Rapport 2026/1 | For Kommunal- og distriktsdepartementet

VISTA ANALYSE

Metodikk egnet for kriterieutformingen i fylkeskommunenes utgiftsutjevning ved begrenset datatilfang

Håndtering av små datasett i fylkeskommunenes inntektssystem

Rasmus Bøgh Holmen, Sarah Eidsmo, Pernille Parmer, Harald Svartsund og Haakon Vennemo

Dokumentdetaljer

Tittel	Metodikk egnet for få enheter i kriterieutformingen knyttet til utgiftsutjevningen for fylkeskommunene. Håndtering av små datasett i fylkeskommunenes inntektssystem
Rapportnummer	Rapport 2026/1
Forfattere	Rasmus Bøgh Holmen, Sarah Eidsmo, Pernille Parmer, Harald Svartsund og Haakon Vennemo
ISBN	978-82-8126-758-9
Prosjektnummer	25-RBH-11
Prosjektleder	Rasmus Bøgh Holmen
Kvalitetssikrer	Pernille Parmer og Haakon Vennemo
Oppdragsgiver	Kommunal- og distriktsdepartementet
Dato for ferdigstilling	12. mai 2026
Kilde forsidefoto	Martin Martz, Unsplash, 24. juni 2023
Tilgjengelighet	Offentlig
Nøkkelord	Fylkeskommunenes inntektssystem; hybridmodeller; maskinlæring; normativ metode; ressurskarlegging; statistisk metodikk; systemevaluering; systemkartlegging; utgiftsutjevning; økonometri

Om Vista Analyse

Vista Analyse AS er et samfunnsfaglig analyseselskap med hovedvekt på økonomisk utredning, evaluering, rådgivning og forskning. Vi utfører oppdrag med høy faglig kvalitet, uavhengighet og integritet.

Våre medarbeidere har meget høy akademisk kompetanse og bred erfaring innenfor konsulentvirksomhet. Ved behov benytter vi et velutviklet nettverk med selskaper og ressurspersoner nasjonalt og internasjonalt. Selskapet er i sin helhet eiet av medarbeiderne.

Forord

På oppdrag fra Kommunalavdelingen i Kommunal- og distriktsdepartementet har Vista Analyse i denne rapporten utredet metodikk egnet for å håndtere at få enheter ligger til grunn for kriterieutformingen knyttet til utgiftsutjevningen i fylkeskommunenes inntektssystem. Prosjektet vurderer hvordan man kan og bør tilnærme seg få beslutningsenheter i fastsettelsen av kriteriene i inntektssystemet. Empiriske målinger inngår ikke i oppdraget.

Oppdraget har blitt ledet av Rasmus Bøgh Holmen. Prosjektmedarbeidere er Sarah Eidsmo, Harald Svartsund, Haakon Vennemo og Pernille Parmer. De to sistnevnte har også vært interne kvalitetssikrere for ulike deler av rapporten.

Det opprinnelige rapportutkastet ble ferdigstilt før julen 2025. Siden har vi foretatt enkelte revideringer i tilknytning til et seminar om revisjon av utgiftsutjevningen i fylkeskommunenes inntektssystem, som ble avholdt 8. mai 2026.

Vår kontaktperson for oppdraget hos Kommunal- og distriktsdepartementet har vært seniorrådgiver Dorottya Bognár-Lahr. Karen Nystad Byrhagen og Lars Tore Rydland har også vært involvert fra departementets side. Departementet har kommet med gode tilbakemeldinger til det tidligere utkastet av rapporten.

Vi ønsker å takke departementet for gode innspill og et godt samarbeid.

Rasmus Bøgh Holmen

Prosjektleder
Vista Analyse AS
12. mai 2026, Oslo

Ordliste

Delkostnadsnøkkel	Kostnadsnøkkel for en tjenestefunksjon
Delkostnadsvekt	Vekten som angir hvor mye et tjenesteområde bidrar til den samlede kostnadsnøkkelen
Delkriterium	Del av et hovedkriterium som er brutt ned i bestanddeler
Hovedprinsipp	Prinsipp som motiverer utgiftsutjevningen i det kommunale inntektssystemet
Hybridmodeller i det kommunale inntektssystemet	Fastsettelse av kriterievekter ved en kombinasjon av statistisk og normativ metodikk, og/eller ressurskartlegginger
Inntektssystemet for kommunal sektor	En finansieringsmodell som fordeler økonomiske midler til primærkommuner og fylkeskommuner med sikte på å sikre likeverdige tjenester på tvers av landet
Indikator i det kommunale inntektssystemet	En statistisk variabel som gir mål på et kriterium
Kontekstuell variabel	En variabel som beskriver egenskaper ved omgivelsene til en enhet
Kostnadsnøkkel	Anslått verdi som uttrykker utgiftsbehov
Kriterieregresjon	Regresjon for å fastsette kriterievektene i det kommunale inntektssystemet
Kriterievekt	Hvor mye hvert kriterium teller i nøkkelen
Kriterium	Konkret og målbar indikator som operasjonaliserer et utgiftsbehov
Normativ metodikk i det kommunale inntektssystemet	Fastsettelse av kriterievekter basert på vurderinger av hva som anses som normativt riktig
Paneldata	Data bestående av flere enheter over flere tidspunkter
Ressurskartlegging i det kommunale inntektssystemet	Fastsettelse av kriterievekter ved å systematisere faktisk ressursbruk
Statistisk metodikk i det kommunale inntektssystemet	Fastsettelse av kriterievekter ved statistisk analyse av relevante data
Tverrsnittdata	Data bestående av flere enheter på ett tidspunkt
Utgiftsutjevning i inntektssystemet	Omfordeling av statlige overføringer for å utjevne forskjeller i kommunale utgifter, slik at alle kommuner får tilstrekkelige ressurser til å tilby likeverdige tjenester

Innhold

Sammendrag og konklusjoner 6

1 Innledning 8
- 1.1 Fastsettelse av kriterier i det kommunale inntektssystemet 8
- 1.2 Utfordringer med å fastsette kriterier for fylkeskommunene 8
- 1.3 Struktur 9

2 Oppbygging av inntektssystemet for fylkeskommunene 10
- 2.1 Kriteriehierarkiet i det kommunale inntektssystemet 10
- 2.2 Vektingshierarkiet i det kommunale inntektssystemet 11

3 Tidligere vurderinger av håndteringen av få enheter 12
- 3.1 Rattsø-utvalget 12
- 3.2 Borge-utvalget 12
- 3.3 Inntektssystemutvalget 13
- 3.4 Ekspertutvalget for fylkeskommunens inntektssystem 13

4 Håndtering av få enheter i andre inntektssystemer 15
- 4.1 Praksiser i andre sektorer 15
- 4.2 Praksis i andre land 16

5 Metodikk for fastsettelse av systemets kriterier 18
- 5.1 Statistisk metodikk 18
- 5.2 Alternativ metodikk 23

6 Systemevaluering 25
- 6.1 Utforming av datasettet 25
- 6.2 Evaluering av modellvalget 27

Referanser 31

Appendiks 33

A Mulige variabler i kriterieregresjonene 34

Sammendrag og konklusjoner

Fastsettelsen av kriteriene i fylkeskommunenes inntektssystem er ikke rett fram, ettersom det er relativt få fylkeskommuner til å kjøre regresjonsanalyser. Mens offentlige utvalg har viet økende oppmerksomhet til problematikken forbundet med få enheter i fastsettelsen av kriteriene i fylkeskommunenes inntektssystem, benytter systemer i andre sektorer og land i stor grad seg av formelbaserte modeller og kostnadsanslag. Vi anbefaler generelt at kriterieanalysene for fylkeskommunene gjennomføres ved lineære paneldataanalyser, der man utnytter tidsdimensjonen i dataene og holder seg til metodikk som er gjenkjennelig for interessentene av inntektssystemet. Valg av modeller og data for fastsettelse av kriterievektene bør skje med bakgrunn i empirisk uttesting og modellevaluering.

