ressursbruk og skalaeffektivitet tas hensyn til av modellen.

2.5 Alternative metoder

Så langt har drøftingen av metode basert seg på DEA. Det er viktig å bemerke at dette bare er en av flere tilgjengelige metoder for effektivitetsanalyse, og at en viktig ulempe med DEA-metoden er at den ikke tillater stokastisk variasjon i data. Følgelig blir den sårbar for målefeil og uteliggere. Vi vil i dette delkapitlet presentere to andre metoder som er alternativer til DEA, og som begge tar hensyn til stokastisk variasjon i data. I begge tilfellene kan modellene inkludere multiple innsatsfaktorer og produkter, men for å gjøre beskrivelsen så enkel som mulig legger vi til grunn en modell med kun én innsatsfaktor. Når dette er brutto driftsutgifter kaller vi modellen en kostnadsfunksjon, som vi kan spesifisere på formen

⁹ DEA-metoden tar utgangspunkt i et produksjonsmulighetssett, og forteller hvor mye som kan produseres av ulike kvanta av produkter og tjenester for en gitt ressursbeholdning.

$$\ln x^k = \ln f(\mathbf{y}^k) + g(\mathbf{Z}^k) + \varepsilon^k \quad (28)$$

hvor x er brutto driftsutgifter.

Den viktigste forskjellen mellom de to alternative metodene SFA og StoNED er hvordan kostnads-funksjonen $f(\cdot)$ modelleres. Det vil si om den beskrives som en *parameterisk* eller *ikke-parameterisk* funksjon. I begge tilfeller antas det at restleddet består av to komponenter: Effektivitet (u) og stokastisk variasjon (v). Mao:

$$\varepsilon^k = u^k + v^k \quad (29)$$

I begge tilfeller estimeres modellen med et sammensatt restledd først, før dette restleddet dekomponeres videre inn i effektivitet og stokastisk variasjon.

2.5.1 Stokastisk frontanalyse

SFA-modellen behandler $f(\cdot)$ som en parametrisk funksjon, noe som innebærer at analytikeren må velge hva slags funksjonsform modellen skal ha. Dette er en ulempe siden det ikke er sikkert at den valgte funksjonsformen passer dataene godt. Man kan få mer fleksible modeller gjennom å velge såkalt fleksible funksjonsformer (som *Translog*-funksjonen; se Christensen mfl., 1973), men disse gir en ulempe gjennom at det må estimeres en lang rekke parametere basert på høyt korrelerte variable. Dette reduserer antallet frihetsgrader i modellen.

En enklere og svært populær funksjonsform er *Cobb-Douglas*-funksjonen. Denne gir modellen:

$$\ln x^k = \alpha + \sum_{m=1}^M \delta_m \ln y_m^k + \sum_{h=1}^H \beta_h Z_h^k + \varepsilon^k \quad (30)$$

SFA kan enkelt estimeres i statistikkprogrammer som for eksempel Stata eller R. I våre analyser benytter vi kommandoen `frontier`, som er en standardpakke i Stata. Vi legger til grunn av effektivitetsleddet u følger en halvnormal fordeling.

2.5.2 Stochastic non-parametric envelopment of data

I StoNED byttes $f(\cdot)$ ut med en ikke-parameterisk funksjon. I dag finnes det ingen standardverktøy for estimering av denne modellen, noe som innebærer at optimeringen må kodes i programmer som GAMS, MatLab eller Python. Derimot kan estimeringen av effektivitetsleddet skje i f.eks. Stata ved å bruke samme rutiner som SFA-modellen. Vår løsningsalgoritme er som følger:

1. Estimer $f(\cdot)$ ved å løse et optimeringsproblem i GAMS.
2. Kvasi-likelihood: Finn $\ln x^k - \ln \hat{f}(\mathbf{y}^k) - \hat{\boldsymbol{\beta}}' \mathbf{Z}^k$ fra steg 1. Benytt denne som avhengig variabel i pakken `frontier` i Stata. Prediker effektivitet.

3. Finn $\ln x^k - \ln \hat{f}(\mathbf{y}^k)$ fra Steg 1. Kjører en vanlig regresjonsanalyse med $\ln x^k - \ln \hat{f}(\mathbf{y}^k)$ som avhengig variabel og Z-variablene som uavhengige variable. Finn parameterestimater og standardfeil for kontekstuelle variable.

Det første steget i StoNED løses ved programmet

$$\begin{aligned} & \min \sum_{k=1}^K (\varepsilon^k)^2 \\ & \text{s.t.} \\ & \ln x^k = \ln \phi^k + \sum_{h=1}^H \beta_h Z_h^k + \varepsilon^k, \forall k \\ & \phi^k = \alpha^k + \sum_{m=1}^M \delta_m^k y_m^k, \forall k \\ & \phi^k \geq \alpha^{k'} + \sum_{m=1}^M \delta_m^{k'} y_m^k, \forall k, k' \\ & \delta^k \geq 0, \forall k \end{aligned} \tag{31}$$

Her angir vektene i kostnadsfunksjonen (δ) marginale kostnader, dvs. at de har samme tolkning som såkalte «skyggepriser» i DEA: De angir hellingen på fronten. Optimeringsproblemet i (31) pålegger at skyggeprisene er ikke-negative, noe som gjør at kostnadsfunksjonen oppfyller egenskapen *monotonitet*. Det andre settet av ulikheter pålegger kurvatur-egenskaper, som i vårt tilfelle betyr at kostnadsfunksjonen er *konveks*.

StoNED er i praksis en DEA-modell som er tilpasset gjennomsnittet av dataene heller enn til de mest ekstreme datapunktene. Følgelig blir standard DEA bare et spesialtilfelle av StoNED-modellen: Vi kan gjenskape DEA-modellen ved å pålegge at restleddene ε skal være ikke-negative for alle enheter i (31). I dette tilfellet blir $\varepsilon = u$, og modellen antar dermed at det ikke er noen stokastisk variasjon (dvs. $v=0$) i dataene.

2.5.3 Paneldatanalyse

Så langt har vi – i likhet med SØF – fokusert på metodikk for estimering av effektivitet basert på tverrsnittsdata. SFA og StoNED kan derimot utvides til å ta hensyn til at kommunedata har en paneldatstruktur, dvs. at den samme kommunen observeres gjentatte ganger over tid.

En viktig fordel med paneldata over tverrsnittsdata er at det er mulig å identifisere uobserverbar heterogenitet; dvs. kjennetegn ved kommunene som påvirker deres produksjon men som ikke enkelt kan kvantifiseres ved hjelp av relevante variable i modellspesifikasjonen. Dette er blitt utnyttet i litteraturen om effektivitetsmåling, hvor det er blitt utviklet modeller som skiller mellom kommuneheterogenitet (som ikke er en del av effektivitetsleddet, men som heller tilskrives variasjon i uforklarte kontekstuelle variable) og effektivitet. I senere tid har litteraturen diskutert et restledd med 4 komponenter; kommuneheterogenitet, persistent effektivitet, tidsvarierende effektivitet og stokastisk variasjon i data. Vi viser til Kumhakar med fl. (2014) for en pedagogisk oversikt over aktuelle paneldatamodeller og hvordan de enkelt kan estimeres ved bruk av flere steg for å dekomponere restleddet.

3 Data

I dette kapittelet vil vi ta for oss datagrunnlaget for effektivitetsanalysene av kommunesektoren, både i form at de dataene som brukes i dag og de som kan bli benyttet i fremtiden. Vi vil starte med å foreta noen overordnede betraktninger, før vi går nærmere inn på de ulike variabelgruppene og variabler i hver enkelte delsektor.

3.1 Overordnede betraktninger

Ideelt sett skal data i en effektivitetsanalyse være komplette, det vil si dekke all ressursbruk og all tjenesteproduksjon for en klart avgrenset sektor. Det må heller ikke være dobbelttelling som gjør at samme informasjon opptrer flere ganger. Variable bør dessuten måle de viktige aspektene uten målefeil. Hvilke data som praktisk sett bør inngå i effektivitetsanalyser av kommunal sektor er likevel gjestand for en rekke kompromisser. I det følgende vil vi gå i gjennom noen av de viktigste av dem. Vi starter med å ta for oss hensiktsmessig utformingen av dataene og generiske problemstillinger knyttet til måling av variablene. Deretter går vi nærmere inn på implikasjonene av hvilke nivå variablene kan måles på og avgrensingen av delsektorene.

3.1.1 Hensiktsmessig utforming av dataene

For det første er det gjerne et spenn mellom dataene man har, som kan framskaffes, som muligens kan framskaffes og som man ideelt sett burde ha hatt. I praksis har man gjerne begrensede ressurser til effektivitetsanalysene og alternativ ressursbruk. Man må dermed foreta noen valg med hensyn til hvor mye vekt man skal legge på å få best mulig data. Like fullt er det ikke tilstrekkelig med solid metodikk for å utføre gode effektivitetsanalyser, hvis dataene ikke er av tilstrekkelig kvalitet.

Datasettet for effektivitetsanalyser av kommunal sektor bør bestå av entydig avgrensede enheter for tjenesteproduksjon. Man bør helst ha lange tidsserier og mange observasjoner. I hver av SØFs effektivitetsberegninger benyttes de fleste av Norges kommuner og de to siste årene, slik at man vil ha relativt mange observasjoner å forholde seg til. I effektivitetsanalyser over flere år kan kommunesammenslåinger eller kommuneoppsplitting håndteres ved at man behandler de nye og gamle kommunene som ulike kommuner, der hver av dem bli stående uten observasjoner i år der en annen kommunestruktur gjelder.

For å unngå feilestimering knyttet til utelatte variabler bør listen av variabler for innsatsfaktorer, produkter og kontekstuelle variabler være mest mulig komplett. Samtidig kan for mange innsatsfaktorer og produkter lede til «dimensjonalitetens forbannelse», der beregningene blir for komplekse til at man får konvergens eller til at man har datakraft til å gjennomføre dem. For mange variabler i modellspesifikasjonen kan også føre til utfordringer med å ha tilstrekkelig mange frihetsgrader, skjønt problemet er mindre i datasett med mange kommuner og eventuelt flere år. Vanlig praksis er derfor å velge ut variabler som har størst økonomisk betydning.

Beregningene av effektivitet for kommunal sektor innebærer såpass mange observasjonsheter og potensielt flere år, så utfordringer med identifikasjon som dimensjonalitet og

antall frihetsgrader er neppe noe problem. Dimensjonalitet kan imidlertid bli en utfordring ved både mange produkter og innsatsfaktorer. Dette gjelder eksempelvis i tilfellet produksjonsindeksen.

I SØFs effektivitetsanalyse korrigerer det ikke for kommunefaste effekter og eventuelt tidsfaste effekter, som kunne ha fange opp forhold som er persistente over tid eller rom. Dette kan løses ved å benytte en paneldataspesifikasjon (jmfør seksjon 2.5.3). Alternativt bør man stille strenge krav til komplett spesifikasjon av kontekstuelle faktorer. Listen av kontekstuelle variabler bør inkludere alle variabler som påvirker produksjonsprosessen og resultatet.

3.1.2 Krav til variabler

Det er ikke tilstrekkelig å inkludere alle de relevante variablene – man må også måle dem riktig. Å klargjøre datasett for frontestimering innebærer typisk mye manuelt arbeid med tilhørende utfordringer og muligheter.

Produktene og faktorinnsats som inngår estimeringsmodellen bør dekke produksjonsaktivitetene som effektivitetsscoren måler og samtidig ikke vedrøre annen produksjon direkte. Med andre ord bør det være et samsvar mellom produksjonen og faktorinnsatsen i estimeringsmodellen. I praksis kan naturligvis dette være utfordrende i enkelte tilfeller, for eksempel når man skille mellom sentraladministrasjon og sektoradministrasjon. Når det gjelder kontekstuelle variabler, forutsetter mange empiriske spesifikasjoner at de har tilsvarende innvirkning på ulike beslutningsenheter. Dette kravet kan brytes om de kontekstuelle variablene har ulik betydning i ulike sammenhenger, eller egentlig er heterogen samling av forskjellige delkomponenter med ulik innvirkning på produksjonen.

Heterogenitet i faktorinnsatsen bør fanges opp av målene som benyttes, hvis man ikke risikerer man at kommuner dominert av produksjonsressuser med relativt lave faktorkostnader straffes unødige i effektivitetsmålingen kontra kommuner med dominert av produksjonsressuser med relativt høye faktorkostnader. Et velkjent eksempelet er når høykompetent arbeidskraft med høyere utdanning og eventuelle god helse, relevant erfaring og talent sidestilles med lavkompetent arbeidskraft med et lavere lønnsnivå. Det gjelder imidlertid også realkapitalinnsatsen, der immobile kapitalformer som bygg, anlegg og tomter har lengre levetid og dermed flere år å tjene inn kapitalverdien sin på enn mobile kapitalformer som immaterial kapital, maskiner, utstyr og transportmidler.

I praksis klarer man å ta høyde for deler av heterogeniteten, men ikke alt. For eksempel har man ofte tilgang til data for arbeidskraftens utdanningsprofil og i neste omgang kanskje også stillingskategorier og utdanningsprofil, mens data på helsetilstand, erfaringsprofil og talent er mindre tilgjengelige. På kapital siden kan man kanskje klare å skille mellom eiendoms kapital og annen kapital, mens videre disaggregering sitter noe lengre inne.

Å ta høyde for kvalitetsforskjeller står også sentralt på produksjonssiden. I dag tar SØF høyde for kvalitet i målingen av effektivitet i enkelte delsektorer, men ikke i andre. Det blir da viktig å utarbeide gode resultat- og kvalitetsmål. Kvalitetsforskjeller kan også spille inn på relevante kontekstuelle faktorer.

Produktene og innsatsfaktorene som er oppgitt som verdier bør deflateres om til fastpristall, som skal gjenspeile volumer heller enn prisendringer. I måling av effektivitet i kommunal sektor i Norge er dette relevant for innsatsfaktorene, som typisk måles som en korrigert utgave av bruttodriftsutgiftene. Som regel bør også kontekstuelle variabler deflateres, jo mindre det er en tolkningsmessig grunn til at man vurderer verdier til å være mer relevant enn volumer. For å unngå systematiske målefeil knyttet til sammensetningen og heterogen prisutvikling knyttet til variablenes delkomponenter bør man deflatere hver

komponent separat. Eksempler på relevante komponenter inkluderer høytudannet arbeidskraft, lavt utdannet arbeidskraft, eiendomskapital, mobil kapital, varer og tjenester.

Prisindekser for sektorielle faktorpriser produseres i dag av Statistisk sentralbyrå og benyttes av Nasjonalregnskapet. Her har man allerede sektorspesifikke deflatorer for ulike kapitalformer og produktinnsatsen. Statistisk sentralbyrå produserer også lønnsindekser som er relevant for arbeidskraften, skjønt skillet mellom høytudannet og lavtutdannet arbeidskraft er noe mer utfordrende å håndtere. Tilsvarende er oppsplittingen av prisindeksene for produktinnsatsen til egne indekser for vareinnsatsen og tjenesteinnsatsen mer komplisert. Samtidig utarbeider Statistisk sentralbyrå deflatorer for bruttoproduksjonen og produktinnsatsen, så vel som oversikter over næringenes kryssløpstabeller, så en oppsplitting vil være mulig.

I SØFs liste over modellvariabler har det blitt foretatt enkelte ukonvensjonelle grep. Det er ikke vanlig å foreta skjønsmessige justering i innsatsfaktorene eller produktene. Innenfor konvensjonell effektivitetsestimering behandler man isteden kontekstuelle variabler separat i modellspesifikasjonen, slik at estimeringsmodellen selv kan bestemme de kontekstuelle variabelenes innvirkning. Videre behandles rene volumkomponenter og kvalitetskomponenter hver for seg, slik at estimeringsmodellen kan avdekke potensiell substituerbarhet eller komplementaritet i produksjonen.

Man kan også stille spørsmåltegn ved at SØF har multiplisert rene volumkomponenter med tilhørende kvalitetskomponenter, uten å inkludere de rene volumkomponentene. Å multiplisere med volumkomponentene er nødvendig for at kvalitetsvariabelene skal tilfredsstille skalaegenskapene til DEA, men når man samtidig ekskluderer de rene volumkomponentene ignorerer man eventuelle krysseffekter mellom kvantitet og kvalitet. En annen mulighet ville ha vært å modellere kvalitetsindikatoren som en kontekstuell variabel, jmfør utfyllende diskusjoner i seksjon 2.1

Multikolaritet mellom variablene kan være en utfordring når koeffisientene til regresjonens forklaringsvariabler og kontekstuelle forhold skal tolkes. Til gjengjeld kan korrelasjonen mellom disse variablene og eventuelle sentrale utelatte variabler bidra til at feilkilden forbundet med utelatte variabler reduseres. For eksempel får utelatelse av data om kapitalinnsatsen mindre betydning samlet sett om de er høyt korrelert med arbeidskraften, som er modellert som en egen innsatsfaktor.

3.1.3 Dataenes dimensjoner

I SØFs effektivitetsberegninger er «år» tidsvariabelen og «kommuner» observasjonseenhetene. For hvert år rapporterer SØF kun effektivitetsscoren for det siste året, men denne estimeres på bakgrunn av de to siste årene. Det forutgående året er inkludert for å kunne betrakte effektivitet over tid, der felles teknologifront er en forutsetning som legges til grunn. En slik fremgangsmåte med endret varibeloppsett over tid er ukonvensjonell, jmfør diskusjoner effektivitetsmåling over tid i delkapittel 2.2 og om paneldata-analyse i seksjon 2.5.3. Fra et myndighetsståsted kan det også oppfattes uheldig at beregninger for tidligere år gjenstas årlig med fornyet datasett, da man kan stå overfor et forklaringsproblem om rangeringene av kommunene endres.

Når det gjelder valg av observasjonseenheter, synes kommuner fornuftig ut, ettersom det er kommunal effektivitet man er opptatt av. Som et alternativ kunne likevel vært interessant å måle effektivitet på institusjonsnivå når det finnes data på dette nivået, og deretter aggregert det opp til kommunalt nivå. Man kunne da fått bedre fram heterogeniteten i tjenesteleveransen innad i kommunene, selv om eventuelle sammensetningseffekter blir neglisjerte. Estimering på institusjonsnivå ville også krevd en form for vekting, for eksempel innsatsfaktorer eller produksjonen. Dersom variablene bak disse vektene endrer seg over tid,

måtte man i en paneldatasammenheng videre tatt stilling til om vil operere med tidsinvariante eller tidsvarierende vekter.

3.1.4 Nivåene på dataene

Variablene som inngår i effektivitetsanalyser rapporteres på ulike aggregeringsnivå, herunder kommunenivå, institusjonsnivå og individnivå. I tillegg rapporteres en del data, som man gjerne skulle hatt tilgang til på kommunenivå, istedenfor på fylkesnivå, skjønt de i prinsippet kan bestilles på et mer disaggregert nivå. Dette gjelder data både levert av Statistisk sentralbyrå og sektorielle tilbydere av statistikk. Et eksempel på dette er private tjenestetilbydere innenfor det kommunale andsvarsdomenet. Data på et mer disaggregert nivå er gjerne bedre egnet til å fange opp hetegogenitet knyttet til produksjonen og omgivelsene.

Å gjennomføre estimeringer på høyere aggregeringsnivå innebærer ikke nødvendigvis at det overordnede bildet endres betydelig, men estimeringsresultatene blir lidende av mer stokastisk støy enn nødvendig. Kontekstuelle faktorer på kommunenivået kan fortsatt bidra med noe forklaringskraft, men typisk vil mye av den forsvinne i aggregeringen. Data på lavere aggregeringsnivå enn kommuner kan fortsatt benyttes i estimering på kommunenivå ved at de aktuelle relasjonene estimeres i et separat trinn. Dette krever imidlertid mer avanserte estimeringsprosedyrer, i hvert fall om estimeringsprosedyrene skal estimeres simultant, hvilket er en forutsetning for ikke å tape estimeringseffektivitet. SØF foretar noen korreksjoner i dataene på mer disaggregert nivå enn kommune, men det gjøres ikke gjennom simultan estimering (jf. kapittel 2.1.3). Dermed introduserer man en estimeringsfeil ved at den kontekstuelle regresjonen ikke tar hensyn til effektivitetsregresjonen. Det vil da ikke være gitt at man vinner mer på å foreta estimeringer på et mer disaggregert nivå enn det man taper på ikke å foreta simultan estimering. Estimering i to trinn innebærer at regresjonskoeffisientenes standardfeil ikke er rett fram å tolke.

Like fullt innbefatter estimering på mer disaggregerte data enn kommunenivå potensiell forbedringer i effektivitetsberegningene. Generelt vil heterogeniteten i sammenhengen mellom produkter, innsatsfaktorer og kontekstuelle variabler komme bedre fram når man tar hensyn til at det er snakk om en rekke institusjoner eller individer, og ikke bare fokuserer på gjennomsnittsbetraktninger eller betraktninger basert på et annet sentralitetsmål, eventuelt supplert med et spredningsmål.

Data på institusjonelt nivå vil i en del tilfeller være offentlig tilgjengelige, mens de i andre tilfeller krever at man søker om dataene eller har spesielle tilganger. Ved enkelte anledninger kan observasjoner både på institusjonsnivå og kommunenivå være prikket av statistikkleverandøren for å ivareta konfidensialitetshensyn. For individdataene er kravene til tilgang og oppbevaring gjennomgående enda strengere. På grunn av personvern hensyn pleier Statistisk sentralbyrå å fjerne observasjon enheter med færre enn tre personer, enten det er snakk om befolkning, sysselsetting eller noe annet. I noen tilfeller fjerner de også andre observasjoner, for at observasjon enhetene med færre enn tre personer heller ikke skal være direkte identifiserbare, skjønt man i prinsippet kan få ta i disse dataene for forsknings- og spesielle utredningsformål.

Det danske motstykke til Kommunal- og moderniseringsdepartementet og Teknisk beregningsutvalgs effektivitetsanalyser for kommunal sektor er Indenrigs- og Boligministeriets Benchmarkingenhed, som utgjør en selvstendig og uavhengig etat under Indenrigs- og Boligministeriet. Indenrigs- og Boligministeriets Benchmarkingenhed fokuserer riktignok ikke på fronteffektivitetsanalyser, men foretar enklere komparative studier av kommuner og administrative regioners oppgaveløsning med henblikk til å synliggjøre forbedringspotensial og beste praksis. Danskene ligger imidlertid relativt langt framme når det gjelder

arbeidet for tilgjengeliggjøring av mikrodata i denne typen sammenlikninger, både gjennom praktisk og juridisk tilrettelegging. Indenrigs- og Boligministeriets Benchmarkingenheds sammenlikningsstudier bygger blant annet på innsamlede personopplysninger, som behandles konfidensielt med datasikring.

I Norge er derimot tilgangen til mikrodata ikke alltid rett fram. Noe stengere lovgiving og håndhevelse av lovgivingen over tid med henblikk på ivareta personvern og andre konfidensialitetshensyn har bidratt til at tilgangen til mikrodata generelt har blitt vanskeligere. Statistisk sentralbyrå anonymiserer ofte data ment for forskning, ikke bare på individnivå, men også på virksomhetsnivå. I de senere år har også enkelte statistikker blitt desentralisert til andre etater med generelt mer områdekompetanse, men også generelt mindre datakompetanse, hvilket har bidratt i samme retning.

På en annen side har fokuset innenfor internasjonal forskning rettet seg stadig mer mot potensialet ved utnyttelse av mikrodata i de senere år. Dette utviklingstrekk har naturligvis også preget norsk forskning, som har et relativt sterkt forskningsmiljø og relativt god statistikk. Statistisk sentralbyrås opprettelse av en egen portal for tilgang til mikrodata har bidratt til noe enklere tilgang til mikrodata ved at det ikke er behov for å søke etter alle data fra prosjekt til prosjekt, skjønt langt fra alle data er tilgjengelig gjennom denne portalen.

3.1.5 Sektoravgrensing

For å oppnå treffsikker og pålitelig effektivitetsmåling for kommunal sektor er det avgjørende å avgrense analysen sektorielt på en god måte. En god avgrensning innebærer at grenselinjene mellom hver delsektor trekkes opp, og at man velger ut hvilke delsektorer man vil foreta effektivitetsmålinger for.

Avgrensingen av hver enkelt delsektor bør foretas på en måte som sikrer at de observasjonene over kommuner og tid er mest mulig sammenliknbare. utfordringer i denne forbindelse kan være tilstedeværelsen av konkurrerende private tilbydere, håndteringen av biaktiviteter i produksjonen, indirekte kostnader og annen uobserverbar heterogenitet i variablene i estimeringsmodellen. Alle kommuner løser lovpålagte oppgaver, men mange av dem kan løses på ulike måter. Desuten kommer valgfri oppgaver inn. Straks disse forholdene varierer over kommuner uten at det tas hensyn til effektivitetsmodellen, vil systemiske målefeil oppstå i estimeringsmålingen. Sekundært bør man forsøke å måle representativ aktivitet innenfor de utvalgte delsektorene, hvilket typisk innebærer at man bør fokusere på de største aktivitetene målt i produksjon og faktorinnsats. I tillegg spiller tilstedeværelsen av gode data inn, inkludert volummål og kvalitetsmål for produksjonen, og mål for faktorinnsatsen og kontekstuelle variabler med sektoriell innflytelse.

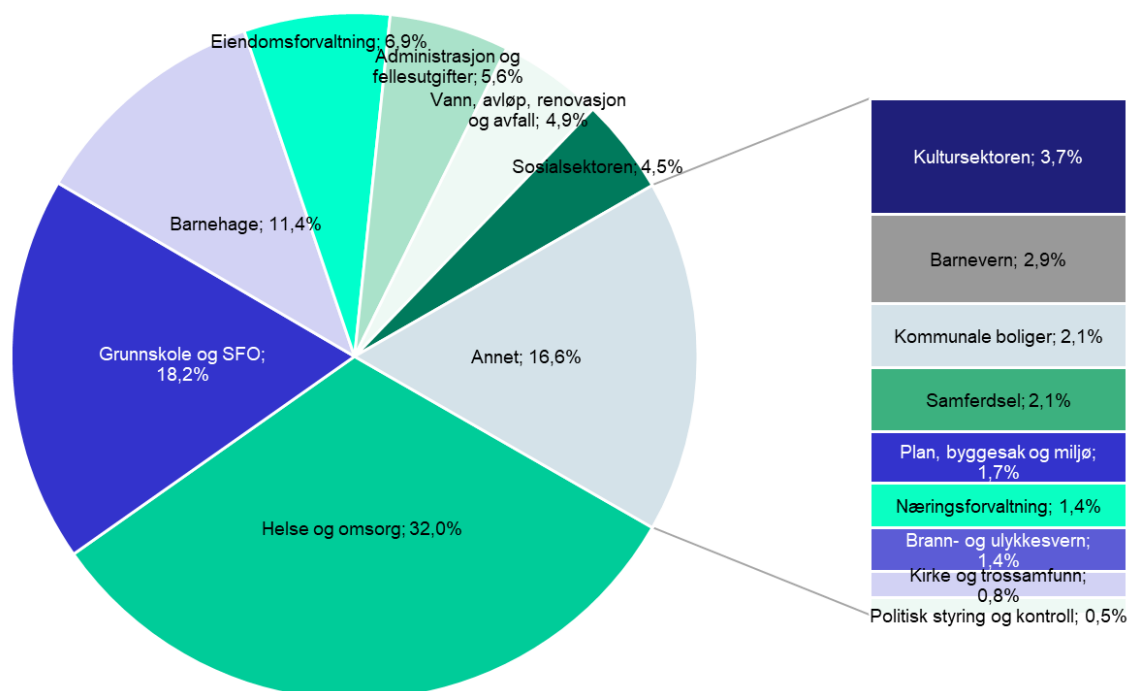
Hvilke kommunale delsektorer som bør løftes fram i den kommunale effektivitetsmålingen velges ut i fra de samme prinsippene, dog med litt annen vektning. Her er det viktig å få dekket en stor og gjerne representativ del av kommunal sektor. Dette innebærer at de største delsektorene bør tas med i effektivitetsanalysen, jo mindre avgrensings- og målingsproblematikk tilsier at effektivitetsmålingen blir for unøyaktige eller upålitelige. Igjen går dette på sammenliknbarhet og god datakvalitet. For eksempel er det ikke gitt at kommuner har felles praksis innenfor kommunale delsektorer med hensyn til organisering, ved at innslaget av kommunale foretak, kommunale aksjeselskap, interkommunale samarbeid og utkontraktering til private varierer. Slike forskjeller kan f.eks. behandles gjennom kategoriske variable i en etteranalyse. Nyttens og kostnadene av samme aktiviteter bør også tilordnes samme sektorer. Her er det blant annet en utfordring med kommunale fellesutgifter, som eventuelt kan håndteres som en kontekstuell variabel. Generelt vil heller ikke produksjonens skalaegenskaper fanges opp skikkelig når tilstedeværelsen av interkommunale samarbeid ignorereres. Legevekt, brannvesen, renovasjon og havnetjenester er eksempel på

kommunale tjenester som mindre kommuner ofte løser gjennom interkommunale samarbeid. I dagens effektivitetsmåling for KMD og TBU konsentrerer SØF seg om tre delsektorer; helse og omsorg, grunnskolen og barnehagen. Avgrensingen kan begrunnes i delsektorenes størrelse, samt avgrensingsproblematikk og datakvalitet, jmfør foregående diskusjon. I andre sammenhenger har SØF også undersøkt effektiviteten i andre kommunale delsektorer, inkludert barnehagen, kultursektoren og skolefritidsordningen. I tillegg fins det en del internasjonal forskning på effektiviteten innenfor rennovasjon, avfall, avløp og vannforsyning. Se for eksempel Hernández-Sancho og Sala-Garrido (2009), Rogge og De Jaeger (2012), Tupper og Resende (2004), Deng, Li og Song (2016) og Pérez-López med flere (2016).