Fastsettelsen av kriteriene i fylkeskommunenes inntektssystem er ikke rett fram, ettersom det er relativt få fylkeskommuner til å kjøre regresjonsanalyser. Hensikten med inntektssystemet er å sette alle fylkeskommuner like i stand til å tilby nødvendige tjenester til sine innbyggere, der kriteriene fanger opp forskjeller i forutsetningene for drift. Generelt er statistisk metodikk godt egnet for å fastslå objektive kriterier, men det er lite formålstjenlig å operere med kriterier som er lite robuste. Når grunnlaget for kriterieregresjoner vurderes til å være for tynt i fastsettelsen av kriterievektene, benyttes ressurskartlegginger, normativ metodikk og hybride modeller.

Mens offentlige utvalg har viet økende oppmerksomhet til problematikken forbundet med få enheter i fastsettelsen av kriteriene i fylkeskommunenes inntektssystem, benytter systemer i andre sektorer og land i stor grad seg av formelbaserte modeller og kostnadsanslag. Tidligere vurderinger viser at få fylkeskommuner gir et svakt statistisk grunnlag for utgiftsutjevningen, noe som har ført til stor vekt på skjønn, metodiske tilpasninger og delvis bruk av alternativ metodikk framfor rene regresjonsanalyser. Behovs- og kostnadsjustering ligger også til grunn både for helseforetakenes inntektsmodell og politiets ressursallokeringsmodell. I Sverige, Sveits og Australia brukes formelbaserte modeller for kostnads- og utgiftsutjevning, med varierende innslag av statistiske analyser, enhetskostnader og behovs- og kostnadsjusteringer.

Vi anbefaler generelt at kriterieanalysene for fylkeskommunene gjennomføres ved lineære paneldataanalyser, der man utnytter tidsdimensjonen i dataene og holder seg til metodikk som er gjenkjennelig for interessentene av inntektssystemet. Jo mindre empirisk testing skulle tilsi noe annet, anbefaler vi at regresjonene gjennomføres ved generaliserte minste kvadraters metode (GLS) med tidsfaste effekter, der man tar høyde for at regresjonenes restledd avhenger av forklaringsvariablenes størrelse (heteroskedastisitet) og forrige rundes restledd (autokorrelasjon). I detaljerte og utforskende empiriske undersøkelser av hva som er hensiktsmessig metodikk for kriterievektene og robusthetssjekker knyttet til vektenes størrelsesorden, vil det være aktuelt med en bredere empirisk tilnærming som også utforsker potensialet i maskinlæring, bayesianske regresjoner og frontestimering. Dersom datagrunnlaget for statistisk modellering ikke vurderes som tilfredsstillende, kan ressurskartlegging, normative vurderinger eller hybride metoder benyttes.

Valg av modeller og data for fastsettelse av kriteriene bør skje med bakgrunn i empirisk uttesting både i form av uttesting og modellevaluering. Indikatorene i kriterieregresjonene bør velges med bakgrunn i at de forklarer (de korrigerte) driftsutgiftene innenfor en fylkeskommunenes tjenester på en robust måte uten høy multikollinearitet (høy korrelasjon mellom forklaringsvariablene) og omvendt kausalitet, samtidig som de er målbare med akseptabel presisjon, gjenkjennelige og tilgjengelige for etterprøving. I paneldataregresjoner for fylkeskommunene vil endringer i

inndelingen av fylkeskommunene være en utfordring som krever at man behandler aggregerte og disaggregerte enheter av samme konstellasjoner av fylker som separate enheter eller harmoniserer tidsserien. Flere verktøy kan benyttes i evaluering av modellvalget for fastsettelsen av kriteriene, herunder vurdering av teststatistikker, sensitivitetsanalyser, metodetriangulering, Monte Carlo-simuleringer og maskinlæring, samt vurdering av gyldigheten og etterprøvingen av vektene ved alternativ metodikk.

1 Innledning

Det kommunale inntektssystemet skal legge til rette for at alle kommuner er i stand til å tilby likeverdige, nødvendige tjenester til innbyggerne, uavhengig av geografiske og demografiske forskjeller. Dette prinsippet er forankret i kommuneloven (§ 2-1 og § 2-2) og reguleres gjennom statsbudsjettet og inntektssystemvedtak (Kommunal- og distriktsdepartementet 2024). Utgiftsutjevningen skal kompensere primærkommuner og fylkeskommuner med høyere kostnader for tjenesteyting.

Vi vil i fortsettelsen innlede med å diskutere fastsettelsen av kriteriene i det kommunale inntektssystemet i delkapittel 1.1, før vi går inn på utfordringene for å fastsette disse for fylkeskommunene i delkapittel 1.2. Vi gir en oversikt over rapportens struktur i delkapittel 1.3.

1.1 Fastsettelse av kriterier i det kommunale inntektssystemet

I utgiftsutjevningen spiller kriterier en sentral rolle. Kriteriene er tallfestede vekter som viser hvor stor betydning ulike kommunekriterier har for kostnadsnivået i kommunal tjenesteyting. Eksempler på kriterier er antall eldre, bosettingsmønster og reiseavstander. Vektene brukes i utgiftsutjevningen til å beregne hvor mye ekstra midler en primærkommune eller fylkeskommune skal få for å kompensere for ufrivillige kostnadsulemper. Jo høyere vekten er, desto større kompensasjon gis for det aktuelle kriteriet.

Kompensasjonen gjennom utgiftsutjevningen gis som en omfordeling innenfor en gitt ramme, der noen fylkeskommuner får et tillegg, mens andre fylkeskommuner får et trekk. Dette innebærer at økt kompensasjon for en fylkeskommune ikke medfører økte statlige overføringer, men snarere en omfordeling av midler fra andre fylkeskommuner.

For primærkommunene brukes statistiske analyser for å identifisere sammenhenger mellom driftsutgifter per innbygger og faktorer som påvirker kostnadsnivået. Vanlig metode er lineære tverrsnittregresjoner, der man ser på data fra ett år for alle kommuner. For kommunene gir dette et godt analysegrunnlag, fordi antallet kommuner er stort.

1.2 Utfordringer med å fastsette kriterier for fylkeskommunene

Fastsettelsen av kriteriene i utgiftsutjevningen er mer krevende for fylkeskommunene. Årsaken er at antallet enheter er begrenset med kun 15 enheter siden 2024, noe som gjør det vanskelig å identifisere klare sammenhenger mellom driftsutgifter og nøkkelfaktorer. Tradisjonelle metoder, som lineære tverrsnittregresjoner, egner seg dårlig på grunn av få datapunkter. Det er også uheldig, hvis små endringer gi store utslag i utgiftsfordelingen. Derfor kan det være nødvendig å benytte alternative metodiske tilnærminger for fylkeskommunene.

Formålet med denne utredningen er å erverve kunnskap om hvilken metodikk som er best egnet til å analysere kostnadsforskjeller på tvers av fylkeskommuner, som i neste instans kan benyttes i utledningen av kriterier og tilhørende vekter i inntektssystemets utgiftsutjevning. Vi tar en bred tilnærming til tematikken, der vi utforsker ulike deler av mulighetsrommet snarere enn å

utelukkende fokusere på tilnærminger som er nært beslektet med dagens praksis. Utredningen er prinsipiell og bør etter vår vurdering følges opp av empiriske undersøkelser.

I vår gjennomgang vil vi ta for oss metodene som benyttes i utformingen av kostnadsnøklene i dagens inntektssystem for fylkeskommunene (Kommunal- og distriktsdepartementet 2023). Per i dag baserer delkostnadsnøklene for videregående opplæring og tannhelse seg på både statistisk analyse og gjennomsnittskostnadsbetraktninger. Innenfor samferdsel benyttes statistisk analyse for kollektivtrafikk og i hovedsak fylkesvei, men normativ metode ligger til grunn for ferger og til en viss grad fylkesvei. Utgiftsutjevningen baseres på et beregnet utgiftsbehov per innbygger for hver fylkeskommune sett opp mot landsgjennomsnittet, beregnet med bakgrunn i objektive kriterier i omgivelsene utenfor deres kontroll og ikke på de faktiske utgiftene. Basert på hver tjenestefunksjons andel av nettodriftsutgifter som inngår i utgiftsutjevningen, vektes delkostnadsnøklene sammen til en samlet kostnadsnøkkel.

1.3 Struktur

Etter denne innledningen i kapittel 1 gjør vi kort greie for oppbyggingen av fylkeskommunenes inntektssystem med fokus på utgiftsutjevningen i kapittel 2. Deretter gjennomgår vi tidligere vurderinger av, og andre inntektssystemers håndtering av få enheter i henholdsvis kapittel 3 og kapittel 4. Videre redegjør vi for mulig estimeringsmetodikk i kapittel 5. Til slutt behandler vi implementeringen av empiriske målinger ved bruk av datasett med få enheter i kapittel 6. I tillegg ha vi redegjort for bredere sett av kandidater for kriterier og kontroller i kriterieregresjonene for fylkeskommunenes utgiftsutjevning i appendiks A.

2 Oppbygging av inntektssystemet for fylkeskommunene

Fylkeskommunenes inntektssystem er dokumentert i Grønt Hefte (Kommunal- og distriktsdepartementet 2025). Systemet består av to hovedmekanismer. Den første er inntektsutjevning, som reduserer forskjeller i skatteinntekter mellom kommuner. Den andre er utgiftsutjevning, som kompenserer primærkommuner og fylkeskommuner med høyere kostnader for tjenesteyting. Utgiftsutjevningen skjer gjennom et hierarkisk system av begreper, som kobler prinsipper, kriterier og vekter til beregningen av fylkeskommunenes samlede utgiftsbehov.

I det følgende gjennomgår vi hierarkiene for kriterier og kriterievekter i det kommunale inntektssystemet i henholdsvis delkapittel 2.1 og delkapittel 2.2.

2.1 Kriteriehierarkiet i det kommunale inntektssystemet

Et hovedprinsipp i utgiftsutjevningen er at det skal kompenseres for ufrivillige variasjoner i utgiftsbehovet, det vil si forhold som fylkene ikke kan påvirke selv. Hensikten med utgiftsutjevningen i det kommunale inntektssystemet er videre å sette primærkommunene og fylkeskommunene i stand til å tilby likeverdige tjenester. Det blir et empirisk spørsmål hvilke variasjoner i utgiftsbehovet det skal kompenseres for. Dette kan for eksempel dreie seg om demografiske forhold, geografiske forskjeller og tjenestespesifikke behov, inkludert variasjoner i etterspørsel, kostnadsintensitet og produksjonsbetingelser. Den kommunale utgiftsutjevningen skal ikke være påvirkelig eller gi vridende insentiver for primærkommunene og fylkeskommunene, samt være objektiv og etterprøvbar.

Hovedprinsippene er overordnede og styrende for utformingen av kostnadsnøklene, men brukes ikke direkte som beregningsvariabler. For fylkeskommunene kan sektorielle kostnadsnøkler knyttes til videregående opplæring, tannhelse og samferdsel. Kostnadsnøklene konkretiseres gjennom kriterier, som er målbare størrelser knyttet til konkrete tjenesteområder. I dagens inntektssystem for videregående opplæring omfatter kriteriene blant annet antall innbyggere i den aktuelle aldersgruppen, indikatorer for søkere til ulike utdanningsprogram og kostnadskrevende program, samt kriteriet om reiseavstand for befolkningen. Innen tannhelse inkluderer kriteriene antall personer i prioriterte aldersgrupper med rett til offentlige tjenester, samt kriterier for voksne med psykisk utviklingshemming og reiseavstand.¨

For samferdsel i fylkeskommunen brukes kriterier for vedlikeholdsbehov for fylkesvei, kollektivtransport i form buss og bane, og båt og ferge. Delkostnadsnøklene bygger på ulike metodiske prinsipper. Fylkesveinøkkelen er basert på en normativ, modellbasert beregning av vedlikeholdsbehov, kollektivnøkkelen bygger på statistiske analyser av utgiftsbehov, mens båt- og fergenøkkelen har en mer hybrid karakter med elementer av både enkle indikatorer og historiske kostnader. Kriteriene er valgt for å reflektere de mest betydningsfulle strukturelle kostnadsforskjellene mellom fylkeskommunene. Kapitalkostnader er ikke en del av den kommunale utgiftsutjevningen med unntak av fylkesveier og fergesamband.

Kriteriene kan videre deles inn i det som kan forstås som underkriterier, der mer detaljerte kostnadsdrivere er relevante. For fylkeskommunene er veinøkkelen og fergenøkkelen eksempler på slike delkostnadsnøkler, der underkriterier inngår. Det er ingen klare underkriterier for tannhelse eller videregående opplæring.

2.2 Vektingshierarkiet i det kommunale inntektssystemet

Både kriterier og underkriterier måles ved hjelp av indikatorer, som er objektive statistiske variabler som inngår i beregningen av utgiftsbehovet. Indikatorene gir kriteriene et konkret og tallmessig uttrykk for hver fylkeskommune, muliggjør beregning av kostnadsnøkler og sikrer systematisk sammenlikning av utgiftsbehov.

Hvert kriterium og underkriterium tildeles en vekt. Kriterievekten reflekterer hvor stor betydning kriteriet skal ha i beregningen av utgiftsbehovet, og i hvilken grad kriteriet antas å forklare systematiske kostnadsforskjeller. Vektene fastsettes som regel på grunnlag av regresjonsanalyser av faktiske utgifter, supplert med bruk av andre metoder der datagrunnlaget er begrenset eller statistisk presisjon lav. Kriterier og vekter kombineres til én delkostnadsnøkkel per tjenesteområde, for eksempel én nøkkel for videregående opplæring, én for tannhelse og én for samferdsel. Delkostnadsnøkkelen uttrykker det relative utgiftsbehovet mellom fylkeskommunene innenfor det aktuelle tjenesteområdet, gitt de valgte kriteriene og vektene. Nøklene uttrykker et normert kostnadsnivå gitt strukturelle forskjeller mellom primærkommuner og fylkeskommuner.

Når delkostnadsnøklene er fastsatt, kombineres de til én samlet kostnadsnøkkel for fylkeskommunene. Vektingen mellom delkostnadsnøklene skjer via delkostnadsvekter, som normalt baseres på tjenesteområdenes andel av fylkeskommunenes samlede netto driftsutgifter. Dette innebærer at videregående opplæring får høyere vekt enn tannhelse, siden utgiftene til videregående opplæring utgjør en større andel av de samlede utgiftene i fylkeskommunene.