Dersom vi tar for oss delsektorenes behandling i kommunale effektivitetsanalyser og deres størrelse, ser vi at det er en klar sammenheng mellom disse faktorene. Dette kommer tydelig fram i figur 3.1, der vi har illustrert fordelingen av bruttodriftsutgifter over hovedsektorer. Vi ser at delsektorene som er inkludert i effektivitetsanalysene for KMD og TBU – helse og omsorg, grunnskolen og barnehagen – også utgjør de største delsektorene innenfor kommunal sektor med vel halvparten av bruttodriftsutgiftene.

Merk at valg av bruttodriftsutgifter i størrelseskategoriseringen ikke tar hensyn til brukerbetalinger som kommer inn i sektorene. Slike brukerbetalinger kan være vesentlige i enkelte sektorer som SFO, eiendom, vann, avløp, rennovasjon og avfall og den kommunale kultursektoren. Størrelseskategoriseringen sier derfor enda mer om hvor viktig hver delsektor er for kommunale budsjetter enn om hvor store sektorene er samlet sett.

De største sektorene målt ved bruttodriftsutgifter – helse og omsorg, grunnskole og barnehage – alle er representert i effektivitetsmålingen for KMD og TBU. I hovedsak er de etterfølgende delsektorene målt i størrelse, også de delsektorene som det har blitt foretatt effektivitetsmåling på. Kommunal eiendomsforvaltning, administrasjon og fellesutgifter, og sosialsektoren er de største kommunale delsektorene som ikke har vært gjenstand for effektivitetsmålinger, mens kommunal kulturtjenester og barnevern er den minste kommunale delsektoren som har vært behandlet i effektivitetsstudier.

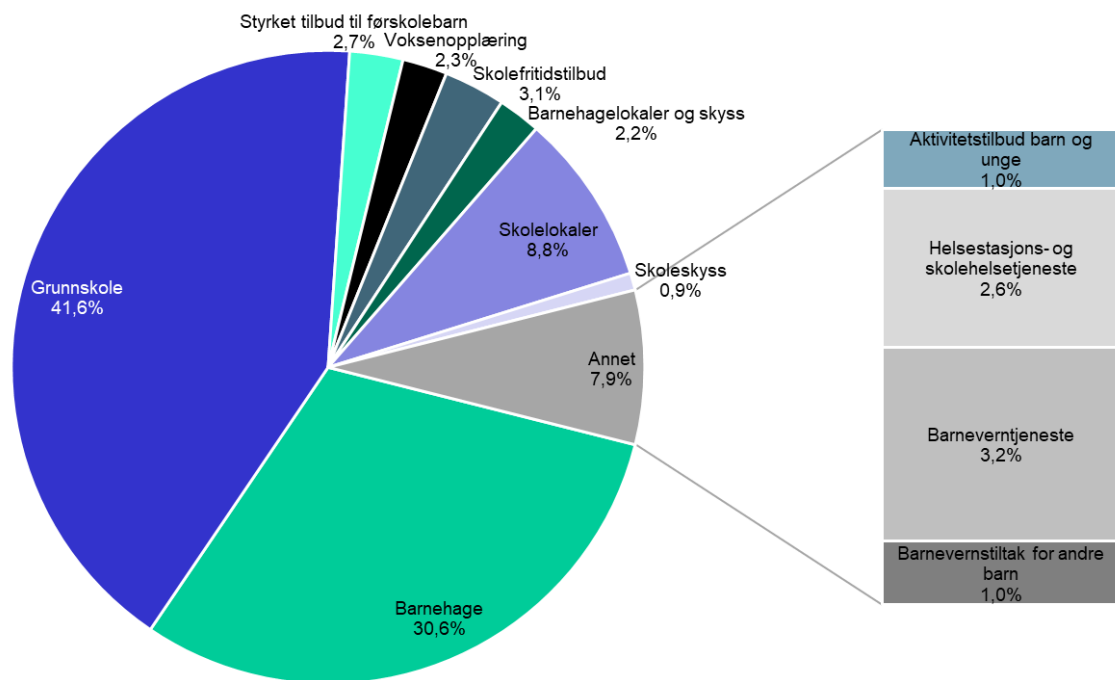


Figur 3.1: Fordeling av bruttodriftsutgifter over hovedsektorer i kommunal sektor. Kilde: Avledet basert på KOSTRA-data fra Statistisk sentralbyrå.

Denne formen for størrelsesrangeringer avhenger naturligvis hvordan sektorene er definerte og avgrensede. I noen tilfeller er til dels ulike aktiviteter gruppert i samme sektor. For eksempel omsorgsboliger og hjemmehjelpstjenester slått sammen til helse og omsorg. Videre er grunnskole og SFO slått sammen til en felles kategori. Det samme er vann, avløp, rennøvasjon og avfall. I andre tilfeller er ulike aktiviteter med en viss samtilhørighet gruppert innenfor ulike kategorier, hvilket spesielt gjelder den kommunale eiendomssektoren.

Til sammenlikningen behandles pleie på institusjon og hjemmesykepleien (som vi har vært inne på) som en delsektor i effektivitetsberegningene til SØF, mens grunnskole og SFO behandles som to forskjellige delsektorer. Det kan argumenteres for at de to pleietjenestene er to ulike måter å tilfredstille det samme behovet på, mens grunnskolen og SFO er to ulike, men komplementære, tjenester. Den nære komplementariteten og avhengighetsforholdet i produksjonen mellom grunnskolen og SFO er likevel et argument for at disse kommunale delsektorene kunne vært behandlet samlet.

Oppsplittingen av hovedsektoren grunnskole og SFO til underliggende sektorer kan ikke leses direkte fra Kostra, siden kostnadskomponenter fra andre delsektorer til dels innbefatter en dobbelttelling og også vedrører de underliggende delsektorene. I figur 3.2 har vi illustrert hvordan bruttodriftsutgifter i de delene av kommunal sektor som retter seg mot barn og opplæring fordeler seg over segmenter. Av denne framstillingen framstår det at grunnskole dominerer kategorien grunnskole og skolefritidsordningen i figur 3.2. Skole- og fritidsordningen utgjør en relativt lav andel av kategorien, men andelen ville vokst noe om man hadde spredt kostnadene for skolelokaler over kostnadssteder, i tillegg til samdriftsfordeler knyttet til eiendom og personal. Skole- og fritidsordningen innbefatter også vesentlige brukerbetaling, slik at sektoren ville vokst om om vi tok hensyn til dette.



Figur 3.2: Fordeling av bruttodriftsutgifter over segmenter i de delene av kommunal sektor som retter seg mot barn og opplæring. Kilde: Avledet basert på KOSTRA-data fra Statistiske sentralbyrå.

3.2 Valg av variabler

Utvalget av variabler har naturligvis stor betydning for hvilke resultater man ender opp med i effektivitetsanalyser. I tabell 2.1 i Borge med flere (2020) og tabell 9.1 i Teknisk beregningsutvalg (2020) listes produktene og innsatsfaktorene som inngår i SØFs effektivitetsanalyser. Variablene framstår i utgangspunktet som rimelige gitt datatilfanget. Likevel savnes det en begrunnelse for valget av variablene, og det er også enkelte rom for forbedringer.

I det følgende går vi nærmere inn på valg av variabler i effektivitetsanalysene for kommunal sektor. Vi starter med noen generiske betraktninger for hver variabeltype på tvers av sektorer, før vi gjennomgår relevante kontekstuelle variabler på tvers av sektor.

Deretter tar vi oss delsektor for delsektor innenfor kommunal sektor. Vi starter med de sektorene som framheves i de offentlige analysene for KMD, inkludert grunnskolen, pleie- og omsorg og barnehager. Deretter tar vi for oss andre kommunale delsektorer som SØF har foretatt analyser av, inkludert skole- og fritidsordningen, barnevernet og kulturelle tjenester.

I vår gjennomgang ser vi det hensiktsmessig å utvide sektorgjennomgangen. Blant delsektorene som ikke er et fokus i SØFs analyser har vi valgt å løfte fram vann, avløp, avfall og rennvasjon i en egen seksjon, idet vi i resten av kommunal sektor ser et særlig potensial for å foreta kommunale effektivitetsanalyser for denne delsektoren. Øvrige delsektorer er behandlet samlet i delkapittelets siste seksjon.

Innenfor rammene på dette prosjektet har det vært for omfattende å frambringe den ideelle variabellisten for hver delsektor innenfor kommunalsektorene eller å drøfte dataene benyttet i analysene av hver delsektor i detalj. Isteden har vi nøyd oss med å framheve noen hovedpoenger for hver sektor. Videre har vi listet opp hvilke variabler som blir og har blitt benyttet i effektivitetsanalyser av de ulike kommunale delsektorerene og pekt på noen alternativer og mulige forbedringer i de sektorielle variabellistene.

3.2.1 Generelle betraktninger om innsatsfaktorer

Som innsatsfaktor i SØFs siste effektivitetsberegninger for kommunal sektor benyttes korrigerede bruttodriftsutgifter. For det første fratrekkes konsernoverføringer, avskrivninger og arbeidsgiveravgift. For det andre deler SØF på sektor-spesifikke lokalt sentralitetsmål, som er ment å reflektere ulempene ved å tilby tjeneste i perifere strøk. Det sistnevnte forholdet vil vi diskutere nærmere i seksjon 3.2.3 om kontekstuelle faktorer.

Ideelt sett burde man benyttet bruttodriftskostnader (som reflekterer faktiske kostnader) og ikke bruttodriftsutgifter (som reflekterer utbetalinger). Brutto driftsutgifter viser de samlede driftsutgiftene inkludert kapitalst og korrigert for dobbeltføringer som fordelte utgifter. Det virker fornuftig å korrigere bort både konsernoverføringer, dobbeltføringer og arbeidsgiveravgift, da disse postene ikke strengt tatt vedrører effektiviteten i tjenesteleveransen direkte. Det virker derimot mer tvilsomt å korrigere bort avskrivningene. Selv om de kan være beheftet med målefeil, bør effektivitetsmålingen ta høyde for kapitalinnsatsen og tilhørende kapitalkostnader.

Det finnes argumenter for å fjerne avskrivningene. Det første gjelder omleggingen av Kostra i 2000, der man blant annet . Den gangen ble eiendomsprisene på den kommunale byggmassen satt nokså ad hoc. Nå har det imidlertid gått over tjue år siden omleggingen, slik at målefeil forbundet med de initielle kapitalverdiene vil ha mindre betydning enn tidligere. Den andre er at byggmassen innebærer forpliktende kapitalkostnader over lengere tid, slik at beslutningsenhetene vil ha vanskligheter for å justere denne faktorinnsatsen for oppnå en mer effektivt ressursinnsats på kort sikt. Dette henger sammen med hvilket

tidsperspektiv man har for effektivitetsmålingen. Et tredje er at det er vanskelig å måle kostnadene i faste priser. Man ønsker ikke at høye eiendomspriser i et område skal påvirke effektivitetsmålingen. I Kostra verdsettes kapital basert på historisk kostnader. Det er for det fjerde tvilsomt at lineære av- og nedskrivninger av realkapitalen reflekterer kapitalens reelle forvitring.

Like fullt er av- og nedskrivninger driftskostnader. Prinsipielt bør det i inkluderingen av delkomponentene i kostnadsvariabelen ikke være avgjørende om kostnadene kostnadsføres direkte eller om kostnadene kapitaliseres for så å kostnadsføres over tid. Nedskrivninger kostnadsføres ikke alltid når kostnaden reelt skjer, men den reflekterer likevel også kapital-slit. I og med at bygg har en levetid på omtrent førti år, er heller ikke utfordringene med de omtrentlige byggverdiene satt i 2000 like betydelige lenger.

En mulighet kunne være å fortsatt ekskludere av- og nedskrivningene knyttet til eiendom, men isteden inkludere arealvariabler knyttet til byggmassen og eventuelt utearealene i estimeringsmodellen. I så fall ville det fort bli mange innsatsfaktorer i modellen. Eventuelt kunne man modellert dem som kontekstuelle variabler isteden.

For eksempel består grunnskolens realkapitalmasse ikke bare av bygg og anlegg, men også mobil realkapital. Merk at mobil kapital innbefatter maskiner, utstyr, transportmidler, immateriell kapital og kultiverte biologiske ressuser. Ifølge Nasjonalregnskapet knytter om lag en fjerdedel av realkapitalmassen i undervisningssektoren seg til mobil realkapital. Av dette var mesteparten immateriell kapital, men resten var i maskiner og utstyr – og i noe mindre grad – transportmidler. Merk at dette estimatet ekskluderer arealer som ikke opparbeidede, som ikke regnes som realkapital.

I omsorgssektoren samlet står derimot bygg og anlegg for nærmere 95 prosent av realkapitalmassen. For øvrig har mobil realkapital som regel en svakere prisutvikling enn bygg- og anlegg, slik og at den vil tillegges relativt mer betydning om man deflaterer realkapitalen med disaggregerte kapitaldeflaterer som deretter legger fastpriskapitalen sammen til en beholdningsvariabel. Konseptet «kapitaltjenester» håndterer i prinsippet forskjeller i årlig avkastningskrav på realkapital på grunn av ulik prisutvikling, levetid og eventuelt risiko. Det synes imidlertid ikke som riktig prioritering nå i videreutviklingen av TBU og KMDs effektivitetsmodell for offentlig sektor.

Det er klart at store investeringer i mobil realkapital kan påvirke kvaliteten på produksjonen i kommunal sektor, for eksempel elevenes læring i tilfellet undervisningssektoren. Hvis en kommune for eksempel investerer betydelig i nye nettbrett med tilhørende programvarer og et stort og moderne bibliotek vil det forventes å komme undervisningen til gode. Denne formen for realkapital kostnadsføres imidlertid over flere år gjennom avskrivninger og eventuelt nedskrivninger, istedenfor å kostnadsføres direkte. Kostnadsføringen over tid gjenspeiler at man ser på utleggene som investeringer som vil kaste av seg over tid. Siden avskrivninger ikke inkluderes i SØFs kostnadsmaal, leder dette til en overvurdering av effektiviteten i kommuner med relativt omfattende mobil realkapitalmasse sammenliknet med andre kommuner. Tilsvarende skjevheter vil også ramme andre delsektorer av kommunal sektor. Det vil imidlertid skje i ulik omgang, ettersom betydningen av mobil realkapital (og bygg- og anleggskapital) varierer over sektor.

I tidligere versjoner av SØFs effektivitetsmodell har sysselsetting blitt benyttet som innsatsfaktor istedenfor bruttodriftsutgifter. Dette innebærer at man neglisjerer kostnader som ikke knytter seg til arbeidskraft og forskjeller i kvaliteten på arbeidskraftene som reflekteres i lønnskostnadene. Skillet mellom ulike former for arbeidskraft vil håndtere noen av de tilknyttede kvalitetsforskjellene. Det kan naturligvis også være regionale lønnsforskjeller, som ikke håndteres om man benytter bruttodriftsutgifter som innsatsfaktorvariabel. Der- som man benytter et sentralitetsmaal som en kontekstuell variabel kan dette problemet forventes å bli mindre.

Bruk av bare sysselsetting som mål for faktorinnsatsen innebærer ikke bare at man neglisjerer realkapitalen, men også produktinnsatsen (altså annen intern faktorinnsats enn arbeidskraften og realkapitalen). I regnskapene involverer produktinnsatsen vareforbruk (varekjøp og endring i varebeholdningen) og andre driftskostnader (tjenestekjøp, innleiekostnader og energikostnader). Disse kostnadskomponentene kommer med om man velger bruttodriftsutgifter (strengt tatt bruttodriftskostnader) som mål for faktorinnsatsen. Bruttodriftskostnadene involverer også tap av fordringer, men dette er strengt tatt en finansiell post i realøkonomisk forstand som ikke bør med, selv om den inngår i driftsregnskapet.

I tidligere effektivitetsmodeller har SØF også skilt mellom utgifter til ulike driftskostnader. Blant annet har man benyttet lønnskostnader som en proxy for kvaliteten på arbeidskraften, selv om den i realiteten i tillegg til utdanning i stor grad avhenger av ansenitet som ikke bare reflekterer økt kvalitet. I tidligere effektivitetsmodeller har man også skilt mellom lønnskostnader og andre driftskostnader. Fordelen med dette er at man da får modellert hvordan innsatsfaktorene substituerer og komplementerer hverandre.

Et annet forhold er felles kommunene fellesutgifter til blant annet administrasjon og infrastruktur som kommer flere kommunale delsektorer til gode. Selv om dette strengt tatt er innsatsfaktor for kommunen, vil det antakelig ikke være hensiktsmessig å behandle det som det i sektorer der kostnadsandelen er lav. Felleskostnadene har i slike tilfeller gjerne begrenset innvirkning på produksjonen, og man kan også stille spørsmålsteget ved at grunnskolen skal straffes i effektivitetsmålingen for ineffektiv administrasjon. Et alternativ vil da være å behandle felleskostnadene som en kontekstuell variabel, om de har betydning for sektoren man ser på. I DEA-modeller der flere sektorer inngår på en gang, er det uproblematisk å la dem inngå som en innsatsfaktor på linje med andre innsatsfaktorer.

3.2.2 Generelle betraktninger om produkter

Delsektorene i kommunal sektor tilbyr ulike typer produkter, ettersom hvilke behov de er ment å tilfredsstille. Det er likevel tilfeller der skillet ikke er helt klart, som tilfredsheten til barn på småskolen og skole- og fritidsordningen, og tilbud av tjenester til eldre som potensielt både kan hjelpes av hjemmehjelpstjenesten eller bo på eldreheim. Konkurransesituasjonen mellom de kommunale tjenesteytere og liknende private aktører varierer også mellom de kommunale delsektorene.

De kommunale delsektorene som SØF tar for seg i sine effektivitetsanalyser innbefatter alle gode volummål for produksjonen. Med unntak av grunnskolen savner vi imidlertid gode indikatorer for produksjonens kvalitet. Kvalitetsindikatorer for produksjonen vil ofte være mindre tilgjengelige enn volummål, fordi de typisk hviler på gjennomføringen og tilgjengeliggjøring av spørreundersøkelser eller på individdata og andre mikrodata. Eksempler er dødelighet eller funksjonsdyktighet for pasienter på sykehjem. Effektivitetsmålinger uten gode kvalitetsindikatorer kan også gi innsikter, men man må være klar over at målte effektivitetsforskjeller ikke nødvendigvis skyldes reelle effektivitetsforskjeller, men også kan skyldes forskjeller i kvalitet. På en annen side taler målerelaterte utfordringer som at produktene er forholdstall, reflekterer kvalitative variabler, kan innta negative verdier og eller er særlig utsatt for målefeil for at man heller bør behandle kvalitetsaspektene som kontekstuelle variabler, jamfør seksjon 2.1.1. I mange tilfeller er heller ikke kvalitetsaspektenes innvirkningen på produksjonsteknologiens front særlig stor i forhold til innvirkningen til volumproduktene. Likevel kan det bygge opp under analysenes troverdighet å ha dem med.

3.2.3 Generelle betraktninger om kontekstuelle variabler

Forutsetningene for å få mye produksjon ut av faktorinnsatsen i kommunal tjenesteproduksjon avhenger betydelig av omstendighetene. Det er derfor viktig å identifisere og ta hensyn

til kontekstuelle variabler i effektivitetsmålingen. Dette gjelder spesielt når man benytter seg av metodikk som konvensjonell DEA, der det ikke kontrolleres for faste effekter over tid.

Mange av de aktuelle empiriske spesifikasjonen for effektivitetsanalyser forutsetter at kontekstuelle variabler har tilsvarende innvirkning på ulike beslutningsenheter, jmfør seksjon 2.1.6 og delkapittel 2.5. Dette kravet kan brytes om de kontekstuelle variablene har ulik betydning i ulike sammenhenger, eller egentlig er heterogen samling av forskjellige delkomponenter med ulik innvirkning på produksjonen.

Mange av de aktuelle kontekstuelle variablene går på tvers av sektorene og angår sosio-økonomiske og geografiske forhold. I dagens målinger utnyttes kontekstuelle variabler i grunnskolen ved at man beregner hvordan faktiske skolepoeng avviker fra de forventede skolepoeng i en såkalt skolebidragsvariabel (jmfør seksjon 2.1.3 og seksjon 2.1.5 for detaljer). De kontekstuelle variablene som inngår i disse regresjonene vedrører utdanning, inntektsnivå og innvandring (Borge, Kråkenes og Nyhus 2020). En slik korreksjon er ukonvensjonell og innebærer redusert estimeringseffektivitet ved at man ikke foretar regresjonen for virkningene av de kontekstuelle variablene simultant med regresjonen for effektivitetsmålingen. Dermed tar de to delregresjonene ikke gjensidig hensyn til hverandre. På en annen side kan unyttelse av mikrodata på de kontekstuelle faktorene gjennom datatilgang forutsette at man estimerer virkningene av de kontekstuelle variablene i et eget steg.

Andre aktuelle forhold aldersstruktur, befolkningstetthet, kvinneandel, andel skilte, sykefravær, næringsstruktur, sysselsettingsgrad, helse og øvrige levekår. Viktigheten av disse kontekstuelle variablene vil åpenbart variere, ettersom hvilken kommunal delsektor det er snakk om. Dersom man har tilgang til mikrodata på hvem som mottar de kommunale tjenester og eventuelt deres nærmeste familie, kan man også foreta sektor-spesifikke tilpasninger på mikronivå i tråd med informasjonen om tjenestenes konsumenter.

Enkelte kontekstuelle variabler er sektor-spesifikke. Sentrale sektor-spesifikke forhold inkluderer innslaget og forholdene knyttet til alternative private tilbydere og forholdene i komplementær tjenesteproduksjon, typisk innenfor kommunal sektor. Det kan også være snakk om biprodukter som ikke anses viktige nok til å inkluderes som produkter i effektivitetsestimeringen eller som måles innen et avgrenset intervall eller innehar andre skalaegenskaper som bryter med DEA-funksjonen.

En kontekstuell variabel som benyttes i SØFs offisielle effektivitetsmålinger for KMD og TBU er sonevariabelen, som diskutert i seksjon 2.1.6. Sonevariablene er et slags invers sentralitetsmål, som måler geografisk avstand til nærmeste institusjoner, hvilket også gjør den sektor-spesifikk. Selv om det er grunn til å tro at sentralitet korrelerer med sektoreffektiviteten for en del sektorer, kan man stille spørsmål med ved inkluderingen av denne kontekstuelle faktoren i modellen for stedsbundne delsektorer, om man er opptatt av kausalitet i målingen. For mobile tjenester utgjør reisekostnader en viktig faktor, slik at begrunnelsen for å ha med sonevariabelen er sterkere. For stedsbundne tjenester kan eiendomspriser, konkurranseforhold og arbeidsmarkedet eventuelt begrunne inkluderingen av variabelen, men den kan også fort tolkes som en variabel som bidrar til å manipulere resultatene i favør perifere strøk og i disfavør sentrale strøk. Særlig gjelder dette når man deler kostnadsvariabelen ukritisk på denne faktoren. Dersom man eventuelt skulle tatt den med, burde den behandles som en egen variabelen i en separat estimeringsstrinn som estimeres simultant (vi referer igjen til delseksjon 2.1.6 for mer om dette).

3.2.4 Grunnskolen

Grunnskolen utgjør den nest største delsektoren innen kommunal sektor, etter pleie- og omorgssektoren. En oversikt over relevante variabler for effektivitetsmålinger i delsektoren

er i gitt i tabell 3.1. Effektivitetsmålingen i grunnskolen har kommet relativt langt i forhold til de andre sektorene når det kommer til måling av produksjonens kvalitet. Dette kommer både uttrykk til skolebidragindikatoren og skolemiljøindikatoren. Samtidig ser vi forbedringspotensialer, som eksempelvis; at omfanget av spesialundervisning eller tilrettelegging i skolen ikke tas hensyn til.

Merk at foreldrenes utdanning, inntekt og innvandrerbakgrunn hensyntas indirekte gjennom skolebidragsindikatoren. Data om spesialundervisning, fremmedspråk mm kan hentes fra Grunnskolens informasjonssystem (GSI). Indikatoren forteller dermed hvordan elevprestasjonene avviker fra det man kan forvente ut i fra kontekstuelle forhold. Den konvensjonelle måten å håndtere dette på ville imidlertid være å beregne virkningene i simultant i en egen regresjon for kontekstuelle variabler, jamfør seksjon 2.1.3.

Skolemiljøvariabelen er et annet eksempel på en variabel som både utgjør en komposittindeks og er reskalert. I første instans er dette en komposittindeks bestående av delindikatorer for mellomtrinnet og ungdomstrinnet. I andre instans består den av tre likt vektete indikatorer for elevenes medbestemmelse, mobbing og trivsel. I tredje instans er de tre delindikatorene igjen basert på svar på tre spørsmål stilt til elevene i elevundersøkelsen. Det er forståelig at SØF har valgt å operere med skolemiljø som ett produkt i DEA-modellen og ikke mange produkter, som kunne bidratt til overdøve betydningene av elevenes skolebidrag på tvers av trinn. Det bør likevel bemerkes at sammensetningen av indeksen synes nokså tilfeldig. Vi henviser til seksjon 2.1.5 for videre diskusjon.

Tabell 3.1: Relevante variabler i effektivitetsanalyser av grunnskolen.

Variabel	Anvendelse	Laveste nivå	Kilde
Innsatsfaktorer			
Bruttodriftsutgifter fratrukket arbeidsgiveravgift og avskrivninger, dividert på sonevariabelen	Nåværende	Insititusjon	Kostra (SSB)
Driftskostnader, annen variant	Potensiell	Insititusjon	Kostra (SSB)
Årsverk for lærere	Tidligere	Insititusjon	Utdanning (SSB og Udir)
Årsverk for assistenter	Tidligere	Insititusjon	Utdanning (SSB og Udir)
Deflaterte lønnskostnader som proxy for lærerkvalitet	Tidligere	Insititusjon	Utdanning (SSB og Udir)
Andre driftsutgifter enn lønnskostnader	Tidligere	Insititusjon	Kostra (SSB)
Realkapital, mobil realkapital og eiendom	Potensiell	Insititusjon	Kostra (SSB)
Produkter			
Gjennomsnittlig skolebidrag eller skolepeong på småskoletrinnet, multiplisert med antall elever	Nåværende	Individ	Utdanning (SSB og Udir)
Gjennomsnittlig skolebidrag eller skolepeong på mellomskoletrinnet, multiplisert med antall elever	Nåværende	Individ	Utdanning (SSB og Udir)
Gjennomsnittlig skolebidrag eller skolepeong på ungdomsskoletrinnet, multiplisert med antall elever	Nåværende	Individ	Utdanning (SSB og Udir)
Gjennomsnittlig skolemiljø for mellomtrinnet og ungdomstrinnet multiplisert med antal elever. Fanger opp elevdemokrati, mobbing og trivsel	Nåværende	Individ	Utdanning (SSB og Udir)
Sektorspesifikke kontekstuelle variabler			
Utdanningsprofil, som sysselsettingsandel med høyere utdanning	Nåværende	Institusjon	Utdanning (SSB)
Inntekstmål, som median personinntekt	Nåværende	Institusjon	Inntekt (SSB)
Innvandringsandel, gjerne dissaggregert	Nåværende	Institusjon	Befolkning (SSB)

Variabel	Anvendelse	Laveste nivå	Kilde
Sysselsetningsgrad av befolkning i arbeidsfør alder eller hele befolkningen	Tidligere	Institusjon	Befolkning og sysselsetting (SSB)
Befolkningstetthet	Tidligere	Kommune	Befolkning og geografi (SSB)
Sentralitetsmål, som sonevariabelen	Nåværende	Institusjon	Geografi (SSB)
Kvinneandel	Potensiell	Kommune	Befolkning (SSB)
Barn av skilte foreldre	Tidligere	Individ	Befolkning og sosialstatistikk - (SSB)
Alderssammensetning	Potensiell	Institusjon	Befolkning (SSB)
Skolestørrelse og gruppestørrelse	Potensiell	Institusjon	Utdanning (SSB og Udir)
Årstimer til særskilt språk- eller spesialundervisning	Tidligere	Individ	Utdanning (SSB og Udir)
Innslag av private skoler	Potensiell	Kommune	Utdanning (SSB og Udir)

3.2.5 Pleie- og omsorgstjenester

Som vi så i seksjon 3.1.5 utgjør pleie- og omsorgssektoren den klart største største delsektoren innen kommunal sektor. Pleie- og omsorgssektoren omfatter ulike typer av tjenester på institusjon og hjemmehjelp og er således ganske kompleks. Om man overkommer tilgangsbARRIERER med datasentitivitet, ligger det et stort potensial for utnyttelse av individdata for helsesektoren, både når det gjelder kvalitet i omsorgstjenesten og relevante kontekstuelle variabler. Her kan blant annet data på individnivå fra Kommunalt pasientregister (KPR) /IPLOS for pleie og omsorgssektoren spille en viktig rolle, eventuelt sammenkoblet med andre individdata i tilfelle kontekstuelle variabler. Helsetilstand, diagnosehyppighet eller funksjonsdyktighet for eldre og andre funksjonsnedsatte vil stå sentralt. Andre mikrodata kan også være relevant, for eksempel om nærhet til funksjonsfriske familiemedlemmer som i praksis kan avlaste kommunene. En oversikt over relevante variabler i effektivitetsanalyser av pleie- og omsorgstjenester er gitt i tabell 3.2 under.