Summen av delkostnadsvektene danner den samlede kostnadsvekten, som sikrer at den samlede kostnadsnøkkelen gir et helhetlig uttrykk for fylkeskommunenes beregnede utgiftsbehov. Det samlede tilskuddet til hver fylkeskommune framkommer ved å multiplisere det beregnede utgiftsbehovet per innbygger med antall innbyggere i fylket. Den ferdige kostnadsnøkkelen brukes deretter i inntektssystemet til å fordele rammetilskudd og utjevne strukturelle kostnadsforskjeller mellom fylkeskommunene.

3 Tidligere vurderinger av håndteringen av få enheter

Vi vil nå gjengi tidligere vurderinger av håndteringen av få enheter i utgiftsutjevningen i fylkeskommunenes inntektssystem og derigjennom hvordan systemet har endret seg over tid. Vi gjennomgår behandlingen til Rattsø-utvalget (NOU 1996:1) i delkapittel 3.1, Borge-utvalget (NOU 2005:18) i delkapittel 3.2, Inntektssystemutvalget NOU (2022:10) i delkapittel 3.3 og ekspertutvalget for fylkeskommunenes inntektssystem (Kommunal- og distriktsdepartementet 2022) i delkapittel 3.4.

3.1 Rattsø-utvalget

Rattsø-utvalget (NOU 1996:1) peker eksplisitt på flere utfordringer ved å beregne kostnads- og kriterienøkler for fylkeskommunene. Utvalget understreker at antallet fylkeskommuner er lavt, noe som gir svært få observasjoner i empiriske analyser. Få enheter gir svakt statistisk grunnlag for å estimere statistisk robuste sammenhenger mellom utgifter og forklaringsvariabler, skille reelle kostnadsforskjeller fra tilfeldige variasjoner, og teste og validere kriterier på samme måte som for primærkommunene. Konsekvensen er at regresjonsbaserte metoder gir usikre resultater, og at små endringer i data kan gi store utslag i nøklene.

Som følge av det svake datagrunnlaget framholder utvalget at fastsettelsen av kriterienøklene for fylkeskommunene i større grad må bygge på faglige vurderinger og skjønn, snarere enn rene statistiske estimater. Rattsø-utvalget er tydelig på skjønnsmessige vurderinger av inkludering av kriterier og vektingen mellom kriterier og mellom sektorer utgjør en metodisk begrensning snarere enn et normativt ønske. Utvalget påpeker også at enkelte fylker kan få utslag som i liten grad reflekterer reelle kostnadsforskjeller og at fordelingsvirkningene blir sensitive for tekniske valg. Dette brukes som et argument for forsiktighet i ambisjonsnivået for hvor presist fylkeskommunenes utgiftsbehov kan måles.

3.2 Borge-utvalget

Borge-utvalget (NOU 2005:18) gjennomgår utgiftsutjevningen i lys av strukturelle endringer, ny forskning og nye data som har kommet til siden Rattsø-utvalget. En kritikk mot det tidligere utvalget var bruken av partielle fremfor simultane analyser, noe som risikerte å undervurdere effekter som virker på tvers av sektorer, eksempelvis bosettingsmønstre. Borge-utvalget imøtekommer dette for primærkommunene ved partielle analyser, i tillegg til å ta i bruk Statistisk sentralbyrås simultane modell for variasjoner i kommunenes utgifter per innbygger, KOMMODE. For fylkeskommunene fantes det imidlertid ingen tilsvarende simultan modell.

Utvalget understreker at analyser av fylkeskommunenes ressursbruk er metodisk krevende fordi det er svært få enheter å analysere, noe som gir liten variasjon i dataene og gjør resultatene fra tradisjonelle regresjonsanalyser usikre. For å håndtere utfordringen med få observasjoner og likevel fange opp kostnadsulemper knyttet til geografi, har utvalget valgt å tilpasse metodebruken for de ulike sektorene. Ved å analysere kostnadsdata for enkeltskoler får man nok observasjoner

til å identifisere skalaegenskaper og effekten av skolestruktur, som deretter kan aggregeres opp for å forklare utgiftsbehovet på fylkesnivå. Utvalgets egne analyser ble gjennomført på fylkesnivå, men de brukte også en analyse fra Borge, Naz og Tovmo (2003) som omgikk problemet med få fylker ved å flytte analysenivået ned til skolenivå. Det endelige forslaget til kriterievariabler og kriterievekter var imidlertid basert på gjennomsnittskostnader på landsnivå og ikke på regresjonsanalysene på skolenivå.

Når det gjelder tannhelsetjenesten, støtter utvalget seg på en ressurskartlegging rettet mot brukere for å vekte alderskriteriene og antall psykisk utviklingshemmede over 16 år. For fylkesveier har utvalget forkastet statistisk metode til fordel for en normativ modell (MOTIV) som beregner behovet basert på tekniske standarder. Det er kun for rutetrafikken under samferdsel, omtalt i utredningen som lokale ruter, at utvalget primært baserer seg på regresjonsanalyser på fylkesnivå, til tross for de metodiske svakhetene, da det her manglet gode alternativer. Utvalget påpeker imidlertid at datagrunnlaget er svakt og at resultatene må tolkes med varsomhet.

3.3 Inntektssystemutvalget

Inntektssystemutvalget (NOU 2022:10), også kjent som Haraldsvik-utvalget, har mandat til å spesifikt se på primærkommunene. Derfor drøftes ikke problematikken rundt få enheter.

Utvalget drøfter Inntektssystemutvalget metodevalg for utgiftsutjevningen og konkluderer med at statistiske metoder bør være hovedgrunnlaget for utvikling og oppdatering av kostnadsnøkler. Det vurderer både simultane og partielle analyser og framhever at slike analyser må fange opp samspill mellom tjenesteområder samtidig som forklaringskraften ivaretas. Normative metoder anses å ha en begrenset rolle, ettersom empirisk baserte statistiske analyser med oppdatert datagrunnlag gir et mer treffsikkert og robust estimat av kommunenes utgiftsbehov. Inntektssystemutvalget påpeker også at valg av datagrunnlag, håndtering av ekstremverdier og bruk av flere årganger er viktige for robuste analyser. Det understreker at kriterier og vekter bør oppdateres jevnlig for å reflektere faktiske variasjoner i ufrivillige kostnadsforskjeller.

3.4 Ekspertutvalget for fylkeskommunens inntektssystem

Ekspertutvalget for fylkeskommunens inntektssystem (Kommunal- og distriktsdepartementet 2022), også kjent som Lunder-utvalget, skulle komme med forslag til et nytt inntektssystem som kunne tre i kraft fra 2024. Dette samsvare med begynnelsen for den nye fylkesinndelingen vi har i dag.

Utvalget diskuterer spesifikt utfordringen bak å bruke statistiske metoder for å fastsette delkostnadsnøkkelen til fylkeskommunen grunnet få enheter. De trekker fram fire hovedutfordringer:

- Enheter med stor innflytelse: Enkelte fylker kan ha uforholdsmessig stor påvirkning på resultater.
- Multikollinearitet: Sterk samvariasjon mellom uavhengige variabler kan gjøre det vanskelig å isolere effekter fra enkelt variabler og begrenser variabler de kan inkludere som kontroll.
- Statistisk signifikans: Færre enheter gir dårligere mål for statistisk signifikans.
- Forklaringskraft: Mål som R^2 kan bli høye selv om modellen treffer dårlig for enkelte fylker.