Tabell 3.2: Relevante variabler i effektivitetsanalyser av pleie- og omsorgstjenester.

Variabel	Anvendelse	Laveste nivå	Kilde
Innsatsfaktorer			
Bruttodriftsutgifter fratrukket arbeidsgiveravgift og avskrivninger, dividert på sonevariabelen	Nåværende	Institusjon	Kostra (SSB)
Årsverk av personell med relevant fagutdanning, korrigert for bosettingsmønster	Tidligere	Institusjon	Omsorg (SSB)
Årsverk av personell uten relevant fagutdanning, korrigert for bosettingsmønster	Tidligere	Institusjon	Omsorg (SSB)
Andre driftsutgifter enn lønnskostnader, korrigert for bosettingsmønster	Tidligere	Institusjon	Kostra (SSB)
Arealressurser og bygningskapital	Potensiell	Institusjon	Geografi (SSB)
Mobil realkapital (utstyr, maskiner, fremkomstmidler med mer)	Potensiell	Institusjon	Kostra (SSB)
Produkter			
Liggedøgn ved institusjoner	Nåværende	Institusjon	Omsorg (Hdir og SSB)
Produkt for liggedøgn i institusjon, korttid	Tidligere	Institusjon	Omsorg (Hdir og SSB)
Produkt for liggedøgn i institusjon., langtid	Tidligere	Institusjon	Omsorg (Hdir og SSB)
Institusjonsbeboere med omfattende bistandsbehov	Nåværende	Institusjon	Omsorg (Hdir og SSB)
Enerom i institusjon	Nåværende	Institusjon	Omsorg (Hdir og SSB)
Timer til hjemmesykepleie	Nåværende	Institusjon	Omsorg (Hdir og SSB)

Variabel	Anvendelse	Laveste nivå	Kilde
Timer til praktisk bistand	Nåværende	Institusjon	Omsorg (Hdir og SSB)
Institusjonsbeboere med omfattende bistandsbehov	Nåværende	Institusjon	Omsorg (Hdir og SSB)
Hjemmeboende pasienter	Potensiell	Institusjon	Omsorg (Hdir og SSB)
Funksjonsendring	Potensiell	Individ	Helse (KPR og IPLOS)
Pasientkategori	Potensiell	Individ	Helse (KPR og IPLOS)
Sektorspesifikke kontekstuelle variabler			
Helsetilstand og funksjonsdyktighet	Potensiell	Individ	Helse (KPR og IPLOS)
Diagnosehyppighet	Potensiell	Institusjon	Helse (KPR og IPLOS)
Demografi (alder og innvandring)	Potensiell	Individ	Helse (KPR og IPLOS)
Sosial profil (utdanning og inntekt)	Potensiell	Individ	Helse (KPR og IPLOS)
Ulykkeforekomster	Potensiell	Institusjon	Helse (KPR og IPLOS) og sosialstatistikk (SSB)
Innslag av private	Potensiell	Kommune	Omsorg (Hdir og SSB) og foretaks- og virksomhetsregisterene (SSB og BREG)
Geografisk avstand til hjemmeboende pasienter	Potensiell	Institusjon	Helse (KPR og IPLOS)
Avstand til funksjonsfriske pårørende	Potensiell	Kommune	Befolkningsdata på individnivå (SSB)

3.2.6 Barnehager

Barnehagetjenesten utgjør den tredje og minste av de tre kommunale tjenestene som inngår i de offisielle effektivitetsmålingene som SØF gjennomfører for kommunal sektor på vegne av KMD og TBU. Også i barnehagene mangler det gode kvalitetsindikatorer utover leke- og uteareal, som alternativt kan betraktes som innsatsfaktorer. Her kunne muligens den nasjonale foreldreundersøkelsen for barnehager ha blitt benyttet, selv om den kun sendes ut annet hvert år. Barnehager er et eksempel på en delsektor der investerings- og vedlikeholdskostnader er viktig og ignorering av kapitalslitet kan føre til misvisende vurderinger. Ellers kan det være et poeng å kontrollere for innslaget av private barnehager i barnehagen og befolkningskarakteristika knyttet til foreldre til barn i barnehagealder i kommunene. I tabell 3.3 har vi gitt en oversikt over aktuelle variabler i effektivitetsmålinger i kommunale barnehager.

Tabell 3.3: Relevante variabler i effektivitetsanalyser av barnehager.

Variabel	Anvendelse	Laveste nivå	Kilde
Innsatsfaktorer			
Driftskostnader		Institusjon	Kostra (SSB)
Årsverk av personell med pedagogisk utdanning		Institusjon	Sysselsetting (SSB)
Årsverk av personell uten pedagogisk utdanning		Institusjon	Sysselsetting (SSB)
Andre driftsutgifter enn lønnskostnader		Institusjon	Kostra (SSB)
Lønnskostnader som proxy for kvaliteten hos personalet		Institusjon	Sysselsetting (SSB)
Dekningsgrad i kommunen		Kommunen	Udir
Leke- og uteareal		Institusjon	Barnehager (SSB og Udir)
Produkter			
Oppholdstimer 0 til 2 år		Institusjon	Barnehager (SSB og Udir)
Oppholdstimer 3 til 5 år		Institusjon	Barnehager (SSB og Udir)
Leke- og uteareal		Institusjon	Barnehager (SSB og Udir)

Variabel	Anvendelse	Laveste nivå	Kilde
Tilfredshetsindikatorer		Individ	Den nasjonale foreldreundersøkelsen (Udir)
Sektorspesifikke kontekstuelle variabler			
Alderssammensetning		Institusjon	Befolkning (SSB)
Innvandringsandel		Institusjon	Befolkning (SSB)
Konkurransesituasjon for private for-profit- og non-profit-barnehager		Institusjon	Barnehager (SSB og Udir)
Antall avdelinger		Institusjon	Barnehager (SSB og Udir)
Private barnehager lokalt		Institusjon	Barnehager (SSB og Udir)

3.2.7 Barnevernet

SØF har foretatt effektivitetsmålinger fra barnevernet. Disse målingene inkluderer volumbaserte mål for produksjon og faktorinnsats, som fremstår relevante og dekkende. Vår vurdering er derfor at en eventuell forbedring at datatilfanget for barnevernet vil knytte seg til mål på kvaliteten og kontekstuelle forhold. For eksempel vil relativt mange underkjenninger det kommunale barnevernets vedtak være en indikere relativt lav kvalitet på produksjonen. Videre er det åpenbart befolkningsstørkuten påvirker barnevernets arbeidsvilkår. I tillegg vil trolig barnevernets produksjon påvirkes av annen kommunal tjenesteproduksjon, da særlig kulturtilbudet rettet mot unge, skole- og fritidsordningen, grunnskolen. Dessuten er det et faktum at den kommunale barnevernstjenesten i mange tilfeller vil være organisert som interkommunale samarbeid, hvilket både kan påvirke den faktiske effektiviteten og medføre målutfordringer så lenge kommuner utgjør observasjonsenhetene. En oversikt over relevante variabler i effektivitetsanalyser av barnevernet er gitt i tabell 3.4.

Tabell 3.4: Relevante variabler i effektivitetsanalyser av barnevernet.

Variabel	Anvendelse	Laveste nivå	Kilde
Innsatsfaktorer			
Driftskostnader	Potensiell	Institusjon	Kostra (SSB)
Årsverk i barnevernet	Tidligere	Institusjon	Barnevernet (Bufdir og SSB)
Andre driftsutgifter til tiltak enn lønnskostnader	Tidligere	Institusjon	Kostra (SSB)
Årsverk med lavere utdanning	Potensiell	Institusjon	Barnevernet (Bufdir og SSB)
Årsverk med høyere utdanning	Potensiell	Institusjon	Barnevernet (Bufdir og SSB)
Produkter			
Undersøkelser av barnevernet	Tidligere	Institusjon	Barnevernet (Bufdir og SSB)
Barn med hjelpetiltak	Tidligere	Institusjon	Barnevernet (Bufdir og SSB)
Barn med omsorgstiltak	Tidligere	Institusjon	Barnevernet (Bufdir og SSB)
Innsendte klager på barnevernet	Potensiell	Kommune	Barnevernet (Bufdir og SSB)
Klager på barnevernet som har fått medhold	Potensiell	Kommune	Barnevernet (Bufdir og SSB)
Sektorspesifikke kontekstuelle variabler			
Foreldrenes utdanning	Potensiell	Institusjon	Barnevernet (Bufdir og SSB)
Alderssammensetningen blant barn	Potensiell	Institusjon	Barnevernet (Bufdir og SSB)
Innvanderandel	Potensiell	Institusjon	Barnevernet (Bufdir og SSB)
Medlemmer og ansatte i frivillige organisasjoner innenfor kultur og idrett	Potensiell	Institusjon	Foretaks- og virksomhetsregisterene (SSB og BREG)

3.2.8 Skole- og fritidsordningen

Skole- og fritidsordningen (SFO) utgjør en annen delsektor som har vært gjenstand for SØFs tidligere effektivitetsmåling. SFO utgjør en tjeneste for barn på småskolen, altså første til fjerde klasse. I tilknytning til noen skoler tilbys også SFO-liknende tjenester for barn ved mellomskolen (femte til sjuende klasse) og fritidsklubber for ungdommer ved ungdomskolen (åttende til tiende klasse), men disse tjenester inngår ikke målingen. Flere utfordringer melder seg ved effektivitetsmålinger innenfor SFO. Disse knytter seg i stor grad til overlapp og interaksjon med andre kommunale delsektorer; da særlig skolen, men også fritidstilbud for barn. Blant annet er det ikke alltid lett å skille på en konsistent måte mellom kostnader knyttet til eiendomskapital og arbeidskraft, som både benyttes i småskolen og på SFO. Vanskeligheter med å skille mellom produksjonen på småskolen og på SFO taler for at man kunne ha målt effektiviteten i delsektorene samlet.

Overlappet og interaksjonen mellom SFO og andre kommunale delsektorer spiller også inn når man skal måle kvalitet i SFO-tjenesten. Blant annet er trivsel og forebygging av mobbing, som reflektert av elevundersøkelsen, også relevant for SFO. Her kunne man sett på forskjeller i disse indikatorene for SFO-barn og barn på småskolen som ikke går på SFO over kommuner, skjønt fordelingen av barna mellom disse gruppe trolig er selektiv ved at barn av ressursvake foreldre tenderer til å være underrepresentert på SFO. Omfanget av tilbudet av fritidsaktiviteter på SFO er en annen relevant kvalitetsindikator for delsektoren. I SØFs effektivitetsmålinger har ikke innebefattet kvalitetsindikatorer for produksjonen. En oversikt over aktuelle variabler i effektivitetsmålinger knyttet til SFO er gitt i tabell 3.5.

Tabell 3.5: Relevante variabler i effektivitetsanalyser av skole- og fritidsordningen.

Variabel	Anvendelse	Laveste nivå	Kilde
Innsatsfaktorer			
Driftskostnader	Potensiell	Institusjon	Kostra (SSB)
Lønnskostnader til SFO	Tidligere	Institusjon	Kostra (SSB)
Andre driftsutgifter enn lønnskostnader	Tidligere	Institusjon	Kostra (SSB)
Produkter			
Antall brukere med heltidsplass	Tidligere	Institusjon	Utdanning (SSB)
Antall brukere med deltidsplass	Tidligere	Institusjon	Utdanning (SSB)
Timer åpent i året	Potensiell	Institusjon	Utdanning (SSB)
Trivsel og forebygging av mobbing	Potensiell	Individ	Utdanning (SSB)
Sektorspesifikke kontekstuelle variabler			
Alderssammensetning	Potensiell	Institusjon	Befolkning (SSB)
Median personinntekt	Potensiell	Institusjon	Inntekt (SSB)
Innvandringsandel	Potensiell	Institusjon	Befolkning (SSB)
Alternativ fritidstilbud for unge, idrett og kultur	Potensiell	Kommune	Foretaks- og virksomhetsregisterene (SSB og BREG)
Utenfor arbeidsstyrken	Potensiell	Kommune	Sysselsetting (SSB)
Privatskolers kommunale markedsandel på hver hovedtrinn	Potensiell	Kommune	Utdanning (SSB)

3.2.9 Kulturelle tjenester

Kommunale kulturelle tjenester kan være utfordrende delsektor å måle effektivitet for, idet delsektoren er fragmentert med ulikt innslag av private tilbydere over segmenter og kommuner. En del kulturelle tjenester tilbys også av andre offentlige aktører enn kommunene eller er kun til stede i en del kommunene, som for eksempel museer og teatre. Like fullt har SØF foretatt flere effektivitetsanalyser på kulturelle tjenester i sine analyser av kommunal sektor. Deres analyser omfatter segmentene biblioteker, kulturskoler og fritidssentre.

For at bruttodriftsutgifter (eventuelt etter korreksjoner) kan brukes som innsatsfaktor, bør produksjonen reflekteres av produktene i estimeringsmodellen, hvilket kan være vanskelig for kommunale kulturtjenester. For noen kulturtjenester kan det også være relevant å ta for seg tilreisende, inkludert innpendlere, turister og yrkesreisende. En annen utfordring ved effektivitetsmåling for kommunale tjenester er at det er få tilgjengelige indikatorer for kvaliteten på produksjonen, som trolig også ville vært kostbart å samle inn. En oversikt over datatilgangen for kulturelle tjenester under kommunal sektor er gitt i tabell 3.6.

Tabell 3.6: Relevante variabler i effektivitetsanalyser av kulturelle tjenester.

Variabel	Anvendelse	Laveste nivå	Kilde
Innsatsfaktorer			
Årsverk i biblioteker	Tidligere	Institusjon	Kulturtilbud (SSB)
Utgifter til medier i biblioteker	Tidligere	Institusjon	Kulturtilbud (SSB)
Årsverk i kommunale kulturskoler	Tidligere	Institusjon	Kulturtilbud (SSB)
Årsverk ved kommunalt drevne fritidssentre	Potensiell	Institusjon	Kulturtilbud (SSB)
Andre driftskostnader i kultursektoren	Potensiell	Institusjon	Kulturtilbud (SSB)
Bruttodriftsutgifter (krever samsvar med produktene)	Potensiell	Kommune	Kulturtilbud (SSB)
Produkter			
Utlån av bøker	Tidligere	Individ	Kulturtilbud (SSB)
Utlån av andre medier	Tidligere	Individ	Kulturtilbud (SSB)
Elever i kommunal kulturskole	Tidligere	Individ	Kulturtilbud (SSB)
Timer åpen årlig per 1 000 innbyggere 6-20 år, kommunale fritidssentre	Potensiell	Institusjon	Kulturtilbud (SSB)
Antall besøk i året per 1000 innbyggere 6-20 år, i kommunale fritidssentre	Potensiell	Institusjon	Kulturtilbud (SSB)
Annen kommunal kinoproduksjon: Kommunale idrettsbygg og idrettsanlegg	Potensiell	Institusjon	Kulturtilbud (SSB)
Annen kommunal kinoproduksjon: Kino	Potensiell	Institusjon	Kulturtilbud (SSB)
Andre kulturproduksjon dominert av kommunal sektor	Potensiell	Institusjon	Kulturstatistikk for øvrig (SSB)
Sektorspesifikke kontekstuelle variabler			
Alderssammensetning	Potensiell	Institusjon	Kulturtilbud (SSB)
Pendlemønstre	Potensiell	Individ	Kulturtilbud (SSB)
Overnattingsaktivitet (relevant om kulturtjenester konsumert av tilreisende)	Potensiell	Institusjon	Overnattingsstatistikk (SSB og Statistikknett Reiseliv)
Medlemmer og ansatte i frivillige organisasjoner innenfor kultur og idrett	Potensiell	Kommunen	Foretaks- og virksomhetsregisterene (SSB og BREG)

3.2.10 Vann, avløp, avfall og renovasjon

Blant delsektorene av kommunal sektor som ikke har vært et fokus i kommunale effektivitetsanalyser skiller «vann, avløp, avfall og renovasjon» seg ut som en delsektor med potensial. Her fins det både gode data og flere effektivitetsstudier i forskningslitteraturen, selv om det er klare vanskeligheter med å ta tilfredsstillende hensyn til de store topografiske forskjellene i et land som Norge. Åpenbart kunne man delt inn i ulike delsektorer, for eksempel vann, avfall og avløp. I vår gjennomgang vil vi behandle dem sammen, både fordi inngangen til disse sektorene likner hverandre, og fordi de deler henførbare indirekte driftsutgifter. En oversikt over datatilganger for denne kommunale delsektoren er gitt i tabell 3.7.

Tabell 3.7: Relevante variabler i effektivitetsanalyser av vann, avløp, avfall og renovasjon.

Variabel	Anvendelse	Laveste nivå	Kilde
Innsatsfaktorer			
Direkte driftsutgifter, avfall	Potensiell	Institusjon	Husholdningsavfall (SSB)
Direkte driftsutgifter, avløp	Potensiell	Institusjon	Husholdningsavfall (SSB)
Direkte driftsutgifter, vann	Potensiell	Institusjon	Husholdningsavfall (SSB)
Henførbare indirekte driftsutgifter	Potensiell	Institusjon	Husholdningsavfall (SSB)
Produkter			
Husholdningsavfall	Potensiell	Institusjon	Husholdningsavfall (SSB)
Andel levert til materialgjenvinning inkludert biologisk behandling	Potensiell	Institusjon	Husholdningsavfall (SSB)
Hentinger av matavfall per år	Potensiell	Institusjon	Husholdningsavfall (SSB)
Årlige timer betjente gjenvinningsstasjoner	Potensiell	Institusjon	Husholdningsavfall (SSB)
Andre forhold som påvirker gebyret (antall)	Potensiell	Institusjon	Husholdningsavfall (SSB)
Fraksjoner som vert henta regelmessig, sorteres ut på sorteringsanlegg og hentes sporadisk (antall)	Potensiell	Institusjon	Husholdningsavfall (SSB)
Henteordning versus gjenvinningsstasjon	Potensiell	Institusjon	Husholdningsavfall (SSB)
Antall innbyggere tilknyttet kommunalt vannverk med tilfredsstillende prøveresultater (antall) for: E.coli-, intestinale enterokokker, farge og PH	Potensiell	Institusjon	Husholdningsavfall (SSB)
Sammensetningen av ulike typer avfall	Potensiell	Institusjon	Husholdningsavfall (SSB)
Sammensetningen av ulike typer håndtering	Potensiell	Institusjon	Husholdningsavfall (SSB)
Antall som får vann (befolkning, husholdninger, sysselsetting, virksomheter, tilreisende og studenter)	Potensiell	Institusjon	Bosetting (SSB) og sysselsetting (NAV og SSB), forekaks- og virksomhets registerene (BREG og SSB)
Sektorspesifikke kontekstuelle variabler			
Eierskap, form og sektor, og deltakelse i kommunalt samarbeid	Potensiell	Institusjon	Foretaks- og virksomhetsregisterene (SSB og BREG)
Kjøreavstander i kommunen for avfallshåndtering	Potensiell	Institusjon	Bosetting (SSB) og kartdata (Karverket og SSB)
Lengden på vannledningsnettet	Potensiell	Institusjon	Bosetting (SSB), kartdata (Karverket og SSB) og data på vannforsyning og avløp (NVE)
Grunnforhold, topografi og samfunnsgeografi	Potensiell	Institusjon	Kartdata (Statens Kartverk)
Antall innbyggere primært forsynt av henholdsvis desinfisert overflatevann, grunnvann og behandlet vann	Potensiell	Institusjon	Husholdningsavfall (SSB)

3.2.11 Andre delsektorer

En fullstendig effektivitetsanalyse av kommunal sektor ville innebåret analyser av en rekke andre delsektorer enn de vi har behandlet overfor, både knyttet til kommunal forvaltning (finansiert over bevilgninger) og kommunal forretningsdrift (selvfinansierende). En oversikt over disse delsektorene er gitt i tabell 3.8, der vi også har listet opp utfordringer forbundet med effektivitetsmåling og oppgitt mulige kilder for behandling av hver delsektor.

Tabell 3.8: Utfordringer og kilder ved effektivitetsanalyser av andre grener kommunal sektor.

Delsektor	Utfordring med effektivitetsmåling	Kilde
Administrasjon	Bidrar til alle sektorer	Kostra (SSB) og KS-statistikk
Brann- og ulykkesvern	Få årlige henvendelser på kommunalt nivå	Brannstatistikk og statistikk på ulykkesvern (DSB) og Brann- og ulykkesvern (SSB)
Eiendom	Fragmentert sektor og uklare målsetninger	Eiendomsforvaltning i kommunesektoren og Kostra (SSB)
Fellesutgifter (driftsutgifter som faktureres samlet, men som berører flere funksjoner)	Bidrar til alle sektorer	Kostra (SSB)
Fysisk planlegging, natur, miljø og nærmiljø	Utfordrende å koble innsatsfaktorer og produkter	Kulturminneforvaltning, miljøforvaltning, byggetillatelse og Kostra (SSB)
Kirke og andre trossamfunn	Kirkens aktiviteter hører generelt ikke inn under kommunal sektor	Kirkelige tjenester og Kostra (SSB), andre trossamfunn (Regjeringen)
Kommunale boliger	Utfordrende å koble innsatsfaktorer og produkter	Kostra og eiendomsforvaltning i kommunesektoren (SSB)
Næringsforvaltning	Utfordrende å koble innsatsfaktorer og produkter	Kostra (SSB) og Foretaks- og virksomhets-registerene (SSB og BREG)
Politisk styring og kontrollvirksomhet	Bidrar til alle sektorer og egenaktivitet	Kostra (SSB)
Samferdsel	Typisk ikke på kommunenivå	Kostra, statistikker på kollektivtrafikk og kommunale veier (SSB)
Sosialsektoren	Fragmentert sektor	Kostra, statistikker innenfor helse og sosiale forhold og kriminalitet (SSB)

4 Empiriske analyser

Vi vil i dette kapitlet teste betydningen av ulike datarelaterte og metoderelaterte valg i effektivitetsberegninger empirisk. Vi begynner med å ta for oss den sektorielle effektivitetsanalysen og målingen av effektivitet over tid. Her har vi valgt å fokusere på grunnskolen, som er en delsektor med relativt godt datagrunnlag og dermed egner seg godt til å illustrere betydningen av ulike forutsetninger. Til slutt tar vi for oss produksjonsindeksen, som indikerer utviklingen i den kommunale tjensteproduksjonen på tvers av delsektorer.

4.1 Effektivitetsanalyse

I vår effektivitetsanalyse av kommunal sektor vil vi fokusere på grunnskolen. Vi har ikke kommentert konsekvensene av alle valg knyttet til metodikk og data i dette delkapitlet. Isteden referer vi til kapittel 2 om metodikk og kapittel 3 om data. Det sistnevnte kapitlet dokumenterer også data som tas i bruk.

I vår utgangsmoell benytter vi bruttodriftsutgifter i faste 2019-kroner som eneste innsatsfaktor, elever på småtrinnet (1. til 4. klasse), mellomtrinnet (5. til 7. klasse) og ungdomstrinnet (7. til 10. klasse) som innsatsfaktorer og ingen kontekstuelle faktorer. Vi referer til seksjon 3.2.4 for mer om datatilfanget for grunnskolen. I modellene har vi videre bootstrappet DEA med variabelt skalautbytte. Som SØF beregner vi effektivitetsscorene for to år (2018 og 2019) under forutsetning om felles teknologifront, men rapporterer kun resultatene for det siste året. Vi har begrenset vårt utvalg til kommuner som har elever både på småtrinnet, mellomtrinnet og ungdomstrinnet, og der de anvendte modellvariablene er tilgjengelige.

Vi starter med å behandle alternative valg av innsatsfaktorer. I de påfølgende analysene trekker vi fra arbeidsgiveravgift og avskrivninger fra bruttodriftsutgiftene når vi måler faktorinnsatsen i tråd med SØFs valg. Vi fortsetter med å ta for oss behandlingen av seleksjon og skalautbytte. Deretter retter vi søkelyset mot betydningen av tjenestekvaliteten. Mot slutten av delkapitlet ser vi på to alternative metoder til DEA, nærmere bestemt SFA og StoNED. Vi avslutter med å ta for oss betydningen av kontekstuelle variabler. I alle DEA-kjøringene med kontekstuelle variabler benytter vi oss av Simar og Wilson (2007) sin estimeringsprosedyre for to-steps-DEA.

Vi vil i våre analyser fokusere på innsatsfaktorbesparende teknisk effektivitet (heretter referert til som teknisk effektivitet), altså ulik utnyttelser av teknologien, og for enkelthets skyld referere til det som effektivitet. Vi vil kort også komme innom skalaeffektivitet i forbindelse med forutseningene om skalautbytte, men dette er ikke noe stort fokus. Vi ser ikke direkte på effektiviteten tilknyttet til faktorbrukens tilpasning til faktorprisene, såkalt allokativ effektivitet, men det kommer inn implisitt når vi benytter aggregerte mål for faktorinnsatsen. Samtidig får man ikke med substituerbarheten og komplementariteten mellom de underliggende faktorene når man kun benytter seg av én faktorinnsats. Vi kommer inn på endring i teknisk effektivitet over tid i delkapittel 4.2.

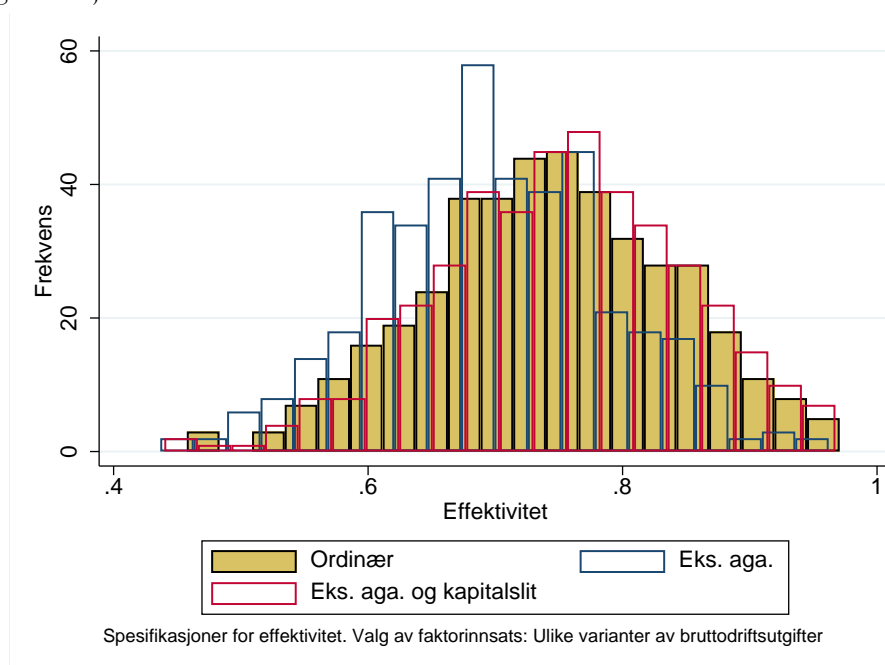
For å illustrere betydningen av ulike forhold, vil for det første illustrere forskjeller i frekvensfordelingene til effektivitetsscorene ved alternative forutsetninger gjennom søylediagrammer. For det andre tar vi i bruk korrelasjonsplott for å illustrere samvariasjonen

mellom effektivitetsscorene beregnet ved alternative forutsetningen. I denne forbindelse har vi også rapportert to korrelasjonskoeffisienter; Pearsons produktmomentskorrelasjonskoeffisient og Kendalls rangskorrelasjonskoeffisient. Pearsons produktmomentskorrelasjonskoeffisient måler lineær korrelasjon utgjør det mest utbredte korrelasjonsmålet og omtales derfor gjerne av mange bare som korrelasjonskoeffisienten. Kendalls rangskorrelasjonskoeffisient, også kjent som Kendalls tau-koeffisient, er derimot et korrelasjonsmål for den ordinale relasjonen mellom observasjonene. Begge koeffisientene er definert i rommet og med fra -1 til og med 1, der -1 angir perfekt negativ korrelasjon, 0 angir ingen korrelasjon og 1 angir perfekt positiv korrelasjon.

4.1.1 Valg av innsatsfaktorer

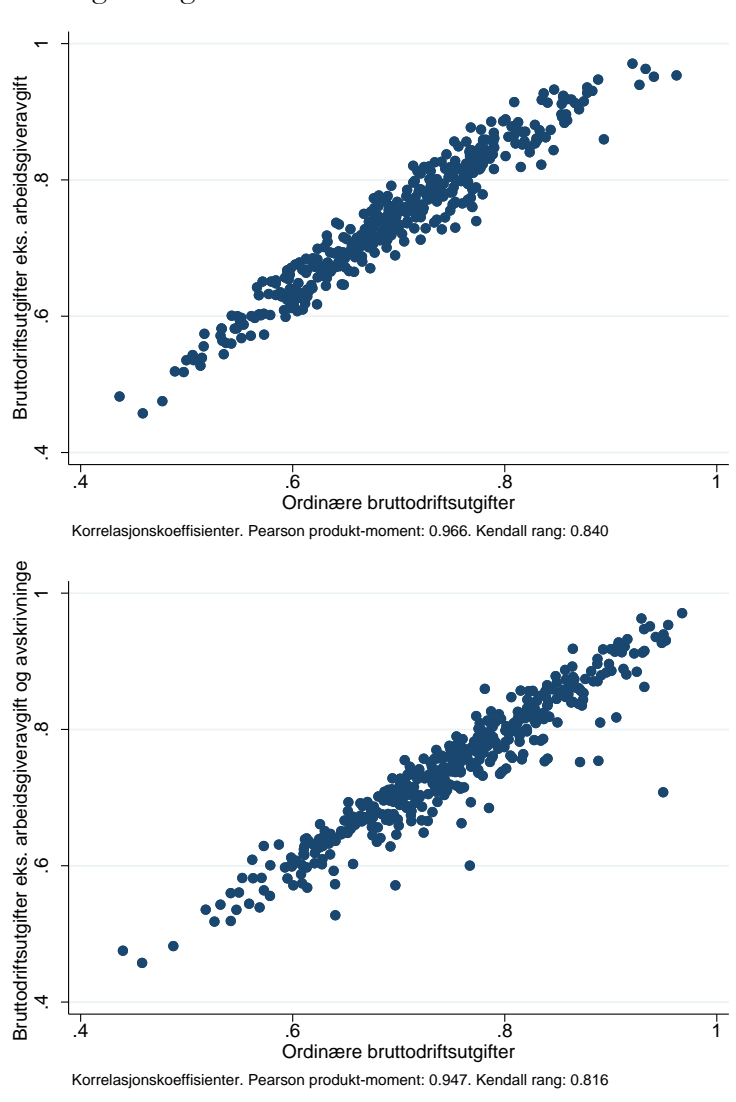
Aktuelle innsatsfaktorer for grunnskolen er ulike varianter av bruttodriftsutgifter og årsverk. Man kunne også tatt for seg arealtilfanget og realkapitalen, eller dekomponert bruttodriftsutgiftene, men det vil vi ikke gjøre her. Valget av innsatsfaktorer er diskutert nærmere i seksjon 3.2.1. Merk at SØF benytter bruttodriftsutgifter fratrukket arbeidsgiveravgift og avskrivninger som sitt mål for faktorinnsatsen. I tillegg deler SØF sin innsatsvariabel på et lokalt inverst sentralitetsmål, som er ment å reflektere ulempene med å være lokalisert i periferien. Vi ser imidlertid på denne korreksjonsvariabelen som en form for kontekstuell variabel og sparer derfor behandlingen av den til seksjon 4.1.5.

Vi starter med å ta for oss hvordan effektivitetsscorene påvirkes av valget av bruttodriftsutgifter som innsatsfaktor uten korreksjoner og bruttodriftsutgifter med fratrukk arbeidsgiveravgift og eventuelt avskrivninger, som illustrert i figur 4.1. Figuren illustrerer at fratrukk av arbeidsgiveravgift trekker effektivitetsfordelingen nedover, mens fratrukk av avskrivninger i tillegg gir en motstridende virkning. Utover dette minner frekvensfordelingene for de ulike effektivitetsscorene om hverandre i form, men det er også en del betydelige forskjeller.



Figur 4.1: Frekvensfordelinger av effektivitetsscore i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA med variabelt skalantbytte for 2018 og 2019. Faktorinnsats: a) Bruttodriftsutgifter, b) bruttodriftsutgifter fratrukket arbeidsgiveravgift og c) bruttodriftsutgifter mot bruttodriftsutgifter fratrukket arbeidsgiveravgift og avskrivninger. Merk at «aga.» står for arbeidsgiveravgift. Produkter: Elever på hvert hovedtrinn. Kontekstuelle variabler: Ingen

I figur 4.2 a) og b) framviser vi hvordan bruttodriftsutgifter korrelerer med sine to søstermål for faktorinnsatsen. Vi ser at fordelignene er relativt tett korrelerte, der korrelasjonen blir noe mindre om man trekker fra både avskrivninger og arbeidsgiveravgift enn når man bare trekker fra arbeidsgiveravgift.

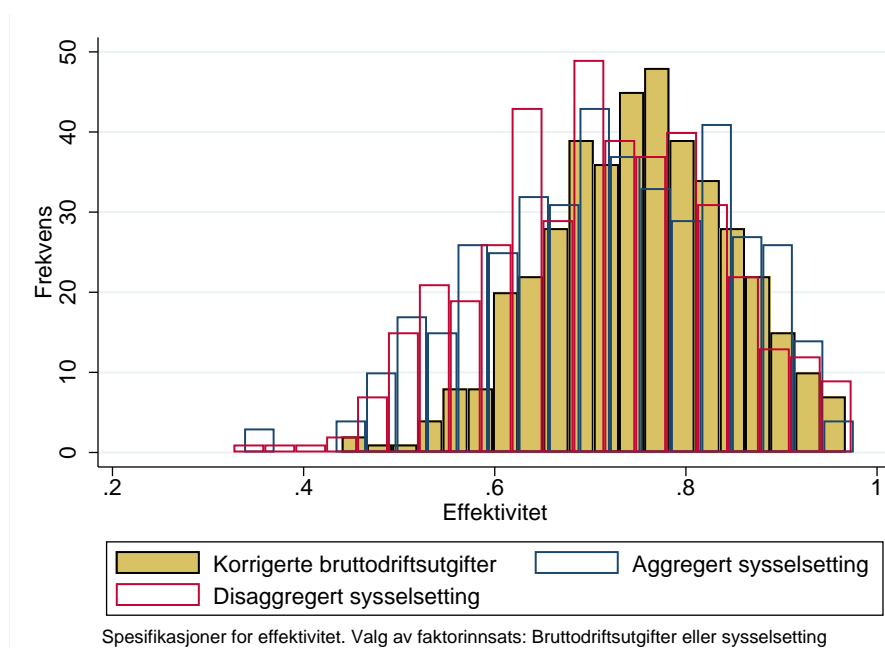


Figur 4.2: Korrelasjonen mellom effektivitetscoren i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA med variabelt skalaubytte for 2018 og 2019.

Bruttodriftsutgifter mot a) bruttodriftsutgifter fratrukket arbeidsgiveravgift og bruttodriftsutgifter fratrukket arbeidsgiveravgift og avskrivninger. Produkter: Elever hvert hovedtrinn. Kontekstuelle variabler: Ingen

Heretter vil vi referere til bruttodriftsutgifter fratrukket arbeidsgiveravgift og avskrivninger som korrigerede bruttodriftsutgifter og fokusere på dette målet for faktorinnsatsen.

Et alternativ til å bruke varianter av bruttodriftsutgifter som innsatsfaktor er som sagt å benytte sysselsetningsmål. I figur 4.3 har vi sammenliknet effektivitetsscorene ved bruk av korrigerede bruttodriftsutgifter som effektivitetsmål med to varianter av sysselsetningsmål, som begge har blitt benyttet tidligere. I den ene varianten har vi benyttet aggregert sysselsetting ved antall årsverk. I det andre tilfellet har vi skilt mellom lærerårsverk og assistentårsverk. Vi ser at frekvensfordelingene for effektivitetsscorene ved begge måtene å måle sysselsetting på er noe mer trukket utover enn ved når korrigerede bruttodriftsutgifter brukes. Forskjellene er like fullt moderate.



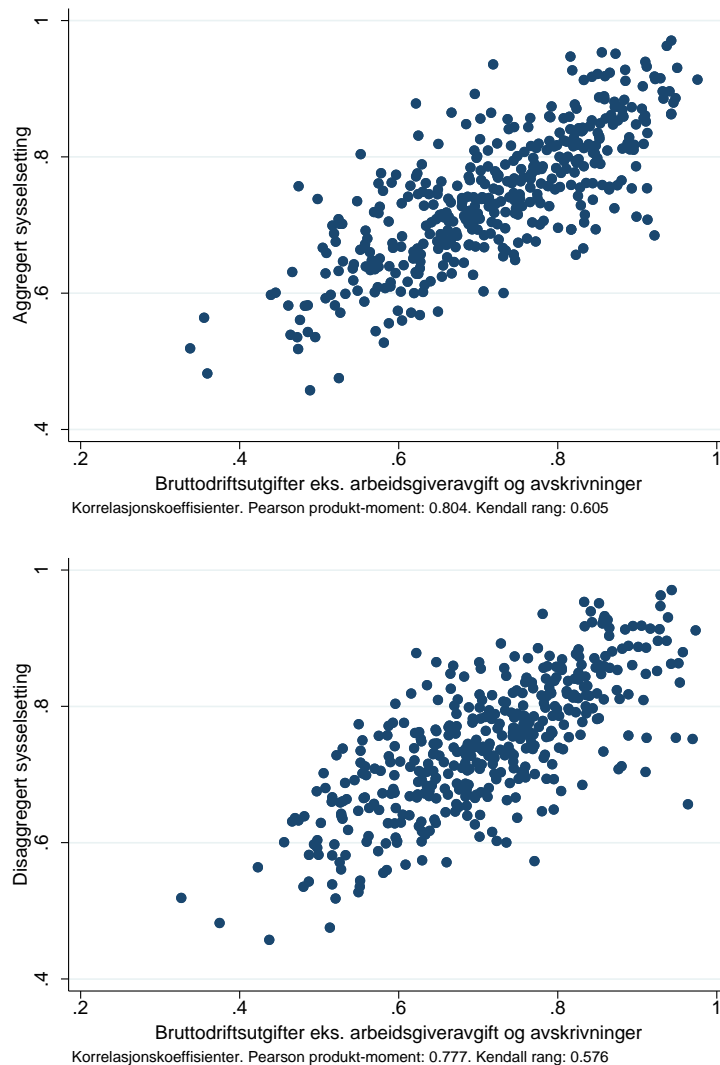
Figur 4.3: Frekvensfordelinger av effektivitetscore i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA med variabelt skalautbytte for 2018 og 2019.

Faktorinnsats: a) Bruttodriftsutgifter fratrukket arbeidsgiveravgift, b) samlede årsverk og c) lærerårsverk og assistentårsverk. Produkter: Elever på hvert hovedtrinn. Kontekstuelle variabler: Ingen

I figur 4.4 a) og b) har vi illustrert korrelasjonen mellom effektivitetscorene estimert ved korrigerte bruttodriftsutgifter og henholdsvis aggregert og disaggregert sysselsetting. Figuren illustrerer at samvariasjonen mellom effektivitetscorene estimert med utgangspunkt i bruttodriftsutgifter og sysselsetting som innsatsfaktorer er vesentlig forskjellige, selv om det fortsatt er betydelig korrelasjon.

I SØFs effektivitetsanalyser over tid for KMD og TBU har forskerne sammenliknet effektivitetscorene estimert ved ulike år, selv om ikke den samme estimeringsmodellen er benyttet i alle år. Dette henger sammen med at man ikke har ønsket å revidere effektivitetscorene for et gitt år etter at publiseringen av den offisielle rapporten om effektivitet i kommunal sektor for vedkommende år. Våre analyser illustrerer imidlertid at en slik framgangsmåte er beheftet med store feilkilder og utgjør en nokså omtrentlig måte å behandle produktivitetsutviklingen på, hvilket vi også har diskutert metodisk i seksjon 2.2.1. I tilfellet grunnskolen benyttet man korrigerte bruttodriftsutgifter som mål for faktorinnsatsen i 2019, disaggregert arbeidskraft som mål i 2018 og aggregert arbeidskraft som mål i 2015 til 2017. Plottene i figur 4.4 a) og b) viser at effektivitetscorene på ingen måte er upåvirket av disse endringene, som dermed også vil komme til uttrykk i SØFs måling av produktivitetsutviklingen over tid, hvilket vi kommer tilbake til i delkapittel 4.2.

Det er også noen forskjeller mellom effektivitetscorene estimert ved aggregert og disaggregert arbeidskraft som faktorinnsats, men de er ikke så store. Merk at valg av disaggregert arbeidskraft kontra aggregert arbeidskraft som mål for faktorinnsatsen innebærer at man tar hensyn til heterogeniteten i arbeidskraften, men en videre overgang til korrigerte bruttodriftsutgifter som effektivitetsmål innebærer at man også tar hensyn til andre innsatsfaktorer. Vi referer igjen til seksjon 3.2.1 for detaljer.



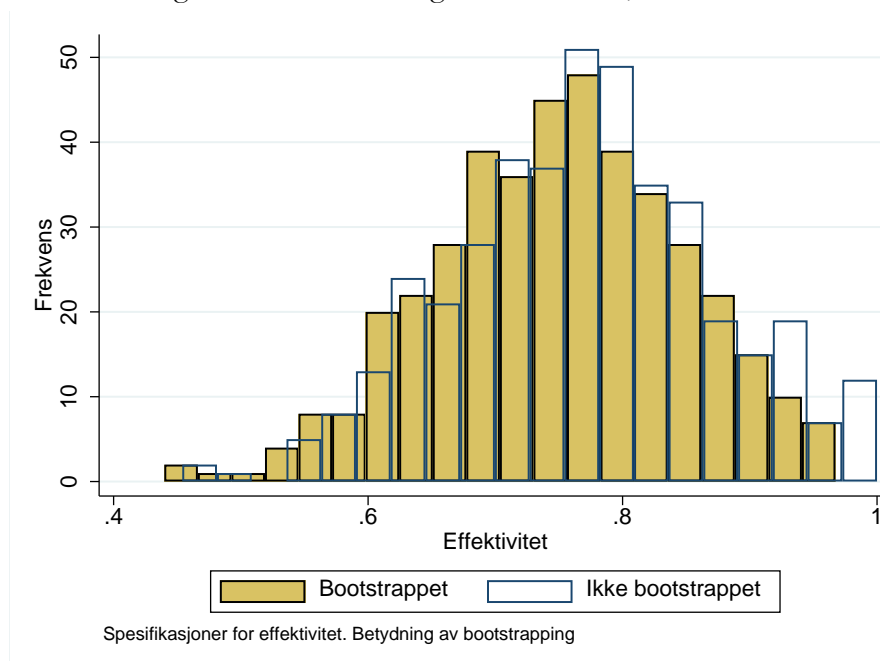
Figur 4.4: Korrelasjonen mellom effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA med variabelt skalaubytte for 2018 og 2019. Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgiveravgift og avskrivninger mot a) aggregert sysselsetting og b) lærerårsverk og assistentårsverk. Produkter: Elever på hvert hovedtrinn. Kontekstuelle variabler: Ingen

4.1.2 Håndtering av seleksjon

Som vi kort var innom i seksjon 3.1.1 og sporadisk gjennom delkapittel 2 innebærer DEA-metoden behov for håndtering av seleksjon både forbundet med urealiserte potensielle observasjoner (utvalgsskjevhet) og forbundet med uteliggere forårsaket av målefeil. Mens utfordringen med ikke-observerte utfall av teknologien gjør seg særlig gjeldende ved små utvalg, er utfordringen med uteliggere generelt større ved store utvalg.

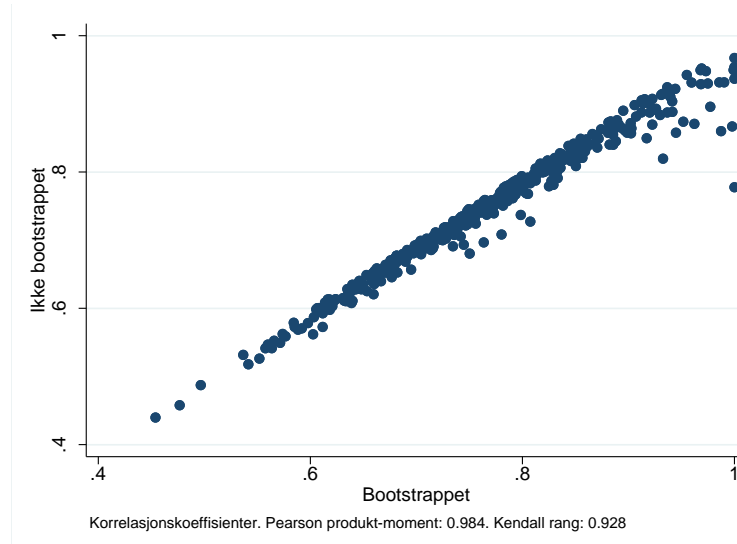
Når det gjelder urealiserte potensielle observasjoner, går problemstillingen på at de observerte produksjonsutfallene for gitt faktorinnsats ikke nødvendigvis reflekter teknologifronten, altså de mest effektive produksjonsutfallene innenfor dagens teknologi. Løsningen på dette er å anslå de ikke-realiserede observasjonene ved å benytte bootstrapping. Dermed forflyttes også teknologifronten utover, som implisitt innebærer effektivitetsscorene reduseres i forhold til når seleksjonsskjevheten ikke håndteres. Som nevnt i innledningen til delkapittel 2.2 begynte SØF med bootstrapping i sine DEA-analyser av offentlig sektor i sine 2019-analyser, mens tidligere analyser ikke er bootstrappet. Igjen kan dette gi utfordringer når effektivitetsscorene sammenliknes over tid.

I figur 4.6 har vi sammenliknet effektivitetsscorene beregnet med og uten bootstrapping. Vi ser at fordelingen forflytter seg mot høyre når man ikke bootstrapper, i tråd med hva man kan forvente. Fordelingene minner for øvrig om hverandre, men er ikke like.



Figur 4.5: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet og ikke bootstrappet DEA med variabelt skalautbytte for 2018 og 2019. Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger
Produkter: Elever på hvert hovedtrinn. Kontekstuelle variabler: Ingen

I figur 4.6 ser vi at effektivitetsscorene mellom estimert ved DEA med og uten bootstrapping er relativt tett korrelerte. Det er likevel vesentlige forskjeller i enkelte tilfeller, så fordelingen vil ikke være lik.



Figur 4.6: Korrelasjonen mellom effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet og ikke bootstrappet DEA med variabelt skalautbytte for 2018 og 2019. Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger
Produkter: Elever på hvert hovedtrinn. Kontekstuelle variabler: Ingen

Når det gjelder håndtering av uteliggere, referer vi til vår korte diskusjon i seksjon 3.1.1 og tilknyttede referanser. Vi vil ikke her rapportere noen eksplisitte uteliggeranalyser eller korreksjoner i utvalget, annet enn at vi ekskludert kommuner med manglende variabelutfall og ekskludere kommuner som ikke har elever på alle tre hovedtrinn (det vil si småtrinnet, mellomtrinnet og ungdomstrinnet). I tillegg har vi behandlet kommuner som har blitt slått sammen eller oppsplittet som separate enheter før og etter sammenslåingen eller oppsplittingen.

Behovet for uteliggeranalyser har likevel vært et tema i våre empiriske analyser tilknyttet dette prosjektet. I våre initale analyser fikk vi nemlig mye lavere effektivitetsscorer enn SØF for grunnskolen. Ved hjelp av nettopp SØF fikk vi avdekket at dette skyldes noen mindre manuelle endringer i Utdanningsdirektoratets utdanningsstatistikk i forbindelse med endring av kommunegrenser og især oppdeling av kommunene Snillefjord og Tysfjord. Ifølge våre informanter ved SØF hadde de liknende utfordringer ved gjennomføringen av sin siste effektivitetsanalyse for grunnskolen.

Dette illustrerer at DEAs sårbarhet for uteliggere og behovet for uteliggeranalyser når denne metodikken benyttes, idet DEA er en deterministisk estimeringsprosedyre som dermed ikke har et stokastisk restledd å håndtere uteliggere med. Effektivitetsscorene ble ikke tilsvarende redusert når vi brukte SFA og StoNED som estimeringsprosedyrer, da disse fanget opp støyen i dataene i sine stokastiske restledd.

4.1.3 Valg av skalautbytte

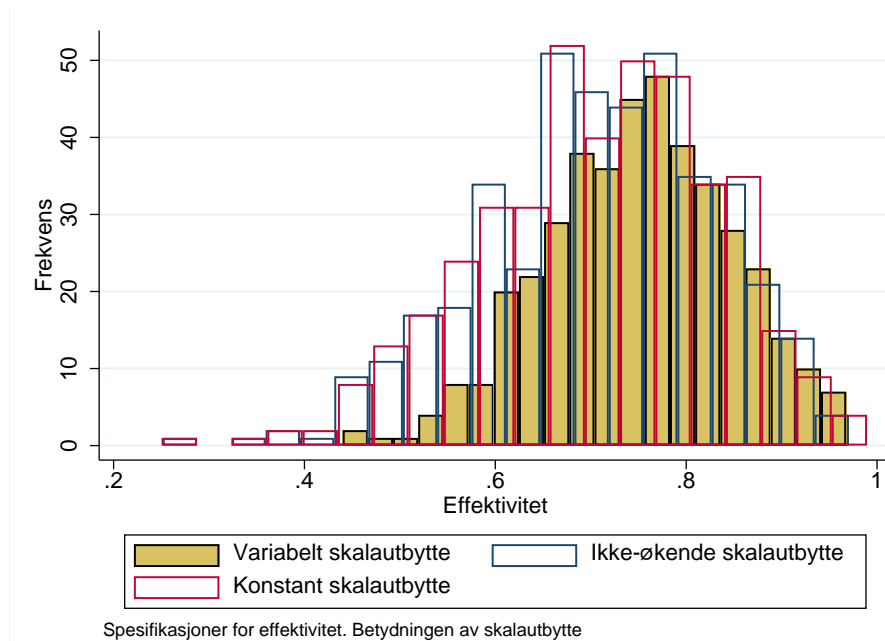
Valg av skalautbytte kan knyttes til hva slags skalaegenskaper man forutsetter at teknologien har, men de relaterte effektivitetsscorene kan også tolkes å svare til ulike effektivitetskonsepter. I vår analyse vil vi holde oss til den sistnevnte tolkningen om alternativ effektivitetskonsepter.

En spesifisering med variabelt skalautbytte for produksjonsfronten kan forstås som den faktiske teknologifronten som gjenspeiles av observasjonene, der de tilhørende effektivitetsscorene refereres til som innsatsfaktorsparende teknisk effektivitet. En spesifisering med konstant skalautbytte over produksjonsfronten tar derimot hensyn til skalaeffektivitet i tillegg til teknisk effektivitet og måler dermed teknisk produktivitet, jfr. figur 2.1. Ikke-økende skalaubytte blir en mellomting av de to forutsetningene, der man tillater fallende skalautbytte, men ikke økende skalautbytte. I SØFs analyser har man holdt seg til variabelt skalaubytte, da man primært har vært interessert i effektivitet og ikke i produktivitet eller skalaeffektivitet.

Merk at man i mange mikroøkonomiske anvendelser antar økende skalaubytte når produksjon går fra null til positiv, før den gradvis avtar og går over til fallende skalautbytte etter hvert som produksjonen øker. En slik teknologi innebærer at ved variabelt skalautbytte vil produksjonsfronten ha økende skalautbytte for små produksjonsvolum, konstant skalautbytte for medium produksjonsvolum (i deler eller punkt av produksjonsfronten) og fallende skalautbytte for store produksjonsvolum. En teknologi med ikke-økende skalautbytte vil være tilsvarende, foruten at man tvinger produksjonsfronten til å ha konstant skalautbytte istedenfor økende skalautbytte, i vårt eksempel for små produksjonsvolum.

I prinsippet skal skalaeffektiviteten implisitt framgå som forholdet mellom effektivitetsscorene ved variabelt og konstant skalautbytte (altså forholdet mellom teknisk effektivitet og teknisk produktivitet). Samtidig skal produksjonsfronten ved variabelt skalaubytte aldri ligge innenfor produksjonsfronten ved konstant skalautbytte. I praksis har vi imidlertid estimert produksjonsfrontene hver for seg der bootstrappingen ikke er harmonisert, slik at disse kriteriene ikke fullt ut vil være oppfylt i vårt eksempel.

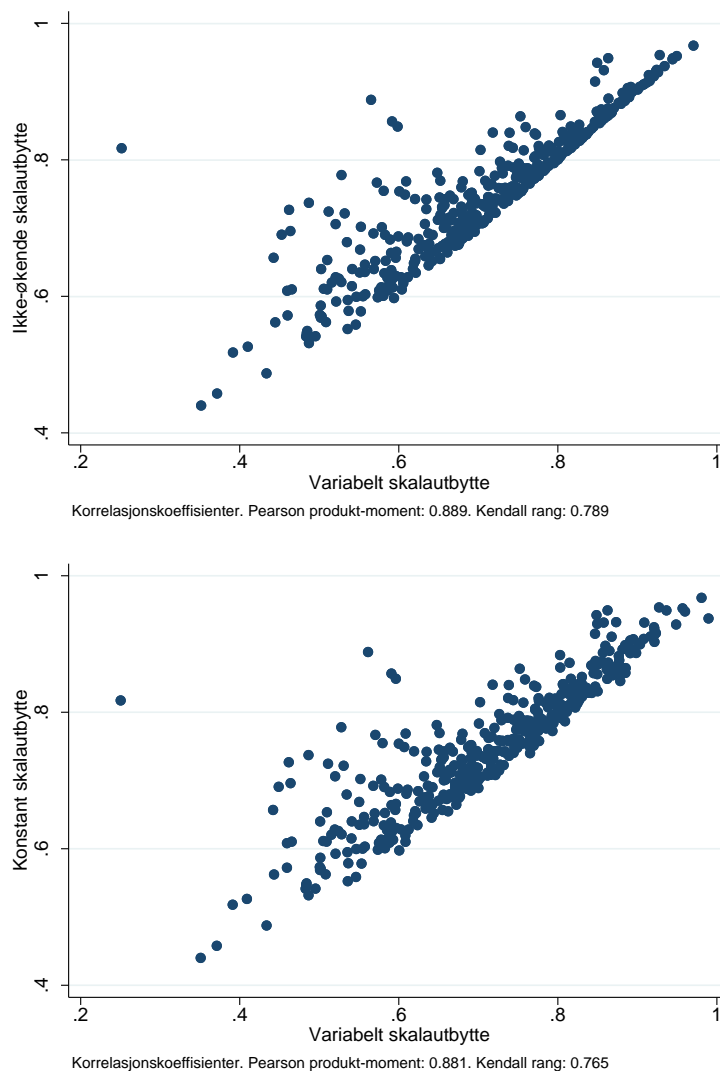
I figur 4.7 har vi illustrert frekvensfordelingene til effektivitetsscorene ved ulike forutsetninger om skalautbyttet. Vi ser at effektivitetsscorene generelt er lavere for de andre skalaforutsetningene enn variabelt skalautbytte. Riktignok er effektivitetsscorene estimert under konstant skalautbytte tilsynelatende de høyeste i enkelte tilfeller, hvilket skyldes at estimeringene av de respektive frontene er foretatt uavhengig av hverandre. I realiteten og ved simultan estimering ville hver observasjonsenhet ikke kunne ha høyere effektivitetscore ved konstant skalautbytte (tilsvarende såkalt teknisk produktivitet) enn ved variabelt skalautbytte (tilsvarer såkalt teknisk effektivitet). Dette skyldes at teknisk produktivitet tilsvarer produktet mellom faktoreffektivitet og skalaeffektivitet, der skalaeffektiviteten (som de andre effektivitetsscorene) per definisjon ligger mellom null og en.