Utvalget er tydelige på at man må tolke resultatene fra statistiske metoder varsomt i arbeidet med fylkeskommunene. For å håndtere utfordringene har de testet modellenes robusthet ved utføre analyse der man utelater enkelte fylker som skiller seg ut og ser om effekten fortsatt er der. Utvalget påpeker allikevel at det ikke er aktuelt å utelate fylker fra den endelige beregningen av nøklene, fordi det svekker datagrunnlaget ytterligere.

For å komme seg rundt få enheter bruker de ikke data i perioden etter 2020 ettersom det kun er 11 enheter da. De ser på tverssnittdata fra 2017, 2018 og 2019. Disse analyserer de mot fylkesstrukturen i 2019, og aggregert opp mot den nye fylkesstrukturen i 2024. For kollektivtransporten (fylkesveiferger) anbefaler utvalget at man ser på buss, bane og båter som en sektor. Utvalgets viktigste argument for å se båt i sammenheng med buss og bane var å unngå å bruke den normative modellen som ble brukt for båter i årene 2022 til 2023, men det spiller også inn på håndteringen av utfordringene med få enheter.. Det mener dette gir bedre forklaringskraft enn å se på buss og bane og båt hver for seg. For fylkesvei anbefaler utvalget en normativ modell fra Senter for økonomisk forskning for å sette delkostnadsnøkkel. Utvalget påpeker at dette også kan gjøres for fylkesveiferger med en normativ modell utviklet av Transportøkonomisk institutt (Rødseth med flere 2022b, se også Svendsen med flere 2017).

4 Håndtering av få enheter i andre inntektssystemer

Det vil være hensiktsmessig å inkludere en gjennomgang av analysemetoder som er brukt i utformingen av utjevningssystemet i andres sektorer og land. I det følgende vil vi kort gjennomgå hvordan problematikken knyttet til få enheter håndteres i andre inntektssystemer. Vi gir innblikk i andre sektorers praksiser i delkapittel 4.1 og andre lands praksiser i delkapittel 4.2.

4.1 Praksiser i andre sektorer

I det følgende vil vi kort gjennomgå inntektsmodellen til helseforetakene og ressursallokeringsmodellen til politiet.

4.1.1 Helseforetakene

Basisbevilgningen fra staten fordeles mellom de fire regionale helseforetakene gjennom en inntektsmodell som skal sikre en behovs- og kostnadsjustert fordeling. Modellen bygger på forslagene i NOU (2019:24). Behovsanalysene forklarer forskjeller i behov gjennom alder, sosioøkonomiske forhold, helserelaterte forhold og bostedsrelaterte forhold. Det er gjennomført separate analyser for somatiske tjenester, psykisk helsevern, rusbehandling og prehospitale tjenester. Analysene er i hovedsak gjennomført på individnivå med data om faktisk tjenesteforbruk, før de estimerte sammenhengene anvendes på befolkningssammensetningen i geografiske enheter slik at behov beregnes på aggregert nivå.

Kostnadsanalysene er derimot gjennomført på helseforetaksnivå ved bruk av minste kvadraters metode med robust estimering av standardfeil. Kostnadsindeksen viser forholdet mellom hvert helseforetak og det nasjonale gjennomsnittet, hvor helseforetakene vektes med sin andel av den samlede aktiviteten. Antall observasjoner i kostnadsanalysene varierer mellom tjenesteområdene: for somatikk er antallet 66, mens andre analyser baserer seg på mellom 20 og 40 observasjoner. Utvalget diskuterer i begrenset grad problematikken knyttet til få observasjoner, utover å påpeke at analysene bygger på et begrenset antall observasjoner.

4.1.2 Politiet

Politidirektoratet innførte i 2022 en ressursallokeringsmodell (RAM) for ressursfordelingen mellom landets 12 politidistrikter. RAM-modellen har som formål å fordele budsjettrammen basert på objektive kriterier som fanger opp forskjeller, herunder forskjeller i geografi, demografi og kriminalitetsbelastning med mer (Finansdepartementet og Justis- og beredskapsdepartementet 2024). RAM-modellen inkluderer også en rekke sosioøkonomiske forhold som er tett knyttet til geografi. Politiet har ikke noe godt mål for beredskap, så politiets beredskapsaktivitet anslås ved ulike kontekstuelle variabler. Modellen inkluderer kriteriet landareal per innbygger for å ta hensyn til geografiske forhold som større avstander og vanskelig fremkommelighet gjennom. RAM-modellen innebærer dessuten grupperinger av store og små regioner med betydning for tilskudd,

der noen fylkeskommuner nødvendigvis er i randsonen for å bli definert som liten eller stor. Vista Analyse med flere (2026) finner at distrikter som har budsjettkniven på strupen, har vært relativt effektive, og vice versa.

4.2 Praksis i andre land

I det følgende redegjør vi kort for praksiser for å utforme kriterier kostnads- og utgiftsutjevning i utvalgte andre land, inkludert Sverige, Sveits og Australia.

4.2.1 Sverige

I Sverige benyttes et formelbasert system for utgiftsutjevningen mellom landets 21 regioner med demografiske, sosioøkonomiske og geografiske kostnadsindikatorer med sektorielle delmodeller (SOU 2024:50a og 2024:50b). Et formelbasert system innebærer her at utgiftsutjevningen er styrt av eksplisitte beregningsregler med faste vekter. Kostnadsutjevningen fastsettes mellom kommuner og regioner (fylkeskommuner) gjennom en standardkostnadsbasert modell. Systemet bygger på beregning av teoretiske standardkostnader for ulike sektorer, basert på gjennomsnittlige kostnadsnivåer i landet og kostnadsdrivere som befolkningssammensetning, bosettingsstruktur og sosioøkonomiske forhold. For regionene er det to modeller – en modell for helse og sykehus og befolkningsforandringer og en modellen for kollektivtransport, som er felles for kommunene og regionene (SCB 2024).

Modellen for helse- og sykehustjenester er behovs- og matrisebasert der man ved matrisebaserte modeller beregner behovene ved hjelp av kryssklassifiserte grupper med egne kostnader. Befolkningen deles inn i grupper basert på blant annet alder, inntekt, utdanning og husholdningstype, og hver gruppe tilordnes en gjennomsnittlig kostnad basert på historiske kostnadsdata fra Region Skåne, Västra Götalandsregionen og Region Norrbotten. I tillegg til resultatet fra matrisemodellen beregnes det tillegg eller fradrag for strukturelle merkostnader knyttet til lønnsnivå, bemanning i distriktsområder og drift av helsetjenester i distriktsområder.

Merkostnadskomponenten for spredt bosetting beregnes ved hjelp av en modell med fiktiv lokalisering av sykehus, helsesentre og ambulansestasjoner. Dette sikrer at beregningene er uavhengige av politiske eller organisatoriske lokaliseringsvalg og dermed ikke kan påvirkes av regionene selv. Blant annet gis det kompensasjon for små sykehus i tilfeller der det er færre enn 100 000 innbyggere innenfor et opptaksområde med 100 kilometers radius (SOU 2024:50a).

Modellen for befolkningsendringer ble innført i 2020 og skulle kompensere for sterk befolkningsvekst- og nedgang. Utjämningskommittén finner ingen dokumentasjon for at befolkningsendringer har gitt merkostnader utover det som allerede fanges opp i eksisterende modeller. De anbefaler derfor å avvikle hele delmodellen (SOU 2024:50a).