Figur 4.7: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA med alternative antakelser om skalautbytte for 2018 og 2019. Faktorinnsats: Bruttodriftsutgifter mot bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger. Produkter: Elever på hvert hovedtrinn. Kontekstuelle variabler: Ingen

I figur 4.8 a) og b) under har vi illustrert samvariasjonen mellom effektivitetsscoren estimert under variabelt skalautbytte og effektivitetsscorene estimert under henholdsvis ikke-økende og konstant skalautbytte. Det er sterk korrelasjon for effektive enheter under variabelt skalautbytte, men korrelasjonene mellom effektivitetsscorene er svakere for enheter som er ineffektive under variabelt skalautbytte. Som man kan forvente er korrelasjonen mellom effektivitetsscorene noe svakere under konstant skalautbytte enn under ikke-økende skalautbytte, men forskjellene er små. Dette tyder på at produksjonsfronten med variabelt skalautbytte primært skiller seg fra produksjonsfronten med konstant skalautbytte ved at den tar hensyn til partier av teknologifronten med økende skalautbytte og ikke ved at den tar hensyn til partier av teknologifronten med avtakende skalautbytte.

Våre funn av empiriske forskjeller mellom teknisk effektivitet og produktivitet viser at SØFs DEA-metodikk kan utvides til å dekomponere kommuneeffektiviteten inn i teknisk effektivitet og skalaeffektivitet. Skalamålet kan gi nyttig tilleggsinformasjon om sammenhengen mellom kommunenes størrelse og deres produktivitet.



Figur 4.8: Korrelasjonen mellom effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA for 2018 og 2019.

Variabelt skalautbytte mot a) ikke-økende skalautbytte og b) konstant skalautbytte

Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger

Produkter: Elever på hvert hovedtrinn. Kontekstuelle variabler: Ingen

4.1.4 Håndtering av tjenestekvalitet

En annet viktig aspekt ved effektivitetsmåling er hvordan man håndterer kvalitetsdimensjonen. Effektivitetsmålingen for grunnskolen ligger relativt langt framme på dette feltet sammenliknet med andre delsektorer innen kommunal sektor med flere gode kvalitetsindikatorer inkludert skolepoeng og skolemiljø, jamfør seksjon 3.2.4.

Vi vil i det følgende i stor grad følge SØF og inkludere mål på tjenestekvalitet som produkter i DEA-modellen. Vi bemerker at dette ikke er i tråd med vår hovedanbefaling fra kapittel 2.1.1, hvor vi problematiserer at variabler med spesielle måleenheter i utgangspunktet ikke er egnet for standard DEA. Vi anbefaler i dette kapitlet å håndtere slike variabler i en ettermodell fram til ev. ny metodikk som er bedre tilpasset til at slike variabler tas i bruk. I de påfølgende analysene velger vi allikevel å behandle både skolepoeng og skolemiljø som produkter for å sikre sammenliknbarhet mot SØFs beregninger.

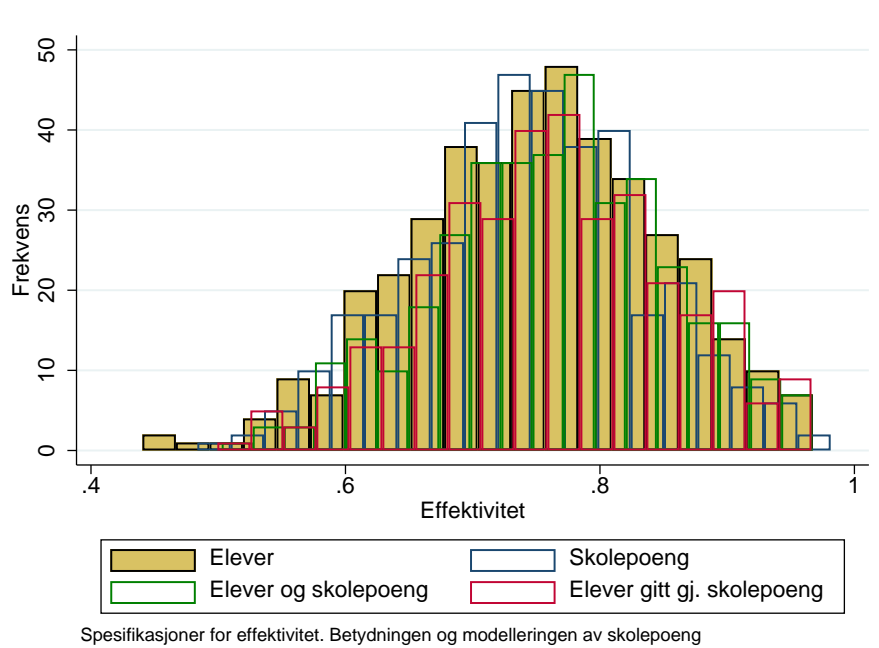
Den fremste kvalitetetsindikatoren for skolesektoren er skolepoeng eller avarter av disse variablene, som gjenspeiler skoleprestasjoner ved avgangen fra hvert av grunnskolens tre hovedtrinn. I Utdanningsdirektoratets og Statistisk sentralbyrås statistikker utregnes skolepoeng på bakgrunn av nasjonale prøver i norsk, engelsk og matematikk.

I en DEA modell kan skolepoeng modelleres som enten et produkt eller en kontekstuell variabel. Skolepoeng er en del av skolens produksjon, hvilket taler for at bør modelleres som et produkt. Det krever imidlertid at skolepoengene bør tilfredsstille DEAs fortsettelser om at dataene som inngår er volumer, hvilket tilsier at man bør operere med samlede skolepoeng per hovedtrinn og ikke gjennomsnittlige skolepoeng per hovedtrinn. Dersom man både inkluderer samlede skolepoeng og elever som produkter tar modellen hensyn til ulike grad av substituerbarhet og komplementaritet mellom kvalitet og volum. Alternativt kan man inkludere skolepoeng per hovedtrinn som kontekstuelle variabler. For å unngå mulikollinearitet med produktene kan man da velge å benytte gjennomsnittlige skolepoeng framfor samlede skolepoeng. Vi referer til delkapittel 2.1.3 for nærmere diskusjon av disse forholdene.

I figur 4.9 har vi illustrert effektivitetsscorene for grunnskolen estimert ved DEA ved tre ulike håndteringer av kvalitetsaspektet sett opp mot en spesifisering der kun elever per hovedtrinn er benyttet som produkter. De tre spesifiseringene er samlede skolepoeng per hovedtrinn alene som produkt, både elever og samlede skolepoeng per hovedtrinn som produkt, og kombinasjonen av elever per hovedtrinn som produkt og gjennomsnittlige skolepoeng per hovedtrinn som kontekstuelle variabler. Som indikert i begynnelsen av dette delkapittelet benytter vi Simar og Wilson (2007) sin prosedyre for estimering av effekter av kontekstuelle variabler i en to-steps-DEA.

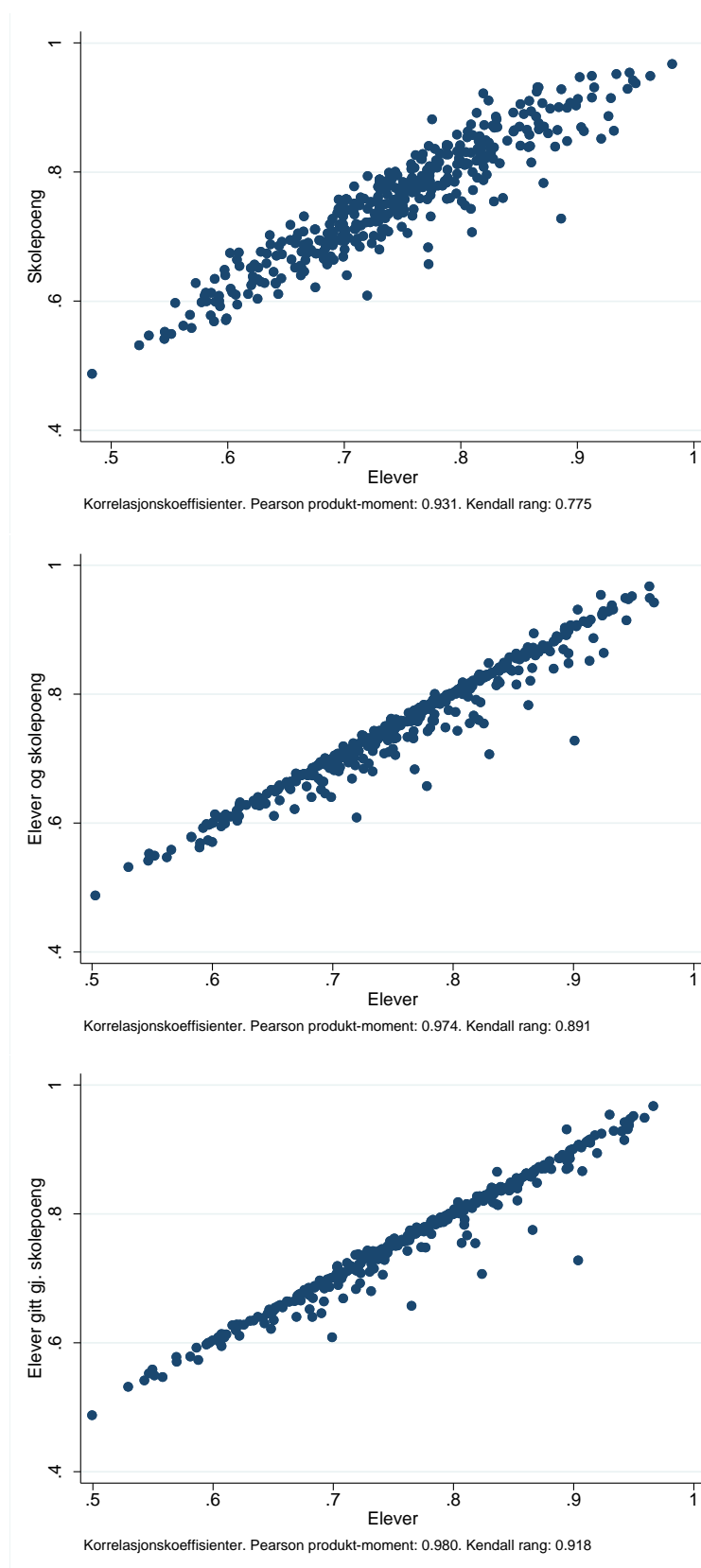
Våre resultater er vist i figur 4.9 viser noen forskyvninger i fordelingen når man tar hensyn til kvaliteten på læringen. Selv om forskyvningene er merkbare, er de ikke radikale. Forskjellen er størst når skolepoengene benyttes som de eneste produktene og man ikke tar hensyn til samspillet mellom volum og kvalitet. Forskjellene i fordelingene av effektivitetsscorene er relativt små mellom spesifiseringen der både elever og skolepoeng er benyttet som produkter og spesifiseringen der elever er benyttet som produkt og gjennomsnittlige skolepoeng er benyttet som kontekstuelle variabler.

Merk at i de påfølgende figurene angir figurteksten variabelnavn dersom de behandles som produkter i modellspesifiseringen (eks. «elever og skolepoeng» innebærer at variabler for antall elever og totale (faktiske) skolepoeng brukes som produkter), mens ordet «gitt» betyr at påfølgende variabler behandles i en regresjonsmodell (eks. «elever gitt gj. Skolepoeng» innebærer at elevtallene inngår som produkter i DEA-modellen mens gjennomsnittlige skolepoeng inngår som kontekstuelle variabler i en regresjonsmodell).



Figur 4.9: Frekvensfordelinger av effektivitetscore i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalautbytte for 2018 og 2019.
 Faktorinnsats: Brutto driftsutgifter mot brutto driftsutgifter fratrukket arbeidsgivergjelt og avskrivninger
 Produkter: a) og d) Elever på hvert hovedtrinn, b) samlede skolepoeng på hvert hovedtrinn og c) Elever og samlede skolepoeng på hvert hovedtrinn
 Kontekstuelle variabler: a) til c) Ingen, d) gjennomsnittlig skolepoeng på hvert hovedtrinn

I figur 4.10 viser hvordan effektivitetsmålingene ved DEA med ulike behandlingene av skolepoeng korrelerer med effektivitetsmålingen ved DEA, der elever utgjør eneste produkt og kvaliteten på produksjonen ikke adresseres.



Figur 4.10: Korrelasjonen mellom effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA for 2018 og 2019.

Faktorinnsats: Brutto driftsutgifter fratrukket arbeidsgivergift og avskrivninger

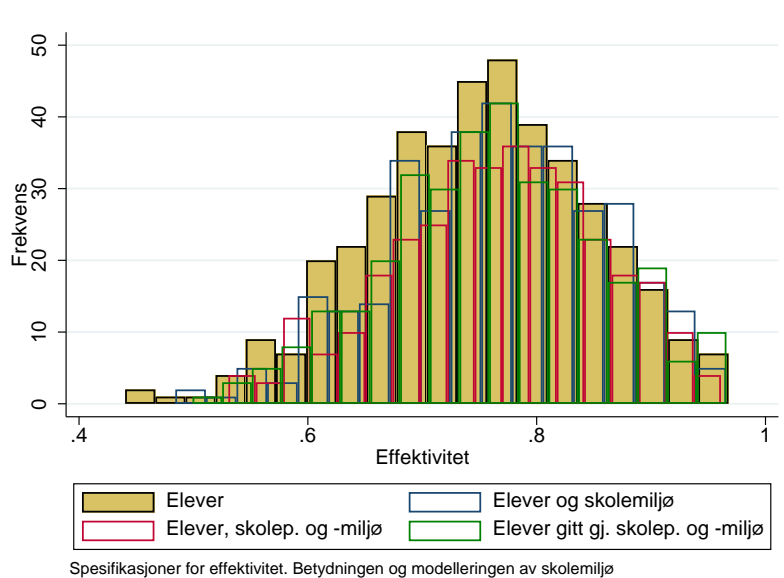
Produkter: Elever på hvert hovedtrinn mot a) samlede skolepoeng, b) elever og samlede skolepoeng (i midten) og c) elever. Kontekstuelle variabler: Ingen mot a) og b) ingen og c) gjennomsnittlig skolepoeng

Figuren bekrefter at effektivitetsscorene stort sett ikke endrer seg radikalt når man tar hensyn til skolepoeng i estimeringen. Forskjellen blir størst når man benytter skolepoeng på hvert hovedtrinn som produkter alene i estimeringen, uten noen form for kontekstuell behandling.

De fleste vil være enige om at grunnskolens mål ikke bare gode skoleprestasjoner, men også sosial læring, skaping av relasjoner og trivsel. SØF benytter derfor også en indikator for skolemiljø i sine regresjoner, noe vi også vil ta inn i det følgende.

Skolemiljø utgjør en komposittindeks basert på spørsmål fra elevundersøkelsen for avgangselever ved mellomtrinnet og ungdomstrinnet. Både spørsmålene og indeksens struktur har endret seg litt over tid. Variabelen vi benytter tilsvarer den SØF har benyttet i sine analyser for 2018 og 2019. Den bygger på tre likt vektete delfaktorer; elvenes medbestemmelse, mobbing og trivsel. Skolemiljøvariabler beregnes både for mellomtrinnet og ungdomstrinnet, og aggregeres deretter opp, vektet etter antall elever på hvert av de to hovedtrinnene. Den gjennomsnittlige skolemiljøvariabelen vi her benytter er normalisert, slik at 100 tilsvarer landsgjennomsnittet, tilsvarende det SØF omtaler som skolemiljøindikator. Vi referer til seksjon 2.1.5 og 3.2.4 for tilknyttede diskusjoner.

I våre regresjoner har vi valgt å modellere skolemiljø på tre ulike måter og sammelikne dem med spesifikasjonen, der antall elever hvert hovedtrinn inngår som de eneste produktene. I den første varianten der skolemiljø inngår, benytter vi skolemiljø multiplisert med antall elever som produkt i tillegg til antall elever. Igjen er multipliseringen med antall elever foretatt for å sikre at variabelen ivaretar DEA-metodens krav til bruk av volumtall. Det bør presiseres her at tolkningen av produktene som volumtall for skolemiljø kan kritiseres for å være mer tvilsom enn summen av skolepoeng (karakterer). I den andre varianten der skolemiljø inngår, benytter vi i tillegg skolepoeng multiplisert med antall elever på hvert hovedtrinn som produkter. I den tredje varianten holder vi oss til kun elever på hvert hovedtrinn som produkter, men modellerer heller skolepoeng på hvert hovedtrinn og skolemiljø som kontekstuelle variabler, målt ved gjennomsnittlige verdier.



Figur 4.11: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalanthytte for 2018 og 2019.

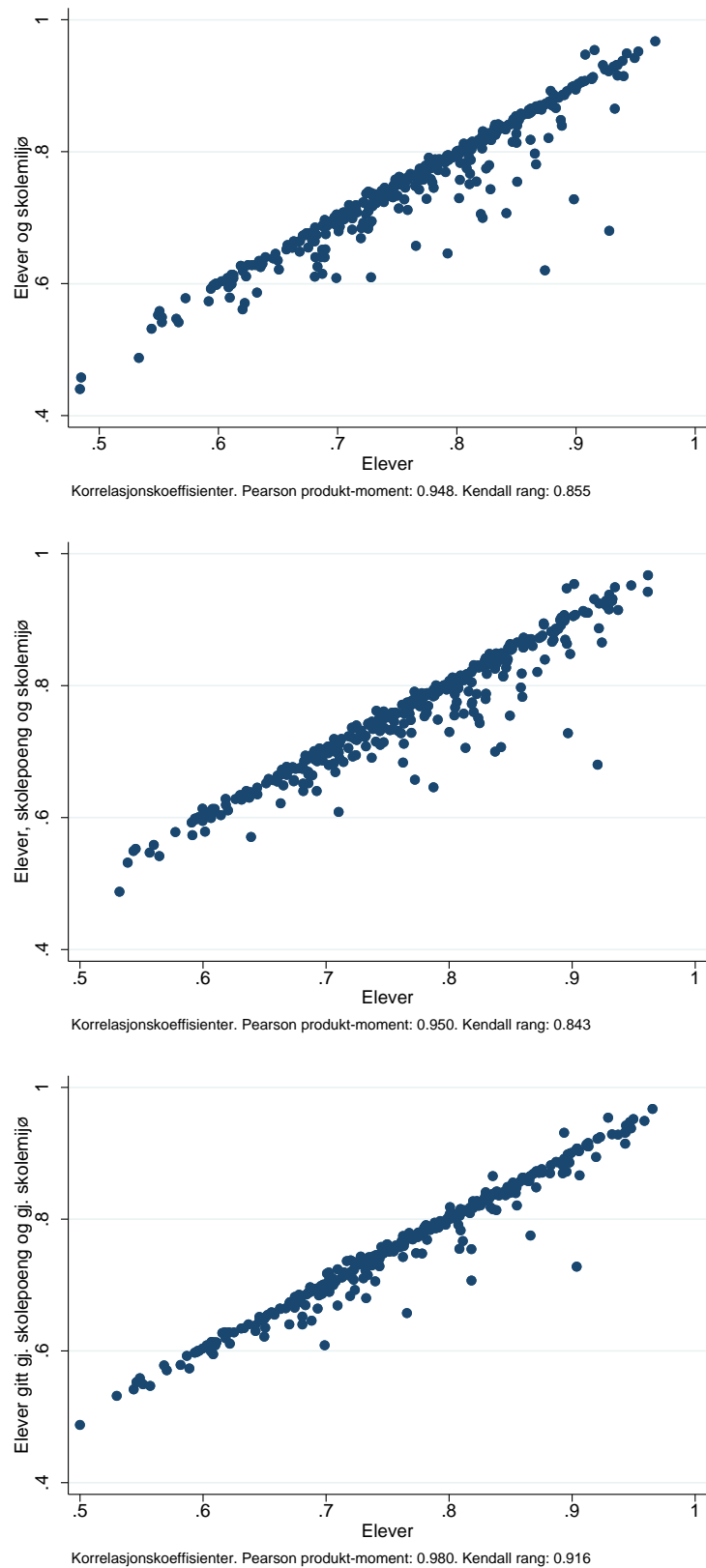
Faktorinnsats: Bruttodriftsutgifter mot bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger

Produkter: a) og d) Elever på hvert hovedtrinn, b) elever på hvert hovedtrinn, skolepoeng på hvert hovedtrinn og skolemiljø og c) elever hvert hovedtrinn og skolemiljø

Kontekstuelle variabler: a) til c) Ingen, d) gjennomsnittlig skolepoeng på hvert hovedtrinn og skolemiljø

Frekvensfordelinger av effektivitetsscorer tilhørende de ulike måtene å behandle skolemiljø er illustrert i figur 4.11. Igjen er det store likhetstrekk. De største forskjellene finner vi mellom modellen med elever som eneste produkter uten kontekstuell behandling og modellen med elever, skolepoeng og skolemiljø som produkter. Vi ser at effektivitetsfordelingen tenderer til å bli jevnere når vi inkluderer flere produkter.

Korrelasjonen mellom effektivitetsscorene tilhørende modellen uten skolemiljø som forklaringsvariabel og de ulike modellene som modellerer skolemiljø framgår av figur 4.12. Jevnt over er korrelasjonene sterke, men vi ser noen avviker der kommuner gjør det vesentlig dårligere relativt sett i modellene der skolemiljø tas inn. Korrelasjonen mellom modellene uten og med skolemiljø er sterkere når variabelen behandles som en kontekstuell variabel, enn når den behandles som et produkt.



Figur 4.12: Korrelasjonen mellom effektivitetscoren i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA for 2018 og 2019.

Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger

Produkter: Elever på hvert hovedtrinn mot a) elever og samlede skolepoeng og skolemiljø og c) elever og skolemiljø og d) elever. Kontekstuelle variabler: Ingen mot a) og b) ingen og c) gjennomsnittlig skolepoeng og skolemiljø

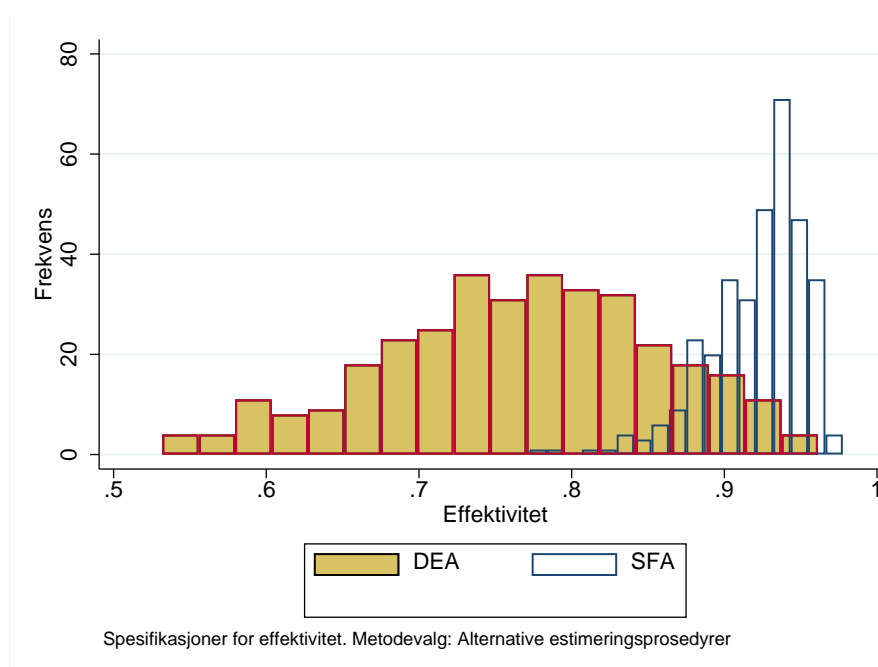
4.1.5 Valg av metode

Vi vil nå ta for oss alternative estimeringsmetodikk til DEA, nærmere bestemt stokastisk frontanalyse og stokastisk ikke-parametriske omhylling av data. Mens DEA åpner for ikke-parametriske produksjonsteknologi og tolker regresjonens uforklarte forskjeller som effektivitetsforskjeller, forsetter SFA en parametriske produksjonsteknologi og tillater stokastisk støy. StoNED er en generalisering av både DEA og SFA, som både tillater ikke-parametriske produksjonsteknologi og stokastisk støy. Merk at produktene og faktorinnsatsen ved SFA er transformert til naturlige logaritmer. Vi referer til delkapittel 2.5 for nærmere teoretiske redegjørelser for SFA, StoNED og alternative estimeringsprosedyrer.

I figur 4.13 har vi belyst hvordan frekvensfordelingen påvirkes av valg av estimeringsprosedyre. I samtlige beregninger er bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger benyttet som faktorinnsats, mens elever og samlede skolepoeng på hvert hovedtrinn og skolemiljø er benyttet som produkter.

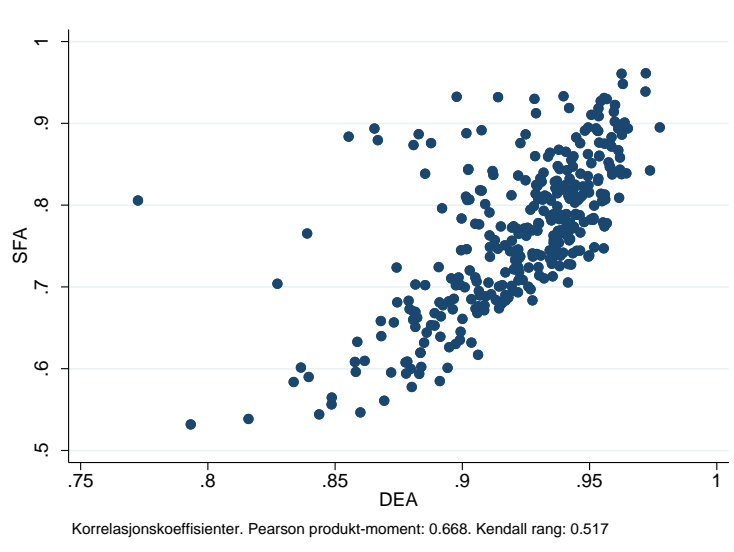
Av figuren er det klart at SFA involverer en frekvensfordeling med langt mer effektive enheter enn DEA. Dette skyldes trolig at SFA tolker en del av variasjonen mellom kommunene som stokastisk støy heller enn ineffektivitet: SFA inneholder et todelt restledd hvor effektivitet kun er en av komponentene, mens DEA tolker alle avvik fra fronten som ineffektivitet. Samtidig bør det påpekes at DEA estimerer fronten ikke-parametriske mens SFA estimerer den parametriske, noe som også kan spille inn på forskjellene mellom metodene.

Videre ser vi at effektivitetsfordelingen ved SFA involverer relativt mange kommuner med høy effektivitet i grunnskolen, med en hale av mer ineffektive kommuner. Dette står i motsetning til effektivitetsfordelingen tilhørende DEA, som i større grad er konsentrert rundt medianeffektiviteten med en jevnere effektivitetsfordeling på begge sider av effektivitetsskalaen.



Figur 4.13: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og a) bootstrappet DEA og b) SFA for 2018 og 2019 med variabelt skalautbytte. Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger. Produkter: Elever og samlede skolepoeng på hvert hovedtrinn og skolemiljø. Kontekstuelle variabler: Ingen

Det er også interessant å se hvordan rangeringene av kommunens effektivitet avhenger av valg av estimeringsmetode. I figur 4.14 under har vi illustrert hvordan effektivitetsscorene beregnet ved DEA korrelerer med effektivitetsscorene beregnet med SFA. Det overordnede bildet er at det er vesentlig korrelasjon, men likevel er større forskjeller enn de som fulgte av de metodiske modifikasjonene av DEA-modellen som vi har tatt for oss så langt. En relevant anbefaling vil være å benytte flere metoder (eks. DEA og SFA) til å måle kommunenes effektivitet. Alle de metodene vi har tatt for oss i dette kapitlet har sine styrker og svakheter, og en bruk av flere metoder vil derfor være en nyttig robusthetssjekk av resultatene.



Figur 4.14: Korrelasjonen mellom effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA for 2018 og 2019 og med SFA. Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgiveravgift og avskrivninger. Produkter: Elever og samlede skolepoeng på hvert hovedtrinn og skolemiljø. Kontekstuelle variabler: Ingen

Vi har også forsøkt StoNED-metoden til å beregne kommuneeffektivitet. En utfordring har vært at det ikke-lineære optimeringsproblemet som ligger til grunn for metoden blir krevende å estimere når utvaglet blir stort, som i tilfellet med kommunene. Samtidig er variablene i problemet bestående av ulike målestokker, noe som har forhindret estimeringsverktøyet GAMS i å finne en løsning på optimeringsproblemet som inngår som steg 1 i løsningsalgoritmen for StoNED. Dette er problemer som kan omgås ved reskalering av variabler og/eller valg av initialverdier i optimeringen, men vi har ikke prioritert en videre utforskning innenfor rammen av dette prosjektet.