Modellen for kollektivtransport beregnes på länsnivå og er felles for kommuner og regioner. Utvalget understreker at det er metodisk krevende å bruke statistiske analyser når antallet observasjoner er så lavt, fordi enkeltregioner kan få stor betydning for resultatene. Av denne grunn benyttes en fast regresjonsbasert modell, der vektene ikke oppdateres årlig, men der man legger vekt på strukturelle, lite påvirkbare variabler som tettstedsbefolkning, pendling over kommunegrenser og gjennomsnittlige avstander (SOU 2024:50a).

4.2.2 Sveits

Kostnadsutjevningen mellom kantonene i Sveits består av to deler – en del for kompensasjon for geografiske ulemper og en for kompensasjon for sosioøkonomiske ulemper. De geografiske ulempene er knyttet til forhold som høyde over havet, bratt terreng og lav befolkningstetthet. De sosioøkonomiske forholdene omfatter fattigdom, aldrende befolkning, innvandring og «kjerneby»-funksjoner (Federal Finance Administration 2025). Kompensasjonen er avgrenset til å dekke ekstraordinære kostnader som enkelte kantoner antas å ha ved finansiering av offentlige oppgaver sammenliknet med det sveitsiske gjennomsnittet. Føderale myndigheter fastsetter en samlet ramme for hvor mye som skal gå til kompensasjon av geografiske og sosioøkonomiske ulemper, og midlene fordeles på grunnlag av indekser (Dougherty med flere 2025).

For de geografiske ulempene sammenlikner man for eksempel andel av befolkning i kantonene som bor over 800 meter over havet med landsgjennomsnittet. Den samlede rammen som er satt av til denne faktoren fordeles mellom kantonene som ligger over landsgjennomsnittet, etter hvor stor deres relative belastning er sammenliknet med de andre kantonene som ligger over snittet. Utjevningen for sosioøkonomiske ulemper beregnes på omtrent samme måte. Man bruker landsgjennomsnittet for de ulike faktorene og standardavviket for å lage en indeks per kanton. Den totale rammen for sosioøkonomiske ulemper fordeles basert på denne indeksen (Eidgenössische Finanzverwaltung 2025). Det benyttes altså ikke regresjonsanalyser basert på faktiske utgifter i beregningene, og vektene justeres ikke årlig, men evalueres periodisk.

4.2.3 Australia

Australia opererer med prinsipper for horisontal og fiskal utjevning. Prinsippene fordeler inntekter mellom delstater basert på skatteinntekter og utgiftsbehov, slik at alle har sammenliknbar evne til å levere tjenester. Metoden er formelbasert med et betydelig datagrunnlag i bunn, med innslag av faglige vurderinger (Commonwealth Grants Commission 2021).

De samler inn data på de totale utgiftene for alle delstatene innen hvert tjenesteområde, og summerer disse for å fastsette et samlet og deretter et gjennomsnittlig utgiftsnivå på nasjonalt nivå. Deretter kalkulerer de «assessed expenses», det vil si det utgiftsnivået en stat ville hatt dersom den fulgte en gjennomsnittlig politikk, justert for strukturelle kostnadsdrivere som staten ikke selv kan påvirke, og forutsatt gjennomsnittlig effektivitet i tjenesteproduksjonen.

Kostnadsdriverne estimeres ved hjelp av regresjonsanalyser og beregninger av gjennomsnittlige kostnader for relevante delpopulasjoner. Metoden anvendes på en rekke tjenesteområder, herunder grunnskole, høyere utdanning, helse, bolig, velferd, tjenester til lokalsamfunn, rettsvesen, veier, transport og øvrige offentlige utgifter. For eksempel benyttes regresjonsmodeller til å anslå kostnader per elev og merkostnader knyttet til elever i mer avsidesliggende områder. Disse enhetskostnadene aggregeres deretter ved å multiplisere dem med antall elever i de aktuelle delpopulasjonene. Ved å basere beregningene på enhetskostnader, fremfor regresjoner av samlede kostnader til for eksempel grunnskole, unngår man metodiske utfordringer knyttet til et begrenset antall enheter i datagrunnlaget (Commonwealth Grants Commission 2025).

5 Metodikk for fastsettelse av systemets kriterier

Dette kapittelet beskriver hvordan metodikken kan ta hensyn til få enheter i fastsettelsen av kriteriene i fylkeskommunenes inntektssystem. I delkapittel 5.1 vurderer vi potensialet som ligger i utnyttelsen av statistisk metodikk i fastsettelsen av kriteriene, og hva slags statistiske metoder som kan egne seg. Deretter vurderer vi alternativ evalueringsmetodikk, som kan benyttes når statistiske metoder er mindre egnet, i delkapittel 5.2.

5.1 Statistisk metodikk

I den følgende vil vi gjennomgå fire former for mulig statistisk metodikk for å fastslå kriterievektene i fylkeskommunenes inntektssystem. Disse inkluderer lineære regresjoner, maskinlæring, Bayesiansk estimering og frontestimering. Vi begynner imidlertid med å diskutere datasettets betydning for potensialet for regresjonsanalyse, herunder datasett med kun en tidsperiode (tverrsnittsdata) og datasett med flere tidsperioder (paneldatasett).

Vi viser til Stock og Watson (2020) for en innføring i økonometriske metoder og Greene (2018) for en omfattende gjennomgang. Videre refererer vi til Basuchoudhary, Bang og Sen (2017) for en nærmere omtale av maskinlæringsteknikker i en økonomisk kontekst.

5.1.1 Regresjoner på tverrsnittsdata

For primærkommunene bestemmes kriterievektene i hovedsak av lineære regresjoner på tversnittdata, altså regresjoner med kun en tidsperiode. Det er flere begrensninger ved å kjøre lineære regresjoner med for få observasjoner. Med få enheter kan det være vanskelig å kjøre selv enkle lineære regresjoner med alle kriteriene for en tjenestefunksjon på en gang. Regresjoner på tverrsnittsdata har derfor lite for seg på fylkeskommunalt nivå, jo mindre det dreier seg om svært få variabler og stabile sammenhenger. Få enheter innebærer begrenset statistisk styrke i estimeringen med få frihetsgrader, som gjør det vanskelig å korrigere for forskjeller i geografi og oppgaveportefølje. Videre vil estimatene ha betydelig varians og usikkerhet med stor sårbarhet for støy i dataene. Tett korrelasjon mellom flere forklaringsvariabler, såkalt multikollinearitet, gjør det vanskelig å skille mellom ulike virkninger på en utfallsvariabel. Siden tversnittanalyser kun utnytter kun variasjonen mellom enhetene, gir de ingen mulighet til å skille mellom varige strukturelle forskjeller og midlertidige utslag over tid.

Et alternativ er å kjøre en enkel regresjon hvor man inkluderer kun sentrale kontrollvariabler. En annen tilnærming er å basere seg på sammenhenger som er dokumentert i datasett med flere observasjoner, som for eksempel data fra de underliggende tjenesteenhetene eller primærkommunene. Alternativ kan man betinge på sammenhenger eller preestimerte variabler som man har funnet fram til i rikere datasett, som for eksempel datasett for de underliggende tjenesteenhetene eller primærkommunene. Uavhengig av hvilken tilnærming man velger, vil det å ikke inkludere alle variablene i samme regresjon innebære et tap av estimeringseffektivitet. Grunnen er at regresjonene da ikke får tatt hensyn til hvordan variablene samvarierer med hverandre.

5.1.2 Regresjoner på paneldata

Et tredje alternativ er å utnytte tidsvariasjon i tillegg til tverrsnittvariasjon mellom fylkeskommunene. Dersom datasettet utvides til et paneldatasett, vil tidsdimensjonen inngå i dataene og man vil øke antallet observasjoner i regresjonene. Dette sporet etter vår vurdering det mest lovende. Et paneldatasett følger de samme fylkeskommunene over tid. Paneldatasett gjør det mulig å kontrollere for årlige endringer, faste forskjeller mellom fylkeskommunene og uobserverte forskjeller mellom enheter som ikke korrelerer med forklaringsvariablene. Dermed gir paneldatasett et rikere informasjonsgrunnlag for å vurdere hvordan ulike kriterier samvarierer med driftsutgiftene, hvilket er spesielt kritisk når antall observasjonsenheter er få.

En begrensing ved paneldataanalyser er at mange av de relevante kriteriekandidatene i inntektssystemet vil være nokså persistente over tid. For eksempel er det lite variasjoner i avstander mellom befolkningen over tid. I så fall blir det begrenset hva man klarer å få ut av tidsvariasjonen. Paneldataanalyser mister noe av sine konkurransefortrinn overfor tversnittanalyser over flere år når tidsdimensjonen er svak, inkonsistent eller lite informativ, eller når analysen primært handler om nivåforskjeller snarere enn utviklingsprosesser. Som diskutert i delkapittel 3.4, ble gjentatte tverrsnittsregresjoner valgt framfor paneldataanalyser av Ekspertutvalget for fylkeskommunens inntektssystem i utvalgets utforskende regresjoner.

5.1.3 Lineær regresjonsanalyse

I sin enkleste form estimeres lineære regresjoner ved hjelp av ordinære minste kvadraters metode (OLS), som forutsetter at restleddene har konstant varians og er ukorrelerte både på tvers av enheter og over tid. Når disse forutsetningene ikke holder vil OLS-estimatene fortsatt være forventningsrette, men ikke effektive. Det kan dreie seg om regresjonens restledd er korrelert med forklaringsvariablene (heteroskedastisitet), restleddet er korrelert med tidligere perioders restledd (autokorrelasjon) eller en kombinasjon. I slike tilfeller er generaliserte minste kvadraters metode (GLS) bedre egnet, ettersom denne metoden eksplisitt tar hensyn til strukturen i restleddene og eventuelle avhengigheter. Ved bruk av lineære regresjonsmodeller anbefales det derfor å benytte GLS fremfor OLS.

Paneldata åpner også for at man kan operere med fylkesfaste effekter og tidsfaste effekter, altså en såkalt regresjon med faste effekter. Samtidig er det ikke hensiktsmessig å inkludere faste enhetseffekter for alle fylkeskommuner i regresjonene, ettersom dette i praksis vil absorbere mye av den strukturelle variasjonen som kriteriene nettopp er ment å fange opp. Datastrukturen tillater også regresjoner med randomiserte effekter, der restleddet dekomponeres i en fylkesspesifikk tilfeldig effekt og et tidsvarierende feilledd. Modellen kan ses som et mellomtilfelle mellom pooled OLS og faste effekter, ved at den utnytter både variasjon over tid innen enheter og mellom enheter, men uten å eliminere all tidsinvariant heterogenitet slik fasteffektsmodellen gjør.

Dersom det er en del variasjon i variablene over tid, er det mulig å skille effektene av ulike forklaringsvariabler i form av kriterievariabler og kontroller (som frie inntekter) fra hverandre, slik at man oppnår mer presise anslag på virkningene av hver kriterievariabel og dermed mer nøyaktige vekter. Inkludering av randomiserte effekter kan vurderes i tilfeller der datamaterialet preges av betydelig støy knyttet til enkeltfylker, og der dette anses som en hensiktsmessig måte å håndtere uobserverte, fylkesspesifikke forhold på.

Når antallet observasjoner er svært begrenset, kan det føre tet brudd på forutsetningen om at restleddene i lineære regresjoner er normalfordelte. Med få observasjoner blir det vanskelig å oppnå en stabil og representativ fordeling av residualene, noe som kan gi skjevheter i estimatene og feil i teststatistikkene. Dermed kan konfidensintervaller og hypotesetester, som legger normalitetsantakelsen til grunn, bli misvisende. Brudd på normalitetsantakelsen i restleddet påvirker i hovedsak inferens i små utvalg. Det kan håndteres ved å benytte robuste standardavvik (standardavvik som justeres for heteroskedastisitet), klustrede standardavvik (standardavvik som tillater korrelasjon innen samme observasjonsklynge, men ikke mellom klynger) eller bootstrap-baserte inferensmetoder.

Det faktum at fylkesinndelingen har endret seg over tid, kompliserer bildet noe. Dette kan likevel håndteres ved å behandle gamle og nye fylker som separate enheter, eller ved å gjøre grep for å balansere datasettet, jamfør delkapittel 6.1.

5.1.4 Maskinlæring

I senere år har maskinlæringsmetoder fått økt utbredelse i empiriske analyser. Maskinlæring omfatter en klasse statistiske og algoritmiske tilnærminger der prediktive sammenhenger estimeres ved at modellen identifiserer mønstre i dataene direkte, uten at sammenhengens funksjonsform eller struktur fastlegges på forhånd av analytikeren. Metodene er særlig egnet til å håndtere komplekse sammenhenger og et stort antall potensielle forklaringsvariabler.

En mye brukt metode i denne sammenhengen er Least Absolute Shrinkage and Selection Operator (LASSO), som kan anvendes til å estimere lineære modeller med mange forklaringsvariabler, inkludert polynomledd og interaksjoner. LASSO påfører en straff på koeffisientenes absoluttverdi, noe som både demper størrelsen på koeffisientene og bidrar til automatisk variabelseleksjon ved at mindre relevante forklaringsvariabler settes til null. Denne egenskapen gjør metoden godt egnet til å begrense overtilpasning i datasett med mange forklaringsvariabler. Samtidig gir LASSO forventningsskjeve estimater av koeffisientene, noe som innebærer at de ikke kan tolkes direkte som marginale effekter eller brukes til inferens uten videre justering. Straffingen gir noe beskyttelse mot store koeffisienter, men metoden er ikke fullt robust overfor ekstreme observasjoner, som fortsatt kan påvirke resultatene.

I situasjoner der både antall enheter og tidsdimensjonen er begrenset, kan det være aktuelt å vurdere metodikk som eksplisitt er utviklet for små datasett. Dette gjelder blant annet shrinkage-baserte regresjonsmetoder utover LASSO, som ridge-regresjon og elastic net, der koeffisientene trekkes mot null på en måte som reduserer variansen i estimatene uten nødvendigvis å eliminere forklaringsvariabler fullstendig. Disse metodene gjør også en avveining mellom å oppnå forventingsrette estimater og å straffe store koeffisienter for å redusere effekten av potensielle utliggere og overtilpasning. Slike metoder kan være særlig hensiktsmessige når forklaringsvariablene er tett korrelerte

For å gjøre koeffisientene fra estimering med shrinkage-baserte metodene sammenliknbare med koeffisientene fra en konvensjonell lineær regresjon og dermed også kriterievektene, kan man transformere dem tilbake til opprinnelig skala. Dette gjøres ved å multiplisere hver koeffisient med standardavviket til den avhengige variabelen og dividere med standardavviket til den respektive forklaringsvariabelen, forutsatt at estimeringen er gjennomført på standardiserte data. Maskinlæring kan også bidra til å stabilisere estimeringen fremfor å gjennomføre hard

variabelseleksjon. En nærmere diskusjon om bruk av maskinlæringsmetoder i valg av variabler er gitt i delkapittel 6.1.