4.1.6 Håndtering av kontekst

Vi vil nå rette søkelys mot hva omstendighetene betyr for effektivitetsmålingene. SØF tar høyde for flere kontekstuelle forhold i sine estimeringsmodeller for KMD og TBU. Til dels er deres håndtering av kontekstuelle variabler det vi vil kalle ukonvensjonelle. I de fleste tilfeller vil likevel trolig det mest vesentlige være at de tar høyde for kontekstuelle variabler. Vi starter med å ta for oss den såkalte sonevariabelen. Dette er en form for et lokalt preferimål, som fanger opp avstander til fra en gitt sone til omkringliggende soner. Sonevariabelen til en institusjon eller kommune øker med hvor perifer beliggenhet den har (jmfør seksjon 2.1.6 for variabelens definisjon). SØF har valgt å dele bruttodriftsutgiftene fratrukket arbeidsgiveravgift og avskrivninger på denne variabelen. Dette målet benyttes i

forbindelse med det kommunale inntektsfordelingssystemet, men vi finner ikke noe klar begrunnelse for hvorfor man skulle gjøre dette i effektivitetsmålingen i kommunal sektor.

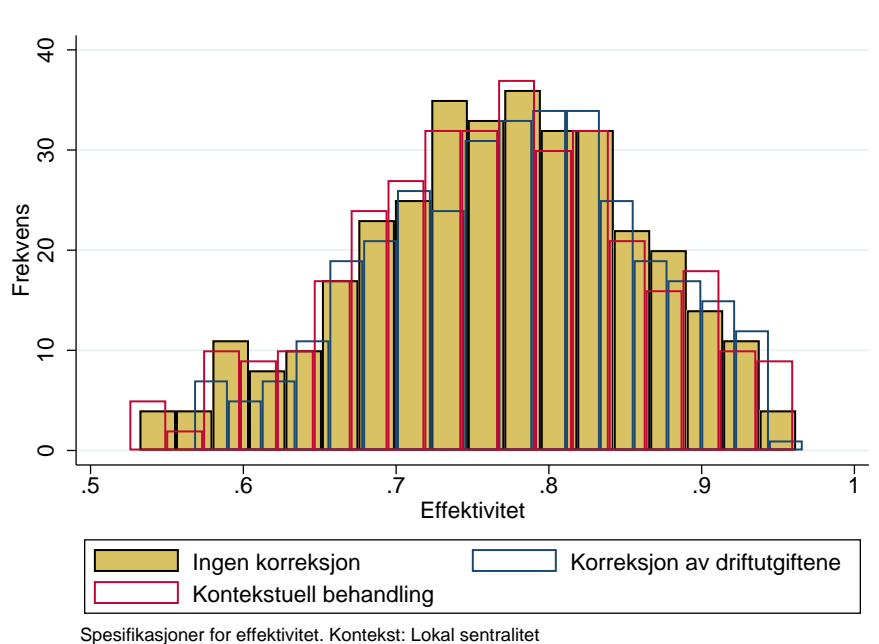
Mens geografiske avstander kan spille en rolle for produksjonen for vareproduksjon og mobil tjenesteyting, er det uklart hvorfor man skulle legge det til grunn for stedbundne tjenester som kommunale grunnskoler er et eksempel på. Man kunne argumentere for at grunnskoler i perifere strøk kan ha en rekrutteringsulempe. På en annen side kan man innvende at grunnskolen i perifere strøk har tilgang på billigere arealer. Ved slike begrunnelser er det likevel grunn til å stille spørsmål både ved utformingen av sonevariabelen, og ved hvordan den håndteres i SØFs estimeringsmodell.

Til tross for at man kan komme med begrunnelser på hvorfor sonevariabelen kan fange opp bakenforliggende omstendigheter finner vi behandlingen til SØF, TBU og KMD av variabelen som uheldig og problematisk. Uten en klar teoretisk begrunnelse for hvorfor man har valgt å dividere på sonevariabelen, fremstår dette mer som en forhåndskorreksjon av dataene, som bidrar til å kommuner i perifere strøk kommer kunstig bedre ut på bekostning av kommuner i sentrale strøk. Ved effektivitetsforskjeller mellom sentrale og perifere strøk blir det uklart hvorfor disse til dels skal tilskrives sentralitetsforskjeller.

Dersom man tar sonevariabelen for god fisk og regner den som en kontekstuell variabel, vil det likevel være ukonvensjonelt å dividere kostnadsvariabelen på variabelen. Den konvensjonelle måten å behandle den på vil derimot være å modellere den som nettopp en kontekstuell variabel. Vi referer til seksjon 2.1.6 og seksjon 3.2.1 for ytterligere diskusjoner av de metodiske og datamessige aspektene knyttet til denne variabelen.

I vår behandling av sonevariabelens empiriske innvirkning på effektivitetsscorene har vi tatt utgangspunkt i en estimeringsmodell med elever per hovedtrinn, skolepoeng per hovedtrinn og skolemiljø som produkter, og bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger som faktorinnsats. I behandlingen av sonevariabelen har vi enten dividert faktorinnsatsen på denne variabelen, slik SØF gjør, eller behandlet den som en kontekstuell variabel i Simar og Wilson (2007) sin to-steps-DEA-modell. Som benchmark har vi benyttet en tilsvarende modell, der sonevariabelen ikke inngår.

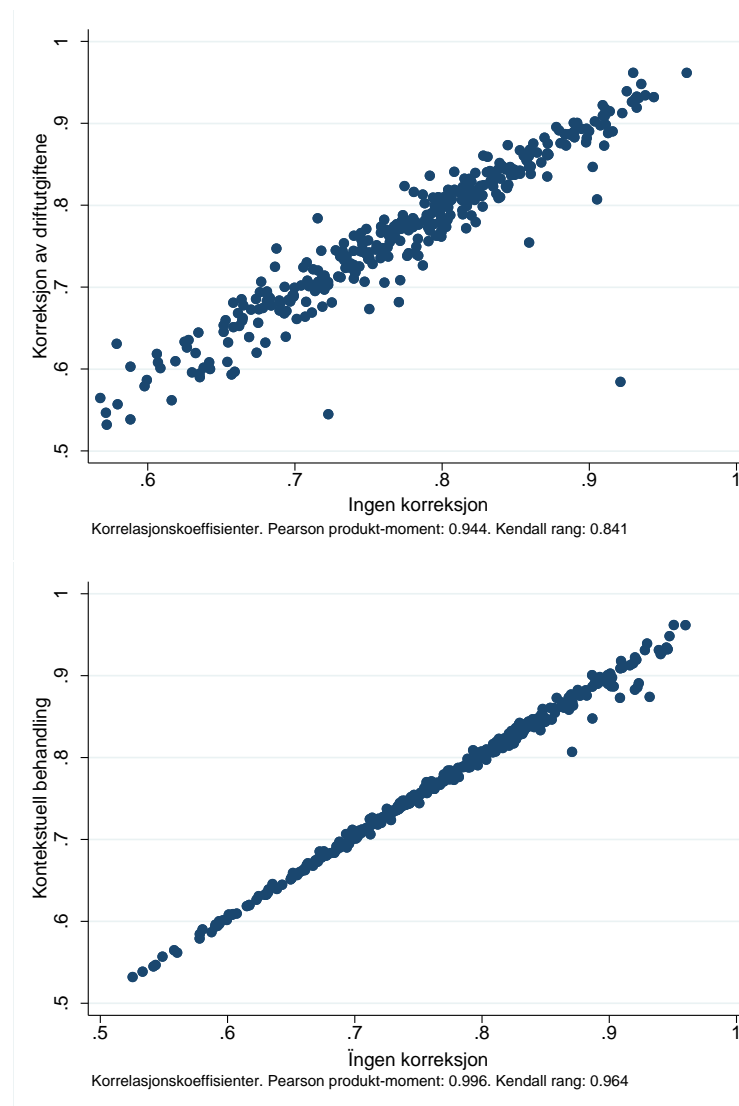
Innvirkningen av behandlingen av sonevariabelen på effektivitetsscorenes frekvensfordeling er framvist i figur 4.15. Vi ser at fordelingen forandres en del ettersom hvilken modell man legger til grunn, uten at den overordnede formen på fordelingen endres vesentlig.



Figur 4.15: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalautbytte for 2018 og 2019. Faktorinnsats: a) og c) Bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger og b) bruttodriftsutgifter mot bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger dividert på sonevariabelen Produkter: Elever og skolepoeng på hvert hovedtrinn, og skolemiljø Kontekstuelle variabler: a) og b) ingen og c) sonevariabelen

Når det gjelder korrelasjonen mellom effektivitetsscorene, endres de også en del, som vist i figur 4.16. Endringene er vesentlig større når man foretar SØFs ad-hoc-messige manipulasjon av kostnadsvariabelen enn når man behandler sonevariabelen som en kontekstuell variabel.

Det er verdt å merke seg at utslaget sonekorrigeringen har på effektivitetsscorene er av omtrent samme størrelsesorden som valget av bruttodriftsutgiftsmål som innsatsfaktor (jamfør seksjon 4.1.1) og inkluderingen av tjenestekvalitetsmål (jamfør seksjon 4.1.4). Det er uheldig at man i de offisielle effektivitetmålingene for kommunal sektor har valgt å foreta et ad-hoc-messig grep knyttet til sonevariabelen, særlig når innvirkningen på resultatene er vesentlig.



Figur 4.16: Korrelasjonen mellom effektivitets-scorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalautbytte for 2018 og 2019.

Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger mot a) bruttodriftsutgifter mot bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger dividert på sonefaktor og b) bruttodriftsutgifter mot bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger

Produkter: Elever og skolepoeng på hvert hovedtrinn, og skolemiljø

Kontekstuelle variabler: Ingen mot a) ingen og b) sonefaktoren

Hvis man kommer opp med en god begrunnelse for hvorfor sonefaktoren kan fange opp kontekstuelle forhold, bør man også vise at den faktisk korrelerer med effektivitetsmålet eller kostnadsålet (som blir tilsvarende). Regresjoner med Simar og Wilson (2007) sin to-steps-DEA-modell for sonefaktoren, som vi presenterte over, viser at dette faktisk er tilfellet. Estimeringsresultatene for den kontekstuelle variabelens innvirkning på ineffektiviteten er gitt i tabell 4.1. Mens effektivitetsmålet vanligvis defineres som et tall mellom 0 og 1, har denne Stata-rutinene «ineffektivitet», den inverse ($1/TE$), som avhengig variabel i trinn-2 regresjonene. Vi ser at koeffisienten er signifikant og positiv. Vi har også gjennomført tilsvarende kjøring med en SFA-modell, der sonefaktoren behandles som en kontekstuell variabel som skifter fronten proporsjonalt innover eller utover. Også denne variabelen gir positive signifikante resultater for sonefaktorens innvirkninger på ineffektiviteten, i dette tilfellet gjennom kostnadsvariabelen.

Tabell 4.1: To-steps-DEA: Innvirkningen av sonevariabelen som kontekstuell variabel på ineffektivitetsscorene (DEA) i grunnskolen i norske kommuner i 2019. Beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalautbytte for 2018 og 2019. Faktorinnsats: Bruttodriftsutgifter mot bruttodriftsutgifter fratrukket arbeidsgivergjift og avskrivninger. Produkter: Elever og skolepoeng på hvert hovedtrinn og skolemiljø. Kontekstuelle variabler: Sonefaktoren. Signifikansnivå: * En prosent, ** fem prosent og *** ti prosent

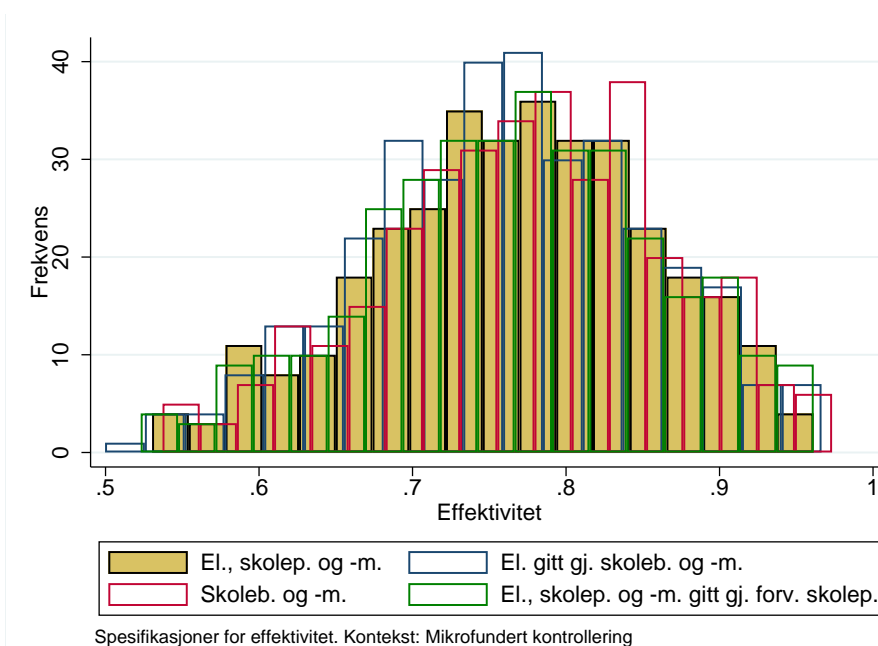
Variabel	Virkning på ineffektivitet
Sonefaktor	1.419*** (0.100)
Konstant	-0.151 (0.104)

SØFs beregninger innbefatter ytterligere kontekstuelle variabler. Også disse er behandlet på en ukonvensjonell måte ved at man har korrigert produktene istedenfor å benytte dem som kontekstuelle variabler. Denne manipulasjonen og det tilsvarende estimeringsopplegget knyttet til behandlingen av de kontekstuelle variablene fremstår likevel mer sofistikert og gjennomtenkt. Borge, Kråkenes og Nyhus (2020) rapporterer nokså upresist at «Elevenes resultater er derfor blant annet korrigert for foreldrenes utdanning, innvandrerbakgrunn og husholdningsinntekt» (side 10, sjettede setning i siste avsnitt). En del flere detaljer knyttet til estimeringen er gitt i Borge, Kråkenes og Nyhus (2018), der også enkelte andre forklaringsvariabler nevnes.

Før effektivitetsestimeringen foretas en estimering av ulike forholds innvirkning på skolepoengene på institusjonsnivå av Utdanningsdirektoratet. På bakgrunn av dette beregnes forventede skolepoeng. Skolebidraget defineres som differansen mellom faktiske og forventede skolepoeng, altså differansen mellom et av produksjonens produkter og en kontekstuell variabel. For at skolebidragsindikatoren skal tilfredsstillende DEAs egenskaper normaliserer SØF den til en indeks lik summen av skolepoeng på landsbasis og differansen mellom faktiske og forventede skolepoeng gitt kontekstuelle forhold, for deretter å multiplisere indeksen med antall elever. Vi referer til seksjon 2.1.3 for detaljer.

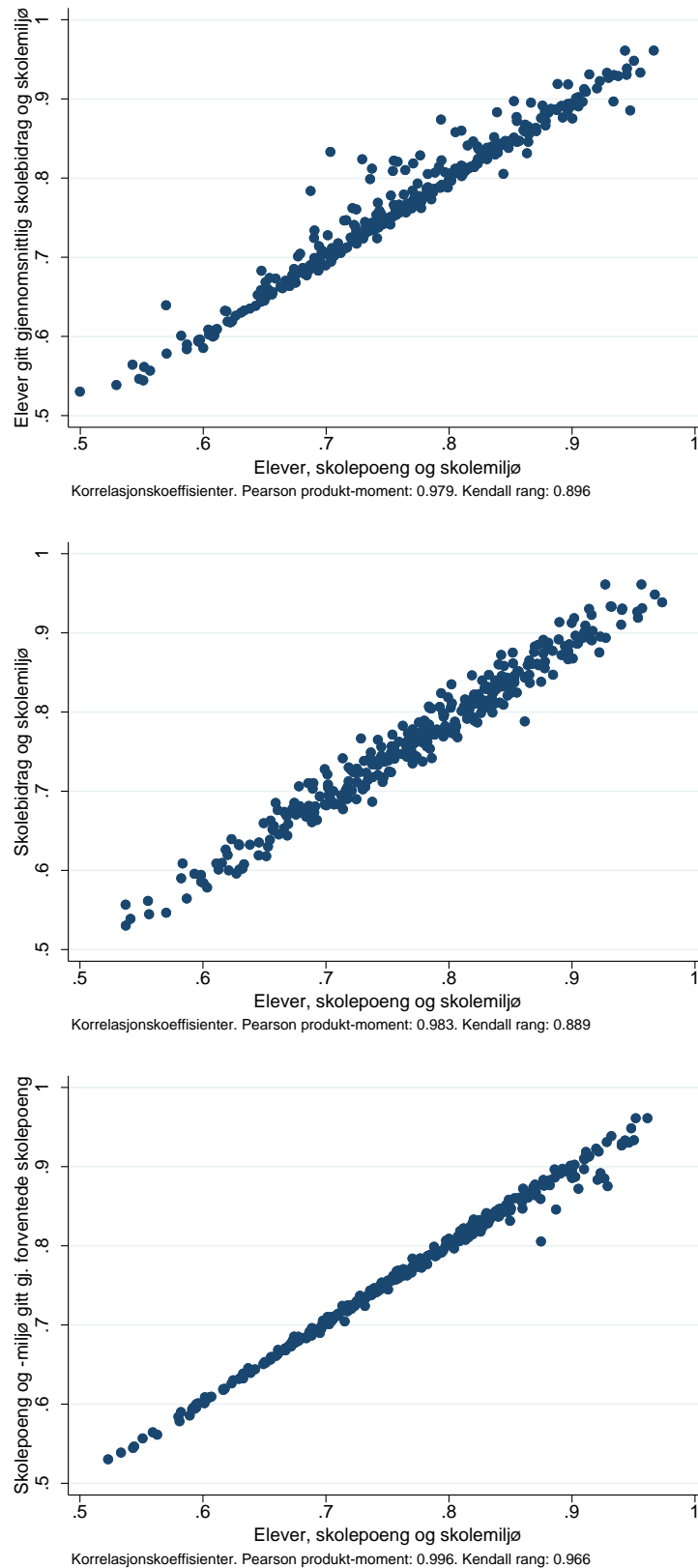
I det følgende vil se på tre alternative behandlinger av de forventede skolepoengene. I benchmark-modellen anvender skolepoeng og elever på hvert hovedtrinn og skolemiljø som produkter, uten at forventede skolepoeng inngår. I den første estimeringsmodellen som tar inn forventede skolepoeng behandler vi elever på hvert hovedtrinn som eneste produkter, mens gjennomsnittlig skolebidrag på hvert hovedtrinn og gjennomsnittlig skolemiljø tas inn som kontekstuelle variabler. I den andre estimeringsmodellen som håndterer forventede skolepoeng benytter vi skolebidrag på hvert hovedtrinn og skolemiljø (alle multiplisert med antall elever) som produkter, uten kontekstuelle variabler. Merk at dette tilsvarer SØFs modellering. I den tredje varianten benytter skolepoeng på hvert hovedtrinn og skolemiljø (multiplisert med antall elever) som produkter og forventede skolepoeng på hvert hovedtrinn som kontekstuelle variabler. Kontekstuelle variabler håndteres igjen med Simar og Wilsons (2007) estimeringsprosedyre for to-trinns-DEA med simultan estimering av stegene.

Innvirkningen av behandlingen av de forventede skolepoengene på fordelingen av effektivitetsscorene er gitt i figur 4.17. Igjen endrer ikke fordelingen seg radikalt når man tar hensyn til forventede skolepoeng, selv om det kan være vesentlige forskjeller for enkelte kommuner. Her spiller det naturligvis inn at Norge er et relativt homogent samfunn, slik at det vil være begrenset hvor stor variasjonen vil være i variablene.



Figur 4.17: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalautbytte for 2018 og 2019. Faktorinnsats: Bruttodriftsutgifter mot bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger. Produkter: a) Elever og skolepoeng på hvert hovedtrinn og skolemiljø, b) elever på hvert hovedtrinn, c) skolebidrag på hver hovedtrinn og skolemiljø og d) elever og skolepoeng på hvert hovedtrinn og skolemiljø Kontekstuelle variabler: a) og c) ingen, b) skolebidrag på hvert hovedtrinn og skolemiljø og d) gjennomsnittlig forventede skolepoeng på hvert hovedtrinn Forkortelser. El.: Elever. Skolep.: Skolepoeng. Skoleb.: Skolebidrag. Skolem.: Skolemiljø.

I figur 4.18 framviser vi hvordan effektivitetsscorene som på ulik måte tar hånd om kontekstuelle forhold gjennom forventede skolepoeng korrelerer med en spesifisering med elever og skolepoeng på hver hovedtrinn, og skolemiljø som produkter, uten kontekstuelle variabler. Vi ser at korrelasjonen er sterkest når gjennomsnittlig forventede skolepoeng per hovedtrinn inkluderes i estimeringsmodellen som kontekstuelle variabler. Samtlige modeller som hensyntar forventede skolepoeng korrelerer like fullt relativt sterkt med benchmark-modellen.



Figur 4.18: Korrelasjonen mellom effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalanthytte for 2018 og 2019. Faktorinnsats: Bruttodriftsutgifter fratrukket arbeidsgivergift og avskrivninger. Produkter: Elever, skolepoeng og skolemiljø, mot a) elever, b) skolebidrag og skolemiljø og c) elever, skolepoeng og skolemiljø. Kontekstuelle variabler: a) mot a) skolebidrag og skolemiljø, b) ingen og c) gjennomsnittlig forventede skolepoeng

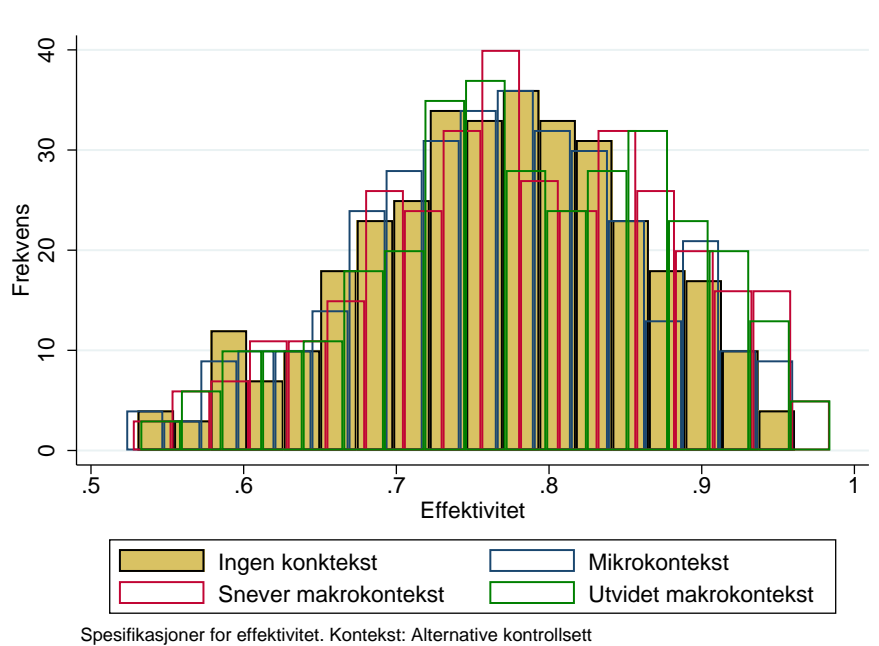
Fordelen med å pre-estimere forventede skolepoeng er at man kan utnytte informasjon på et mer disaggregert nivå enn kommuner. Ulempen er at man mister estimeringseffektivitet ved å foreta estimeringen i to trinn, fremfor å estimere innvirkningen de kontekstuelle variablene har simultant. I tillegg forutsetter førstnevnte estimeringsmodell at de kontekstuelle variablene virker sammen gjennom skoleprestasjoner og ikke på gjennom skolemiljø. Den sistnevnte estimeringsmodellen innebærer at endringer i kontekstuelle variablene i første instans forskyver teknologifronten innover eller utover, og at produksjonsfronten i neste instans omdefineres. Vi referer til seksjon 3.1.4 for mer inngående diskusjon av dette temaet. Innsikten som er viktig å ta med seg inn i vår videre analyse er at det ikke er gitt hvilke av de to måtene å håndtere kontekstuelle variabler på som gir riktigst resultat.

I det følgende tar vi for oss to DEA-spesifikasjoner, der de kontekstuelle variabler tas høyde for i Simar og Wilsons (2007) sin to-steps-DEA-prosedyre. I den første varianten benytter vi sysselsettingsandel med høyere utdanning, befolkningandel første- og andre-generasjons innvandrere (definert som barn av to innvandrere) og median personinntekt i millioner norske kroner som forklaringsvariabelene. Disse er valgt, da de fremheves av Borge, Kråkenes og Nyhus (2020). I den andre varianten har vi i tillegg inkludert andre potensielt relevante kontekstuelle forhold. Disse inkluderer en indikator for befolkningstetthet,¹⁰ kvinneandel, befolkningsandel utenfor arbeidsstyrken i arbeidsfør alder mellom 20 og 66 år og andel av befolkningen i arbeidsfør alder mellom 20 og 66 år.

Idet begge disse variantene estimeres på kommunenivå med kontekstuelle kontroller på kommunenivå, vil vi referere til de respektive pakkene av kontekstuelle variabler som henholdsvis snever makrokontroll og utvidet makrokontroll. I uttestingen av betydningen av disse kontekstuelle korreksjonene benytter vi for alle estimeringsmodeller bruttodriftsutgifter fratrukket arbeidsgiveravgift og avskrivninger som innsatsfaktor, og elever og skolepoeng per hovedtrinn, og skolemiljø som produkter. Som benchmark benytter vi en variant uten kontekstuelle variabler. I tillegg inkluderer vi varianten fra forrige sammelikning, der gjennomsnittlig forventede skolepoeng på hvert hovedtrinn benyttes som kontekstuelle variabler. Med et teoretisk utgangspunkt utgjør vår foretrukne variant for kontekstuell forhåndskorreksjon på mikronivå. Vi vil referere til denne konteksten som mikrokontekst.

En sammenlikning av frekvensfordelingene til effektivitetsscorene tilhørende disse estimeringsmodellene er vist i figur 4.19. Igjen er fordelingene nokså like. Det er like fullt synlig at effektivitetsscorene i snitt er høyere når makrokontekster er byttet, særlig i tilfellet der den utvidede makrokonteksten er benyttet. Dette kan forstås som at forskjellene i effektivitetsscorene jevner seg ut når man tar ulike bakenforliggende forhold inn i betrakningen.

¹⁰ Operativ definisjon: *indikator for befolkningstetthet* = $1000 * (\text{befolkningsandel i tettsted} * \frac{\text{personer bostatt i tettsted}}{\text{tettstedsareal i kommunen i km}^2} + \text{befolkningsandel utenfor tettsted} * \frac{\text{personer bosatt i spredbygde strøk}}{\text{annet landareal i kommunen i km}^2})$

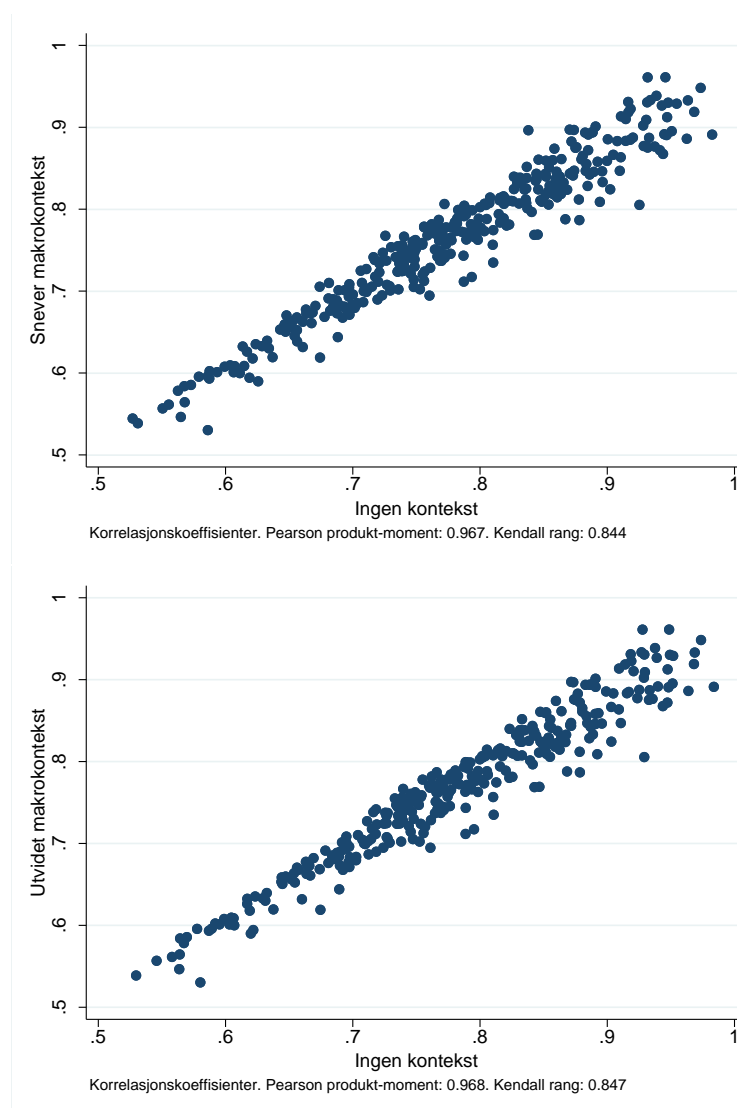


Figur 4.19: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalautbytte for 2018 og 2019.

Faktorinnsats: Brutodriftsutgifter fratrukket arbeidsgivergift og avskrivninger

Produkter: Elever og skolepoeng på hvert hovedtrinn og skolemiljø. Kontekstuelle variabler: a) ingen, b) gjennomsnittlig forventede skolepoeng (mikrokontekst), c) snever makrokontekst og d) utvidet makrokontekst

I figur 4.20 har vi illustrert hvordan DEA-kjøringene med makrovariabler på kommunenivå som kontekstuelle variabler korrelerer med grunnsesifikasjon uten kontekstuelle variabler. Vi ser at den lineære korrelasjonen (Pearsons produktmomentskorrelasjonskoeffisient) med grunnsesifikasjonen er omtrent den samme i spesifikasjonene med makrokontekster, som vår foretrukne spesifikasjon med mikrokontekst. Kendalls rangskorrelasjonskoeffisient med grunnsesifikasjonen er imidlertid litt lavere for regresjonene med makrokontekster enn for våre foretrukne regresjon med mikrokontekst, hvilket innebærer at den ordinale relasjonene mellom observasjonen er noe svakere for de førstnevnte regresjonene.



Figur 4.20: Frekvensfordelinger av effektivitetsscorer i grunnskolen over norske kommuner i 2019 beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalautbytte for 2018 og 2019.

Faktorinnsats: Brutto driftsutgifter fratrukket arbeidsgivergift og avskrivninger

Produkter: Elever og skolepoeng på hvert hovedtrinn og skolemiljø. Kontekstuelle variabler: Ingen mot a) snever makrokontekst og b) utvidet makrokontekst

I tabell 4.2 har vi tatt for oss signifikansen (med p-verdier på ti prosent eller mindre) til de kontekstuelle variablene i regresjonene med makrokontekster. Tabellen kan ses opp mot resultatene Borge, Kråkenes og Nyhus (2018) sine resulater om kontekstuelle variablers innvirkning på grunnskolepoeng og snitt nasjonale prøver åttende trinn på kommunenivå. Dette er rapportert i rapportens tabell V1.1 i Vedlegg 1 om korreksjon av læringsutbytte.

Vi starter med å ta for oss de kontekstuelle variablene som inngår i vår snevre makrokontekst. I tabell 4.2 har sysselsettingsandelen med høyere utdanning har negativ innvirkning på ineffektiviteten i begge regresjonene. I Borge, Kråkenes og Nyhus finner positive og signifikante effekter av de som har lang høyskole og/eller universitetsutdanning på skoleprestasjoner, men ingen av kort høyskole og/eller universitetsutdanning eller videregående utdanning. Videre finner vi ingen signifikante virkninger medianen for personinntekt på ineffektiviteten. Dette kan tolkes som at inntektsnivået ikke påvirker skoleprestasjoner utover virkningene som går i gjennom utdanning. Befolkingandelen med første- og andre-generasjons innvandrere (definert som barn av to innvandrere) har ikke en signifikant

virkning på ineffektiviteten i våre DEA-kjøringer. Muligens kunne utfallet blitt et annet om vi hadde begrenset oss til første generasjons innvandrere og/eller ikke-vestlige innvandrere. Borge, Kråkenes og Nyhus (2018) finner at flyktningsindeksen fra Grønt hefte virker negativt inn på grunnskolepoengene, men inkluderer ikke variabelen i sine regresjoner på snittet i nasjonale prøver åttende trinn.

Av øvrige variabler er vår indikator for befolkningstetthet, kvinneandel og befolningsandel utenfor arbeidsstyrken i arbeidsfør alder mellom 20 og 66 år negative i vår DEA-kjøring med en utvidet makrokontekst. Videre finner vi at andelen av befolkningen i arbeidsfør alder mellom 20 og 66 år har en negativ virkning på ineffektiviteten i grunnskolen. Riktignok er det en del multikollinearitet mellom de kontekstuelle variablene som knytter seg til hvordan hver av dem samvarierer med sentralitet, noe som kan bidra til insignifikante resulater. Borge, Kråkenes og Nyhus (2018) inkluderer ikke kvinneandel, andel i arbeidsfør alder eller befolkningstetthet i sine regresjoner. De finner imidlertid at Grønt heftes indeks for uføre har negativ og signifikant innvirkning på grunnskolepoeng, dog ingen signifikant virkning på snittet nasjonale prøver åttende trinn.

Forfatterne finner heller ingen signifikante virkninger fra særnorsk opplæring eller morsmålsundervisning på skoleprestasjoner, hvilket kan henge sammen med at variablene både reflekterer utfordringer og at disse tas tak i. Andel elever med spesialundervisning på ungdomsskolen har derimot en negativ og signifikant virkning på snittet på nasjonale prøver i åttende klasse, mens andel skilte og separerte i alderen 16 til 66 år ikke har noen virkninger (disse inngår ikke i regresjonene på grunnskolepoeng).

Tabell 4.2: To-steps-DEA: Innvirkningen av kontekstuelle makrovariabler på ineffektivitetsscorene i grunnskolen i norske kommuner i 2019. Beregnet ved innsatsfaktororientert og bootstrappet DEA og variabelt skalautbytte for 2018 og 2019. Faktorinnsats: Bruttodriftsutgifter mot bruttodriftsutgifter fratrukket arbeidsgivergjift og avskrivninger. Produkter: Elever og skolepoeng på hvert hovedtrinn og skolemiljø. Signifikansnivå: * En prosent, ** fem prosent og *** ti prosent

Variabel	Smalt kontrollsett	Bredt kontrollsett
Sysselsettingsandel med høyere utdanning	-0,973*** (0,350)	-0,701** (0,344)
Befolingandel første- og andregenerasjons innvandrere	-0,306 (0,338)	-0,121 (0,726)
Median personinntekt i millioner norske kroner	-0,075 (0,263)	0,275 (0,328)
Indikator for befolkningstetthet		-0,083 (0,001)
Kvinneandel av befolkningen		-2,736 (1,997)
Befolningsandel utenfor arbeidsstyrken i arbeidsfør alder mellom 20 og 66 år		0,795 (0,582)
Andel av befolkningen i arbeidsfør alder mellom 20 og 66 år		-3,536*** (1,237)
Konstant	1,510*** (0,075)	5,208*** (1,345)

Analysene av de kontekstuelle variablene er av bredere interesse enn innvirkningen på effektivitetsscorene. De forteller også noe om betydningen av hver enkelte variabel. Vi har derfor foretatt de kontekstuelle analysene med snever og bred makrokontekster med SFA også. I SFA-spesifikasjonene modelleres de kontekstuelle variablene uten noen form for

transformasjon, slik at endringer i de kontekstuelle variabelene kan tolkes som et proporsjonalt skift av fronten innover eller utover (skjønt fronten trolig også redefineres i noen grad). Dette står i motsetning til innsatsfaktorene og produktene.

4.2 Produktivitets- og effektivitetsendring

4.2.1 Modellspesifikasjon i grunnskoleanalysen

Vi har kjørt en analyse av produktivitetsutviklingen for perioden 2015-2019 for å illustrere hvilken informasjon en standard Malmquistanalyse kan bidra med. I SØF sin analyse av grunnskolen varierer modellspesifikasjonen noe mellom de ulike årene. Vi har ikke hatt tilgang til data for alle årene med samtlige av disse spesifikasjonene. Vi har ikke vurdert det som hensiktsmessig å bruke mye ressurser på å skaffe data for en best mulig modell over flere år, men i stedet lagt vekt på en enkel og forståelig modellspesifikasjon som kan vise metodens potensiale og hovedtrekkene i produktivitetsutviklingen de siste årene. Dette er dermed ikke ment som en analyse med gode estimater for produktivitetsutviklingen for grunnskolen i enkeltkommuner.

Som innsatsfaktorer har SØF hatt årsverk og øvrige utgifter i 2015-17, utvidet til to kategorier årsverk og øvrige utgifter i 2017-18 og kun brukt korrigerede brutto driftsutgifter i 2018-19. Korrigeringen er både for avskrivninger og for bosetningsstruktur. I vår illustrasjonsmodell her bruker vi kun (ukorrigerede) brutto driftsutgifter.

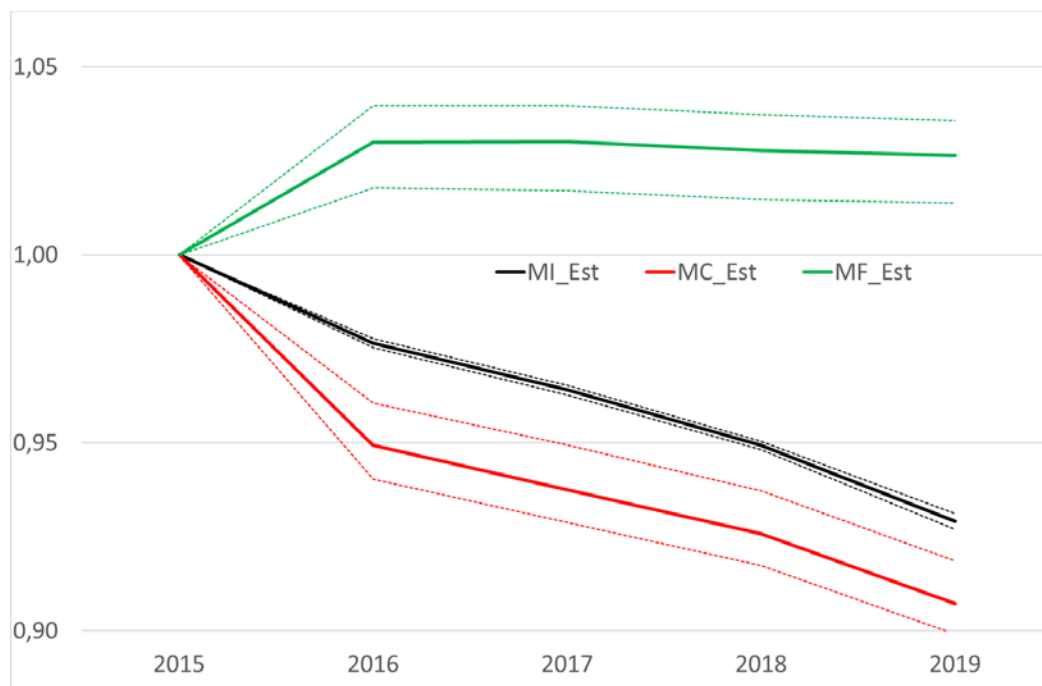
Læringsmiljø, målt som gjennomsnittlig indikatorer på sjuende trinn multiplisert med antall elever i grunnskolen, er brukt i SØFs modeller i alle år. Totalt skolebidrag, målt som gjennomsnittlig skolebidrag ganget med antall elever, er med på to trinn i perioden 2015-17 og på tre trinn i 2017-2019. I vår modell er det kun brukt ukorrigert elevtall på tre skoletrinn.

Både SØFs modell og vår modell bruker DEA med variabel skalautbytte til å estimere fronten. I SØFs analyser er det antatt felles teknologi for to år av gangen, slik at effektivitetsestimaterne for år t er basert på alle observasjoner både i år t og i år $t-1$. I vår analyse er det antatt årlige teknologier men uten mulighet for teknisk tilbakegang, slik at estimatene for år t er basert på alle observasjoner fra år 2015 til år t .¹¹

Mens SØF har benyttet bootstrapping til å skjevhetekorrigere estimatene og beregne statistisk usikkerhet kun for siste analyse 2018-19 har vi her benyttet bootstrapping for alle år.

¹¹ En kunne alternativt ha antatt at teknologifronten også kunne vist tilbakegang, ved at for eksempel normkrav hadde medført at det som ble gjort i 2015 ikke lenger var mulig i 2019, men det har vi her sett bort fra.

4.2.2 Malmquistindeksen og dens komponenter

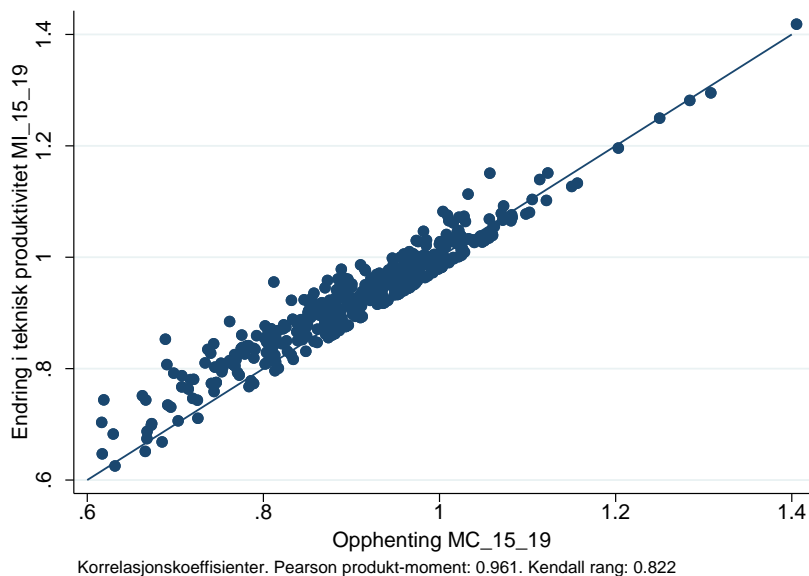


Figur 4.21: Malmquistindeksen MI og dens komponenter frontskjift MF og opphenting MC. Estimater og 95% konfidensintervall for snittet over kommuner.

Den svarte heltrukne linjen (MI_Est) i Figur 4.21 viser estimert gjennomsnittlig produktivitetstutvikling for de observerte kommunene. Tallene er uttrykk for produktiviteten i det aktuelle året sammenliknet med produktiviteten i 2015. Resultatene tyder på en tilbakegang på en til to prosent i gjennomsnittlig produktivitet hvert år, til sammen ca syv prosent over perioden 2015-19. De svarte stiplede linjene angir 95% konfidensintervall for estimatene, og det smale båndet mellom linjene antyder at produktivitetstutviklingen er relativt presist anslått.

Som vist i avsnitt 3.2.3 kan Malmquistindeksen dekomponeres i to hovedkomponenter: En komponent MF måler frontendring, dvs. hvor mye felles teknologiendring bidrar til endring i produktivitet. Den andre komponenten MC måler opphenting, dvs. hvor mye det bidrar til produktivitetstendringen at en kommune nærmer seg fronten. For hver enhet vil en da kunne skrive $MI = MF \times MC$, men dette vil ikke nødvendigvis gjelde for av disse målene.

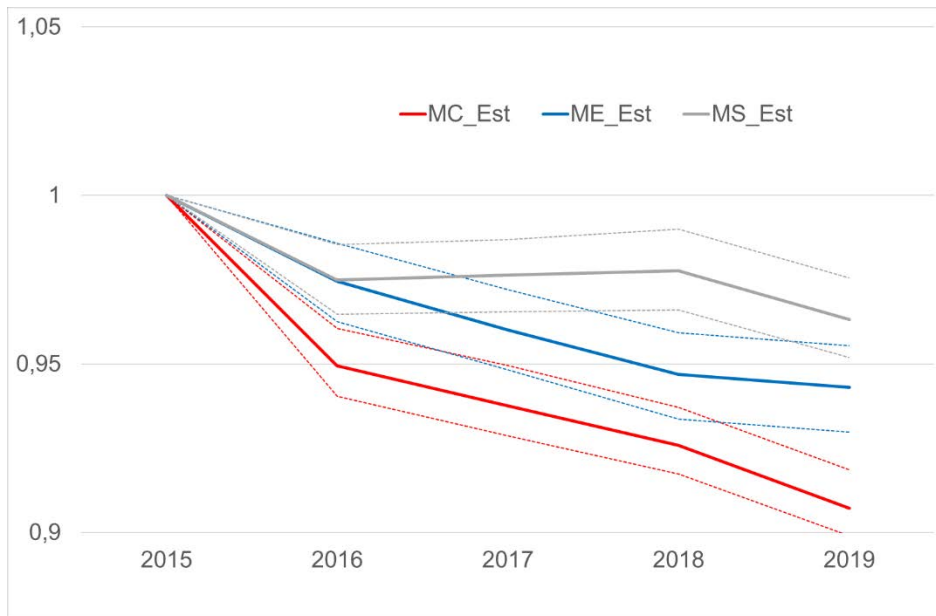
I figur 4.21 gir den gjennomsnittlige estimerte frontendringen (MF_est) et svakt positivt bidrag til produktivitetstendringen, i hovedsak ved en bedring på rundt 3 prosent fra 2015 til 2016. I årene etter er det i snitt liten endring i den estimerte teknologien. Opphengkomponenten (MC_est) viser derimot en jevn tilbakegang gjennom hele perioden, med en samlet nedgang i malmquistindeksen på ca 9% fra 2015 til 2019. Igjen er de stiplede linjene et 95% konfidensintervall rundt estimatet og en ser at presisjonen er mye lavere for MF og MC enn for produktivitetstutviklingen som helhet MI.



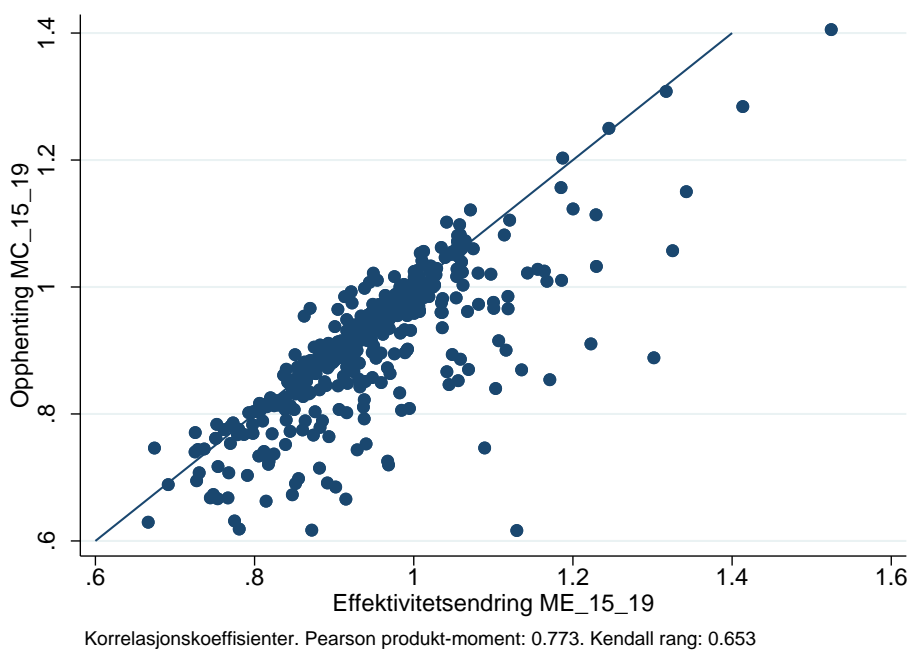
Figur 4.22: Den estimerte Malmquistindeksen MI og komponenten opphenting MC fra 2015 til 2019. Samvariasjon MI og MC for alle observerte kommuner.

Nedgangen i teknisk produktivitet målt ved Malmquistindeksen MI er derfor i all hovedsak et resultat av at opphenting målt ved MC har vært negativ i gjennomsnitt. I figur 4.22 er samvariasjonen mellom MI og MC vist for hver kommune for endring fra 2015 til 2019. Også her ser en at det aller meste av variasjonen i MI skyldes endring i MC, dvs. hvor langt bak fronten den enkelte kommune har beveget seg. De fleste kommuner har en tilbakegang målt ved at opphenting er negativ ($MC < 1$), og langt de fleste av de samme kommunene har også en tilbakegang i total teknisk produktivitet målt ved MI.

Opphentingskomponenten kan som vist i avsnitt 2.2.3 videre dekomponeres i en del som skyldes endring i teknisk effektivitet (ME) og en del som skyldes endring i skalaeffektivitet ME. Igjen har vi at for hver enhet vil $MC = ME \times MS$. Gjennomsnittlig indeksverdi for disse tre er vist i figur 4.23 med sine respektive 95% konfidensintervall. Sammenlignes den blå linjen for estimert ME med den grå for MS ser en at mesteparten av tilbakegangen skyldes nedgang i teknisk effektivitet over perioden, selv om begge komponentene viser tilbakegang. Figur 4.24 viser tilsvarende en klar sammenheng mellom de estimerte ME og MC for de enkelte kommunene, men samvariasjonen ikke er like sterk her som i figur 4.22.

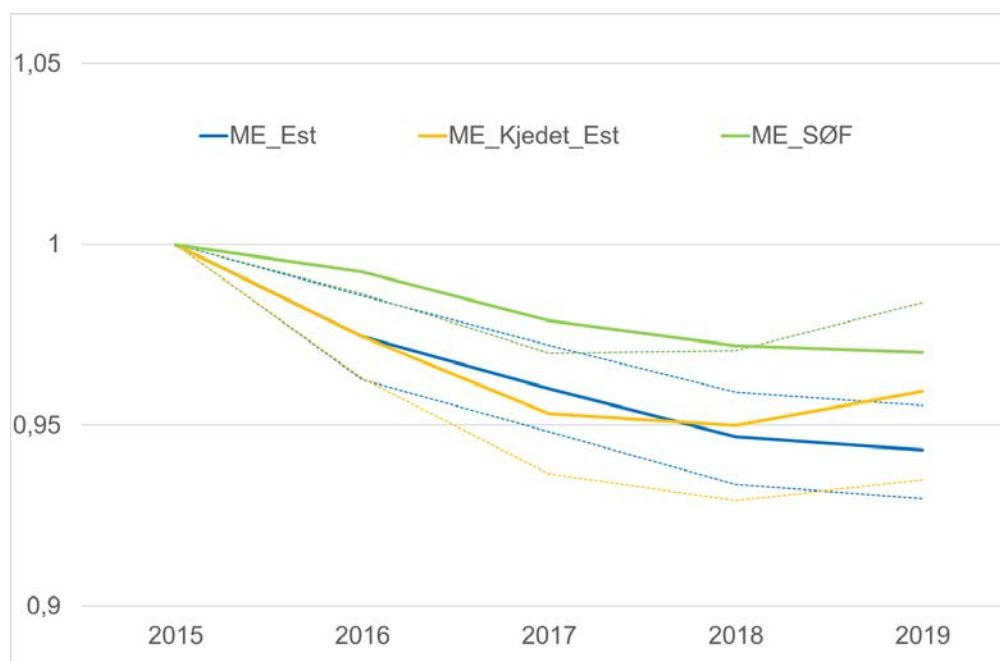


Figur 4.23: Opphøningskomponenten MC i Malmquistindeksen og dens videre dekomponering i effektivitetsendring ME og skalaeffektivitetsendring MS. Estimer og 95% konfidensintervall for snittet over kommuner.



Figur 4.24: Opphøningskomponenten MC og komponenten effektivitetsendring ME fra 2015 til 2019. Samvariasjon MC og ME for alle observerte kommuner.

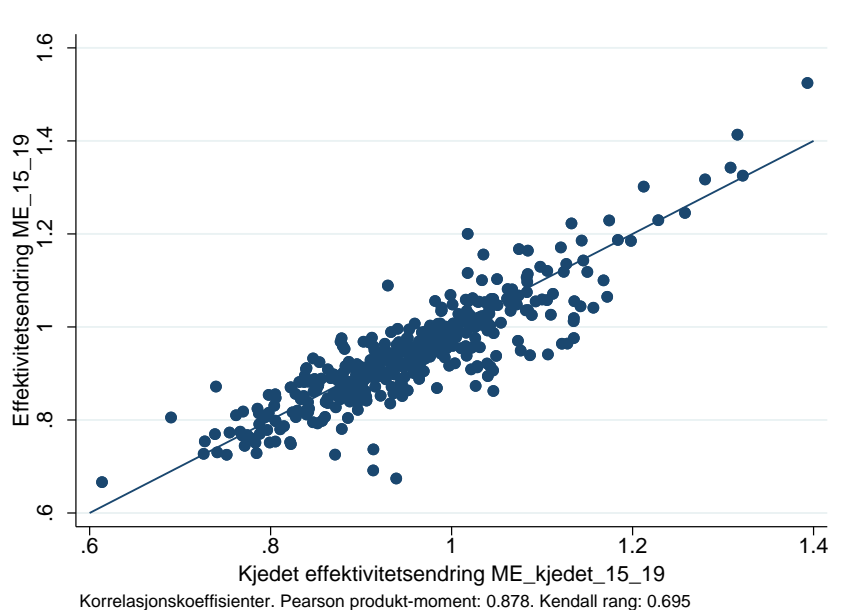
4.2.3 Sammenligning med SØFs metodikk



Figur 4.25: Effektivitetsendringkomponenten ME i Malmquistindeksen sammenliknet med effektivitetsendring som kjedning av parvise årlige sammenlikninger ME_kjedet. Estimater og 95% konfidensintervall for snittet over kommuner. Effektivitetsendring ME_SØF slik det er rapportert av TBU.

I vår analyse er teknologien eller fronten definert for hvert år separat, selv om en i denne analysen antar at det ikke kan være teknisk tilbakegang. I SØFs analyse er derimot frontene definert parvist av det aktuelle året og året før og så kjedet sammen med resultatene fra tidligere analyser. Effekten av denne antakelsen på endring i teknisk effektivitet er vist i figur 4.25, der vår metodikk er vist ved den blå linjen (ME_est) og SØFs kjedete parvise metodikk (men med våre data og modellspesifikasjon) er vist ved den gule linjen (ME_Kjedet_Est). Fra 2015 til 2016 er det bortimot fullt samsvar fordi vi da kun har to år i analysen etter begge metodene. Etter hvert er det mindre samsvar ettersom det er flere års observasjoner som skiller, men resultatene er ikke veldig sprikende. Som vist i figur 4.26 er det også for de enkelte kommunene et visst samsvar mellom resultatene for endring i teknisk effektivitet fra 2015 til 2019 i de to metodikkene, men det er mange enheter som viser klart forskjellig utvikling. Det er f.eks. en del kommuner som har fremgang i effektivitet etter den ene metodikken mens de har tilbakegang etter den andre.

Vi har i figur 4.25 også vist gjennomsnittlig endring i effektivitet slik disse er beregnet til de årlige TBU-rapportene. Også deres analyser viser effektivitetstilbakegang gjennom årene 2015-2019, men betydelig mindre enn i våre analyser. Vi har ikke hatt tilstrekkelig tilgang til de årlige dataene til å reprodusere deres analyser og har kun brukt resultatene. Unntatt i siste periode (2018-19) har disse ikke vært bootstrappet og har derfor ikke informasjon om statistisk usikkerhet og konfidensintervaller. Variabelspesifikasjonene er heller ikke de samme over tid eller mellom deres modell og vår. Det er derfor ikke overraskende at det er mindre samsvar mellom deres resultater og våre for endring av effektivitet gjennom perioden.



Figur 4.26: Effektivitetsendringkomponenten ME i Malmquistindeksen sammenliknet med effektivitetsendring som kjedning av parvise årlige sammenlikninger ME_kjedet fra 2015 til 2019. Samvariasjon ME og ME_kjedet for alle observerte kommuner.

Samlet er det mye informasjon i Malmquistindeksen utover det som en finner når en kun ser på endring i teknisk effektivitet. For det første vil en kunne si noe om hvor mye mere tjenester en får ut av hver enhet ressursbruk (endring i produktivitet), og dernest hvor mye av denne endringen som skyldes endring i mulighetene (frontskift) i skalaeffektivitet og i teknisk effektivitet. Denne informasjonen er ikke tilgjengelig i SØFs analyser hittil. Når det også er endrede modellspesifikasjoner og antakelser over tid er det vanskelig å ha like stor tiltro til resultatene som i en Malmquistanalyse med konstante antakelser. Analysene er ikke veldig tenkisk krevende. Prisen er at en må bruke en noe enklere modellspesifikasjon som lar seg reproducere bakover i tid.

4.3 Produksjonsindeksen

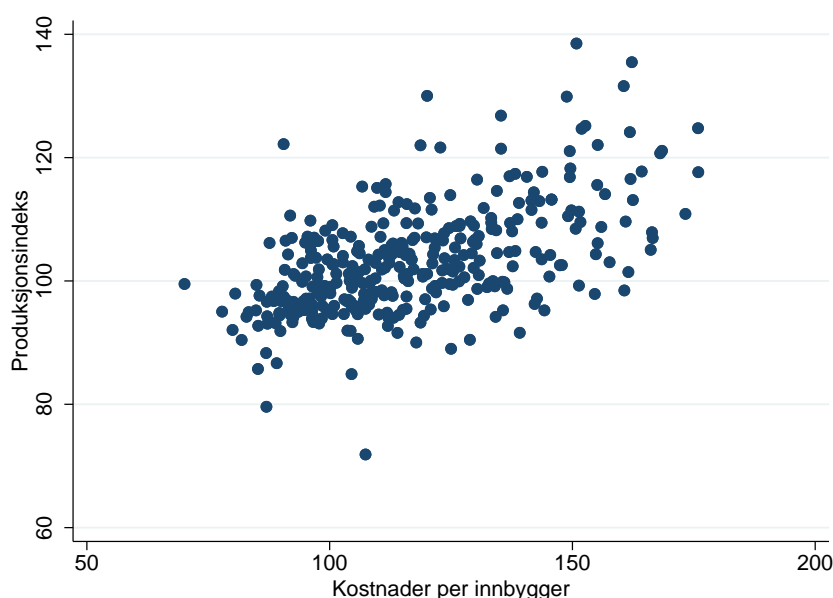
I dette kapitlet ser vi nærmere på produksjonsindeksen for 2019, slik den foreligger beregnet av SØF, samt den alternative produksjonsindeksen beregnet ved hjelp av DEA (jf. kapittel 2.3.3). Det er viktig å påpeke at dette kapitlet kun er ment å illustrere bruken av disse metodene, og ikke noe endelig fasit for måling av kommuneøkonomi. Det er spesielt to viktige forbehold:

1. *For det første* benytter våre sammenlikninger korrigerede brutto driftsutgifter fra KOSTRA til å måle ressursinnsatsen. Vi forstår at dette kan gi noen avvik mot produksjonsindeksen, bl.a. fordi korrigerede driftsutgifter ikke tar hensyn til betalinger til private barnehager samtidig som disse inngår i produksjonsindeksen. Det sentrale for oss er å sammenlikne metodene (som begge er basert på samme datagrunnlag), ikke å gi gode målinger på kommuneeffektivitet. Vi har derfor ikke ønsket å bruke begrensede ressurser på en utvidet innhenting og kvalitetssikring av datagrunnlag for driftskostnader.
2. *For det andre* består produksjonsindeksen av en lang rekke variable med ulik skala/måleenheter. Dette skaper problemer for DEA-metoden, slik vi beskrev i kapittel 2. Men det er også grunn til å stille spørsmål om hvor godt dette håndteres av den gjeldende produksjonsindeksen. I denne applikasjonen gjør vi ingen tilpassing av

metodikken til spesielle variable, men benytter standard DEA på det foreliggende datagrunnlaget for beregning av produksjonsindeksen. Fokuset er utelukkende på å sammenlikne de to metodene basert på det foreliggende datagrunnlaget.

Vi starter med å se nærmere på SØFs produksjonsindeks. I Kommunal- og moderniseringsdepartementet (2020) viser figur 7.3 en positiv sammenheng mellom produksjonsindeksen og korrigert inntekt. Det framgår ikke klart av publikasjonen, men vår forståelse er at det er snakk om korrigert inntekt *per innbygger*. Siden korrigerte bruttodriftsutgifter forventes å være korrelert med korrigert inntekt regner vi med å kunne gjenskape denne figuren basert på våre data. I hht. SØFs framstilling skalerer vi kostnader per innbygger slik at det vektete gjennomsnittet over alle kommuner er 100.

Som forventet viser figur 4.27 at det er en positiv sammenheng mellom produksjonsindeksen og kostnader per innbygger. Siden vi oppfatter at kriteriene i indeksen er normalisert med hensyn til kommunenes innbyggertall tolker vi forholdet mellom variablene som et mål på produktivitet. Det fremgår ikke tydelig i SØFs dokumentasjon av produksjonsindeksen hvorvidt den er normalisert med innbyggertallet, men siden resultatene blir langt mer plausible når vi legger kostnader til innbygger for effektivitetsmålingen velger vi å fokusere på disse resultatene i den videre sammenlikningen.

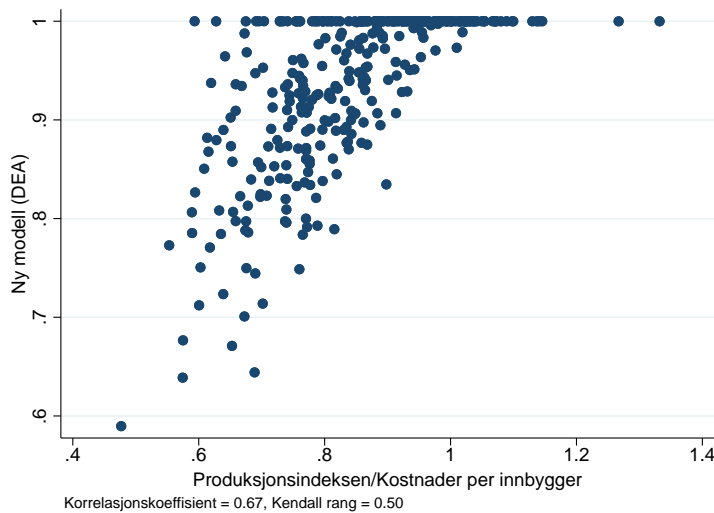


Figur 4.27: Sammenhengen mellom produksjonsindeksen og korrigerte brutto driftsutgifter per innbygger. Kilde: SØFs 2019-beregninger og KOSTRA.

En sentral forskjell mellom de to produksjonsindeksmetodene diskutert i kapittel 2.3.3 – SØFs indeks og en alternativ indeks basert på DEA-metoden – er at SØFs ikke benytter noen referansefront, mens DEA beregner beste mulige produktivitet gjennom å benytte en kommune mot alle andre kommuner: Den nye (DEA-baserte) produksjonsindeksmetoden finner for hver kommune det beste forholdet mellom vektete produkter og korrigerte brutto driftsutgifter som er mulig basert på en sammenlikning av alle kommuner i datasettet. Dette kan forstås som et effektivitetsmål som ligger mellom 0 og 1, hvor 1 beskriver at kommunen har maksimal produktivitet.

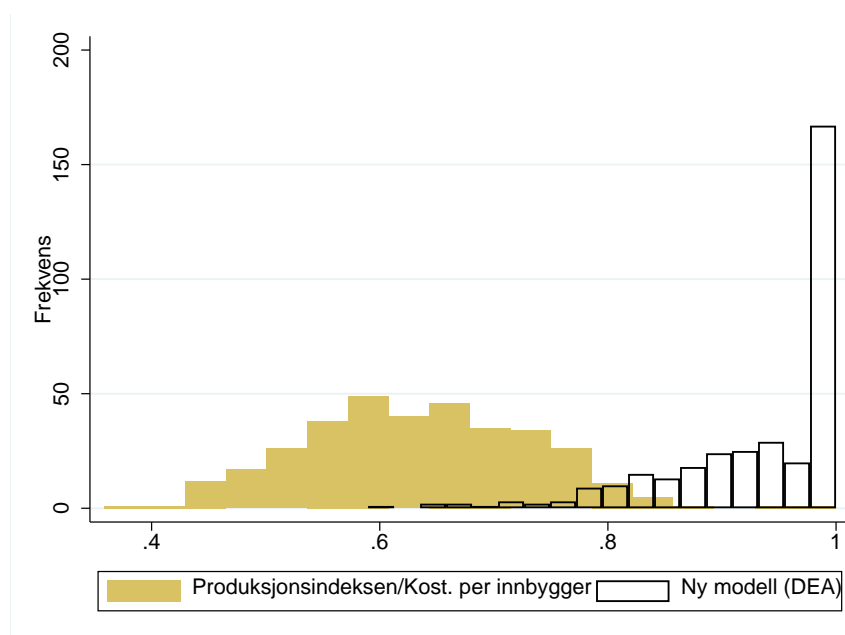
Figur 4.28 beskriver forholdet mellom DEA-indeksen og SØFs produksjonsindeks delt på kostnader per innbygger. Det framgår at DEA-modellen er positiv korrelert med SØFs produktivetsmål (produksjonsindeksen/kostnader per innbygger). Samtidig viser Figur

4.28 at det er mange kommuner med optimal produktivitet i DEA-analysen (dvs. med effektivitetsscore lik 1) som har svært heterogen produktivitet målt etter SØFs produksjonsindeksmetode. En forklaring på dette kan nettopp være at SØFs produksjonsindeksopplegg ikke har noen klar referansenorm for beregning av effektivitet. Samtidig er DEA-estimeringen i dette tilfellet beheftet med dimensjonalitetens forbannelse (dvs. at det er mange variable i modellen sett opp mot antall observasjoner), noe som gjør at svært mange kommuner får en effektivitetsscore lik 1.



Figur 4.28: Sammenhengen mellom DEA-produksjonsindeksen og SØFs produksjonsindeks delt på korrigerte brutto driftsutgifter per innbygger.

For å kunne gi en nærmere beskrivelse av forskjeller mellom de to produksjonsindeksmetodene er det hensiktsmessig å velge en referansenorm, også for SØFs metodikk. En mulighet er å benytte Corrected Ordinary Least Squares eller COLS (Winsten, 1957), som i vårt tilfelle innebærer å velge enheten med høyest produktivitet som benchmark for andre enheter, samt å beregne effektivitet som forholdet mellom hver kommunes produktivitet og den høyeste produktiviteten målt i utvalget. I likhet med DEA gir dette en deterministisk effektivitetsmodell.



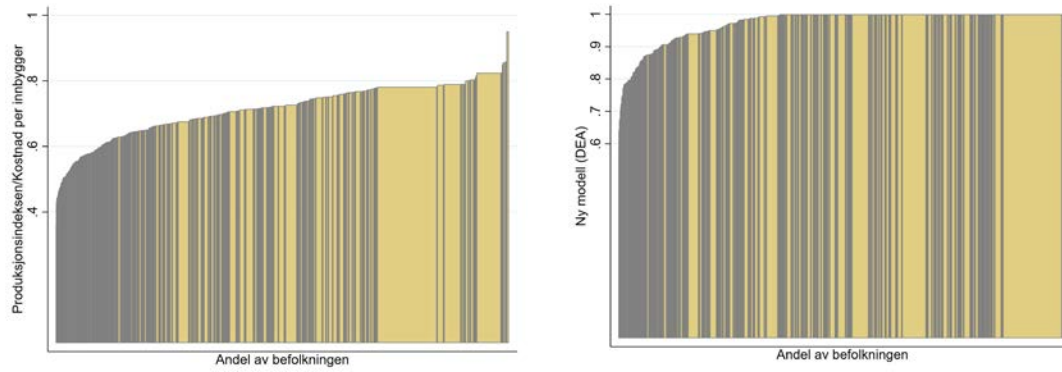
Figur 4.29: En sammenlikning mellom kommuneeffektivitet målt ved DEA og SØFs produksjonsindeks basert på COLS-estimatoren.

Figur 4.29 viser sammenlikningen mellom de to effektivitetsmålene som er avledet av produksjonsindeksen. Vi ser at COLS-metoden gir effektivitetsscore som i større grad er normalfordelte – noe som kan sees i sammenheng med at produksjonsindeksen er innrettet mot å måle avvik fra det gjennomsnittlige produksjonsnivået.¹²

DEA vurderer rundt 50 prosent av kommunene til å være effektive (dvs. har en effektivitetsscore lik 1), og kommunene fremstår som langt mer effektive ved denne målingen. Dette tyder på at det vil være fordelaktig å øke utvalgsstørrelsen (eks. øke antall år i analysen fra ett til flere år) for å muliggjøre større spredning i effektivitetsscorene. Alternativt vil det være mulig å redusere antall variable benyttet i analysen¹³. Disse alternative kan testes ut i det videre arbeidet med produksjonsindeksen.

¹² Mer at resultatene endres dramatisk dersom man ikke legger kostnader per innbyggere til grunn men heller benytter produksjonsindeksen delt på kostnader i effektivitetsmålingen. I dette tilfellet gir COLS-metoden trolig en urealistisk benchmark for måling av kommuneeffektivitet, og det store flertallet av kommuner får en svært lav effektivitetsscore.

¹³ Det bør bemerkes at vi har multiplisert alle produkter (dvs. delindekser) i produksjonsindeksen med kommunenes innbyggertall før estimering med DEA, da vi tolker det dit hen at alle delindekser er blitt normalisert med innbyggertallene i SØFs beregninger av produksjonsindeksen. Formålet med vår korreksjon er å regne produktvariablene om fra forholdstall til volumtall. Dersom man legger de ukorrigerede indeksene til grunn gir DEA en langt større spredning i effektivitetsscorene.



Figur 4.30: Kummulative effektivitetsscorer. Kommunenes andel av den samlede befolkningen angis ved bredden på stolpene.

Til slutt ser vi på hvordan effektivitetsscorene basert på DEA og COLS varierer med kommunenes størrelse, målt ved innbyggertall. Figur 4.30 presenterer alle effektivitetsscorene fra minst til størst. Bredden på stolpene beskriver kommunens innbyggertall relativt til nasjonalt innbyggertall. Begge metodene peker på at større kommuner er blant de mest effektive, mens små kommuner er overrepresentert blant de mest ineffektive beslutningsenhetene.

5 Konklusjon og diskusjon

Teknisk beregningsutvalg for kommunal og fylkeskommunal økonomi (TBU) har bedt om at det skal gjøres en vurdering av dagens metoder for å analysere effektiviteten i kommunene og effektivitetsutviklingen i kommunene som Senter for økonomisk forskning (SØF) gjør årlig for TBU.

SØF har opp gjennom årene levert et solid arbeid i sine analyser av effektivitet i kommunesektoren, i all hovedsak basert på det som har vært standard innenfor effektivitetslitteraturen. Likevel har det i analysene vært foretatt tilpasninger som fremstår som noe mer ad hoc og derfor mindre gjennomskubare og etterprøvbare. SØFs analyser avviker fra standard metodikk innen effektivitetsanalyser på en del punkter.

Dette gjelder bl.a. behandling av kontekstuelle variabler og endring over tid. Vi anbefaler mer bruk av standardmetoder. DEA er primært egnet for volummål, men SØF blander data med ulike målestokker. Dette fører til at effektivitetsberegninger ikke blir korrekte og vanskelige å tolke. Behandling av kvalitet og sammensetning er et vanskelig punkt. Mange aktuelle variabler har spesielle måleenheter som ikke egner seg for DEA som er basert på volumdata og ikke data basert på forholdstall eller diskrete kategorier. Også korreksjoner av driftskostnadene kan spille inn på effektivitetsmålingen. Vi anbefaler bruk av disse som kontekstuelle variabler i en tilleggs- eller ettermodell som kortsiktig løsning, men anbefaler videre forskning på temaet.

Vi finner empiriske forskjeller mellom teknisk effektivitet og produktivitet. Det viser at SØF kan ha nytte av å utvide modellopplegget til å dekomponere teknisk effektivitet og skalaeffektivitet. Det sistnevnte kan gi nyttig informasjon om hvordan kommunestørrelse påvirker produktiviteten.

Spesielt for analyser av effektivitet over tid ser vi ikke SØFs vinduanalyse som hensiktsmessig da effektivitetstallene ikke er sammenliknbare når modellen skifter over tid. Våre beregninger viser at dette har til dels betydelige innvirkninger på resultatene. Kjeding av effektivitetsscore viser da en sammensatt effekt av modell- og effektivitetsendringer. Malmquist-indeksen for produktivitetsendring over diskret tid er standarden i litteraturen og vi anbefaler at denne tas i bruk, da med en enkel variabelspesifikasjon som lar seg bruke for noen år tilbake. Malmquist-indeksen gir også informasjon om produktivitetsutviklingen i tillegg til effektivitetsutviklingen.

SØFs opplegg basert på produksjonsindeksen gir ikke et korrekt mål på effektivitet fordi opplegget gir en indeks for produktivitet, ikke effektivitet, slik begrepene vanligvis brukes. Man trenger en referanseproduktivitet for å beregne effektivitet. Produksjonsindeksen har også flere uønskede egenskaper, som bl.a. konstant bytteforhold mellom produkter. Vi foreslår en mulig alternativ produksjonsindeks som tilfredsstillende vanligvis krav til effektivitetsmåling, men vil ikke anbefale dette som hovedtilnærming nå. Det virker hensiktsmessig å ta en grundig gjennomgang av variablene i produksjonsindeksen og i hvilken grad dagens opplegg er vellykket i å konvertere dem til en felles målestokk. En anbefaling i forbindelse med DEA, vil være å i stor grad benytte tradisjonelle volummål i definisjonen av modellen for produktfunksjonen, og heller korrigere for mer komplekse variabler i en ettermodell.

DEA-metoden er enkel å forstå og solid etablert som en hovedmetode i effektivitetsmåling internasjonalt. Vi anbefaler at denne fortsatt benyttes som hovedmetode, og at usikkerheten fortsatt beskrives ved bruk av bootstrapping. Benchmarking i betydningen partielle forholdstall er ikke spesielt egnet til effektivitetsanalyse. En utvidelse av metodeopplegget er å anvende stokastiske metoder (SFA; StoNED) og/eller paneldatametodikk til å beriket analysene og belyse robustheten til resultatene. Våre beregninger viser at resultatene fra stokastiske modeller avviker sterkt fra den deterministiske DEA-modellen. Paneldata kan gi ny innsikt, bl.a. ved å gjøre det mulig å skille mellom effektivitet og andre forskjeller mellom kommunene som har betydning for ressursbruken.

Løpende statistikk og analyser av effektivitets- og produktivitetsutviklingen i kommunesektoren har stor samfunnsmessig verdi. Vi anbefaler at arbeidet fortsetter, men med større vekt på standard metoder.

Vi har gjennomgått en del kilder til data for ressursbruk og tjenesteproduksjon i kommunesektoren, og dette bør kunne gi grunnlag for rikere modeller i flere delsektorer. Likevel, enkle modeller er oftest mest gjennomsiktede og forståelige.

Etterprøvnbarhet og reproduserbarhet er viktig for resultatenes legitimitet, i tillegg til å gjøre det enklere med evalueringer som denne. Det bør derfor etableres ordninger for deponering eller offentliggjøring (f.eks. på web) av datagrunnlaget både i rå form og i bearbejdede form slik at det kan for fremtiden er enklere å gjøre reproduksjonsstudier eller alternative antakelser og metoder.

I tillegg er det ønskelig å stimulere til videre forskning, f.eks. på hvordan kvalitetsaspekter kan bli tatt hensyn til, særlig i ikke-parametriske eller semi-parametriske metoder. Det er dessuten ønskelig å stimulere forskning om hvordan en skal behandle interkommunalt samarbeid, betydningen av bosettingsmønster og ikke minst utvikling av flere kvalitetsmål basert på mikrodata.

Referanser

- Andersen, P. og Petersen, N. C. (1993). A procedure for ranking efficient units in data envelopment analysis. *Management science*, 39(10), 1261-1264.
- Banker, R. D. (1996). Hypothesis tests using data envelopment analysis. *Journal of productivity analysis*, 7(2-3), 139-159.
- Battese, G. E., og T. J. Coelli (1992). Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. *Journal of Productivity Analysis* 3: 153–169.
- Borge, L.-E. og P. Tovmo (2009): Ny produksjonsindeks for kommunene, SØF-rapport nr. 06/09.
- Borge, L.-E., Kråkenes, T. og O.H. Nyhus (2018). Effektivitet i kommunale tjenester: Analyser for 2015-2016. SØF-rapport nr. 01/18.
- Borge, L.-E., Kråkenes, T. og O.H. Nyhus (2020). Effektivitet i kommunale tjenester: Analyser for 2017-2018. SØF-rapport nr. 05/20
- Borge, L.-E., T. Falch og P. Tovmo (2001): Produksjonsindeks for kommunale tjenester, Rapport, ALLFORSK.
- Borge, L.-E., O.H. Nyhus og P. Tovmo (2011): Bedre måling av tjenesteproduksjonen i kommunene, SØF-rapport nr. 06/11
- Bradford, D.F., Malt, R.A., Oates, W.E. (1969) The rising cost of local public services: some evidence and reflections. *National Tax Journal*, 22, 185-202.
- Caves, D.W., L.R. Christensen og W.E. Diewert (1982) The economic theory of index numbers and the measurement of input, output, and productivity, *Econometrica*, 50, 1393-1414.
- Chambers, R.G., Chung, Y. og R.Färe (1998) Profit, directional distance functions and Nerlovian efficiency, *Journal of Optimization Theory and Application*, 98, 351–364.
- Charnes, A., W. W. Cooper og E. Rhodes. (1978) Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429–444.
- Christensen, L.R., Jorgenson, D.W og L.J. Lau (1973) Transcendental logarithmic production frontiers. *Review of Economics and Statistics*, 55, 28-45.
- Daraio, C., & Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: A probabilistic approach. *Journal of Productivity*.
- Daraio, C., & Simar, L. (2007). Conditional nonparametric frontier models for convex and nonconvex technologies: A unifying approach. *Journal of productivity analysis*, 28(1-2), 13-32.
- Deng, G., Li, L. og Song, Y. (2016). Provincial water use efficiency measurement and factor analysis in China: based on SBM-DEA model. *Ecological Indicators*, 69, 12-18.
- Farrell, M. J. (1957) The measurement of productive efficiency of production.” *Journal of the Royal Statistical Society, Series A*, 120(III), 253–281.
- Farrell, M.J. og Fieldhouse, M. (1962) Estimating efficient production functions under increasing returns to scale. *Journal of the Royal Statistical Society Series A (General)* 125, 252-267.
- Färe, R. og D. Primont (1995) *Multi-Output Production And Duality: Theory And Applications*. Kluwer Academic Publishers, Boston
- Färe, R. og G. Karagiannis (2017) The denominator rule for share-weighting aggregation. *European Journal of Operational Research*, 260, 1175-1180.
- Färe, R. og V. Zelenyuk (2003) On aggregate Farrell efficiencies. *European Journal of Operational Research*, 146, 615-620.

- Førsund F.R. (2017) Measuring effectiveness of production in the public sector. *Omega: The International Journal of Management Science*, 73, 93-103.
- Førsund, F. og L. Hjalmarsson (1979) Generalised Farrell measures of efficiency: An application to milk processing in Swedish dairy plants. *Economic Journal*, 89, 294-315.
- Grifell-Tatjé E og Lovell CAK (1995) A note on the Malmquist productivity index. *Econ. Lett.* 47(2):169–175.
- Hampf, B. og K.L. Rødseth (2019) Environmental efficiency measurement with heterogenous input quality: A nonparametric analysis of U.S. power plants. *Energy Economics*, 81, 610-625.
- Hanson, T. (2019) Estimating output mix effectiveness: An applied scenario approach for the armed forces. *Omega: The International Journal of Management Science*, 83, 39-49.
- Helsedirektoratet (2021) Kostnader, produktivitet og økonomisk status i spesialisthelsetjenesten. SAMDATA Spesialisthelsetjeneste. Rapport IS-3011
- Hernández-Sancho og Sala-Garrido, R. (2009). Technical efficiency and cost analysis in wastewater treatment processes: A DEA approach. *Desalination*, 249(1), 230-234.
- Hollingsworth, B. and Smith, P.C.(2003). The use of ratios in data envelopment analysis. *Applied Economics Letters* 10:733-5.
- Johnson, A.L. og T. Kuosmanen (2012) One-stage and two-stage DEA estimation of the effects of contextual variables. *European Journal of Operational Research*, 220, 559-570.
- Kerstens, K. og I. Van de Woestyne (2011) Negative data in DEA: a simple proportional distance function approach. *Journal of the Operational Research Society*, 62, 1413-141.9
- Kommunal- og moderniseringsdepartementet (2020) Rapport fra Det tekniske beregningsutvalg for kommunal og fylkeskommunal økonomi. URL: <https://www.regjeringen.no/contentassets/7e94c055b37041a68652e488d5e78b97/rapport-fra-det-tekniske-beregningsutvalg-for-kommunal-og-fylkeskommunal-okonomi-november-2020-til-publisering.pdf>
- Kumbhakar, S.C., Lien, G. og J. B: Hardaker (2014) Technical efficiency in competing panel data models: a study of Norwegian grain farming. *Journal of Productivity Analysis*, 41, 321-337.
- Kuosmanen T. og M. Kortelainen (2005) Measuring eco-efficiency of production with data envelopment analysis. *Journal of Industrial Ecology*, 9, 59-72.
- Malmquist, S. (1953) Index numbers and indifference surfaces. *Trabajos de Estadística* 4, 209-242.
- Olesen, O.B., Petersen, N.C. og V.V. Podinovski (2015) Efficiency analysis with ratio measures. *European Journal of Operational Research*, 245, 446-462.
- Olesen, O.B., Petersen, N.C. og V.V. Podinovski (2017) Efficiency measures and computational approaches for data envelopment analysis models with ratio inputs and outputs. *European Journal of Operational Research*, 261, 640-655.
- Pérez-López, G., Prior, D., Zafra-Gómez, J. L., & Plata-Díaz, A. M. (2016). Cost efficiency in municipal solid waste service delivery. Alternative management forms in relation to local population size. *European Journal of Operational Research*, 255(2), 583-592.
- Perlic, B. og R. Ekren (2020) Dokumentasjon av skolebidragsindikatorer for grunnskolen. Indikatorer beregnet for Skoleporten 2020. Notater 2020/16. Statistisk Sentralbyrå. URL: Dokumentasjon av skolebidragsindikatorer for grunnskolen. Indikatorer beregnet for Skoleporten 2020 (ssb.no)
- Rogge, N. og De Jaeger, S (2012). Evaluating the efficiency of municipalities in collecting and processing municipal solid waste: A shared input DEA-model. *Waste management*, 32(10), 1968-1978.
- Shen, Y., Ruan, D. Hermans, E., Brijs, T., Wets. G. og K. Vanhoof (2011) Modeling qualitative data in data envelopment analysis for composite indicators. *International Journal of System Assurance Engineering and Management*, 2, 21-30.
- Simar, L. og P.W. Wilson (1998). Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models. *Management Science*, 44, 49-61.

- Simar, L. og P.W. Wilson (2007) Estimation and inferences in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136, 31-64.
- Schmidt, P., og R. C. Sickles (1984). Production frontiers and panel data. *Journal of Business and Economic Statistics* 2: 367–374.
- Torgersen, A. M., Førsvund, F. R. og Kittelsen, S. A. (1996). Slack-adjusted efficiency measures and ranking of efficient units. *Journal of Productivity Analysis*, 7(4), 379-398.
- Balaguer-Coll, M.T., Prior, D. og E. Tortosa-Ausina (2010) Decentralization and efficiency of local government. *Annals of Regional Science*, 45, 471-601.
- Tulkens, H. og P. Vanden Eeckaut (1995) Non-parametric efficiency, progress and regress measures for panel data: Methodological aspects. *European Journal of Operational Research*, 80, 474-499.
- Tupper, H. C. og Resende, M. (2004). Efficiency and regulatory issues in the Brazilian water and sewage sector: an empirical study. *Utilities Policy*, 12(1), 29-40.
- Winsten, C.B. (1957) Discussion on Mr. Farrell's paper. *Journal of the Royal Statistical Society, Series A*, 120(III), 282–284.
- Zofio, J.L. og A.M. Prieto (2006) Return to dollar, generalized distance function and the Fisher productivity index. *Spanish Economic Review*, 8, 113-138.

Transportøkonomisk institutt (TØI) Stiftelsen Norsk senter for samferdselsforskning

TØI er et anvendt forskningsinstitutt, som mottar basisbevilgning fra Norges forskningsråd og gjennomfører forsknings- og utredningsoppdrag for næringsliv og offentlige etater. TØI ble opprettet i 1964 og er organisert som uavhengig stiftelse.

TØI utvikler og formidler kunnskap om samferdsel med vitenskapelig kvalitet og praktisk anvendelse. Instituttet har et verrfaglig miljø med rundt 90 høyt spesialiserte forskere.

Instituttet utgir tidsskriftet Samferdsel på internett og driver også forskningsformidling gjennom TØI-rapporter, artikler i vitenskapelige tidsskrifter, samt innlegg og intervjuer i media. TØI-rapportene er gratis tilgjengelige på instituttets hjemmeside www.toi.no.

TØI er partner i CIENS Forskningscenter for miljø og samfunn, lokalisert i Forskningsparken nær Universitetet i Oslo (se www.ciens.no). Instituttet deltar aktivt i internasjonalt forsknings-samarbeid, med særlig vekt på EUs rammeprogrammer.

TØI dekker alle transportmidler og temaområder innen samferdsel, inkludert trafiksikkerhet, kollektivtransport, klima og miljø, reiseliv, reisevaner og reiseetterspørsel, arealplanlegging, offentlige beslutningsprosesser, næringslivets transporter og generell transportøkonomi.

Transportøkonomisk institutt krever opphavsrett til egne arbeider og legger vekt på å opptre uavhengig av oppdragsgiverne i alle faglige analyser og vurderinger.

Besøks- og postadresse:

Transportøkonomisk institutt
Gautstadalléen 21
NO-0349 Oslo

22 57 38 00
toi@toi.no
www.toi.no