Det kan også diskuteres om man bør fortsette å basere estimeringen på konvensjonelle regresjonsanalyse minste kvadraters metode, som baserer seg på gjennomsnittsbetraktninger, eller om det ville ha vært bedre å ta utgangspunkt i metodikk som straffer store avvik. Dette kommer an på om hvorvidt man er opptatt av å vekte alle fylkeskommunene likt, eller om man vil vekte betydningen av fylkeskommuner som avviker mye fra resten. Dersom hybride modeller brukes i fastsettelse av kriterievektene, kan eventuelle kriterieutformingen avvikende fylkeskommuner heller håndteres med normative metodikk eller ressurskartlegginger. Vi kommer til disse metodiske tilnærmingene i delkapittel 5.2.

5.1.5 Bayesiansk regresjonsanalyse

En annen tilnærming er bayesiansk regresjonsestimering, der man legger til grunne sannsynlighetsfordelinger for parameterne, som deretter modifiseres av dataene til posteriorfordelinger. I motsetning til konvensjonell regresjonsanalyse, der parameterne behandles som faste, men ukjente størrelser, modelleres parameterusikkerheten eksplisitt i bayesianske modeller. Dette kan være særlig hensiktsmessig i små datasett, der datagrunnlaget alene gir svakt identifiserte sammenhenger og stor usikkerhet i punktestimatene.

Priorer et grunnleggende konsept ved bayesiansk estimering. Med en prior menes en sannsynlighetsfordeling som uttrykker hva man antar om en parameters verdi før dataene observeres. Prioren kan reflektere tidligere empiri, teoretiske forventninger eller generelle antakelser om fortegn og størrelsesorden, og fungerer som et utgangspunkt for den bayesianske analysen. En informativ prior er en prior som inneholder relativt mye informasjon og dermed legger tydelige føringer på hvilke parameterverdier som anses som plausible. Slike priorer kan være basert på tidligere studier, institusjonell kunnskap eller sterke teoretiske forventninger, og vil ha merkbar innflytelse på posteriorfordelingen, særlig når datagrunnlaget er begrenset. I motsetning til dette gir svakt informative priorer kun milde begrensninger og er primært ment å stabilisere estimeringen uten å dominere resultatene.

Ved å benytte svakt informative priorer, eller priorer som reflekterer rimelige antakelser om fortegn, størrelsesorden eller relative forskjeller mellom kriterievariabler, kan man oppnå mer stabile estimater uten å påtvinge modellen harde restriksjoner. Bayesianske modeller kan videre utvides til hierarkiske spesifikasjoner, der parametere tillates å variere mellom fylkeskommuner eller over tid, samtidig som estimatene delvis trekkes mot et felles gjennomsnitt. Slike strukturer innebærer en eksplisitt avveiing mellom at estimatene skal være forventningsrette og ha lavest mulig varians. De kan være særlig egnet når antall observasjonsenheter er begrenset og når man samtidig ønsker å ta hensyn til uobserverte og fylkesspesifikke forhold. Bayesianske modeller er særlig egnet når datagrunnlaget er begrenset. Dette skyldes at de kan kombinere eksisterende kunnskap i form av priorer med observasjonsdata, hvilket gjerne gir relativt stabile og realistiske estimater.

5.1.6 Frontestimering

Strengt tatt innbefatter flere av kriterieregresjonene i inntektssystemet en spesifikasjonsfeil ved at de ikke tar har hensyn til produksjonsprosessen knyttet til kommunal tjenesteproduksjon. For videregående skole har man indikasjoner på produksjonens volum i elevtallet, som likevel er utenfor kommunenes kontroll. Det kontrolleres like fullt ikke for kvaliteten i tjenesten. Nivået på (de korrigerte) bruttodriftutgiftene per innbygger skyldes ikke alltid kostnadsnivået, men kan også ha med nivået på produksjonen å gjøre. Dette kunne man håndtert ved å inkludere produksjonen som kontroller i de lineære kriterieregresjonene. Disse funksjonene kan samtidig inkludere kriteriene som forklaringsvariabler.

Det vil imidlertid være enda mer nærliggende å ta høyde for at ikke alle fylkeskommunene er på produktivitetsfronten og at det er forskjell mellom beste praksis og effektiviteten i form at avstanden til beste praksis. Dette kan gjøres med såkalte frontestimeringsmetoder som dataomslutningsanalyse (DEA), stokastisk frontanalyse (SFA), Stochastic Nonparametric Envelopment of Data (StoNED) og kernellregresjoner. Frontestimeringsmetodene er ikke egnet for tverrsnittsdata med svært få observasjoner, men er aktuelt for paneldatasett.

Kommunal- og distriktsdepartementet og Teknisk beregningsutvalg for kommunal og fylkeskommunal økonomi bruker allerede DEA i sine effektivitetsanalyser av kommunal sektor. Metoden er ikke-parametrisk og deterministisk, hvilket innebærer at den ikke forutsetter en bestemt funksjonsform og ikke inkluderer et restledd for støy. SFA er derimot en parametrisk og stokastisk metode, hvilket innebærer at den forutsetter form på produksjonsfunksjonen og kostnadsfunksjonen. StoNED er generalisering av begge metodene, som både har DEAs styrke med ikke-parametrisk front og stokastiske støyledd. Kernelregresjoner er en ikke-parametrisk metode som bruker en vektningsfunksjon (kalt kernelfunksjonen) til å gi nærmere observasjoner større vekt enn fjernere observasjoner når man estimerer gjennomsnittlig verdi av en utfallsvariabel for et gitt nivå på forklaringsvariabler.

Effektivitetsestimering av kostnadsfunksjoner vil ved inkludering av kontekstuelle variabler i en totrinnsregresjon estimere akkurat de vektene man er på jakt etter i inntektssystemet. Felles for alle disse metodene er at de kan håndtere innflytelser av omgivelsene. Helst bør estimeringen av effektiviteten og kriterienes innvirkning på (de korrigerte) bruttodriftsutgiftene estimeres simultant for ikke å tape estimeringseffektivitet. Ved bruk av SFA og StoNED kan man typisk finne virkningene av omgivelsene på kostnadene direkte i regresjonene. Når det gjelder DEA, er Simar og Wilsons (2007) totrinnstilnærming for estimering av omgivelsenes innvirkning et aktuelt alternativ. I denne tilnærmingen estimeres kostnadsteknologien (eller produksjonsteknologien) i første steg, mens omgivelsenes innvirkning på kostnadsineffektiviteten estimeres i andre trinn. Effekten på faktiske kostnader kan deretter utledes ved å skalere ineffektiviteten tilbake til kostnadsnivået, slik at de estimerte koeffisientene implisitt gir marginale effekter på kostnader gjennom deres effekt på avstanden til kostnadsfronten. Vi viser til Rødseth med flere (2022a) for en mer detaljert gjennomgang av frontestimeringspotensialet til kommunal sektor.

Dersom produksjonen og produksjonsinnsatsen bestemmes samtidig, kan det oppstå såkalte simultane forventningsskjevheter i estimeringen ved at kausale sammenhengene mellom kostnadene og produksjonen går begge veier. Dette kan håndteres på flere måter, der den mest utbredte i dag innenfor konvensjonell økonometri er kontrollfunksjonsmetodikk. Vi viser til Van Beveren (2012) og Rødseth, Kuosmanen og Holmen (2025) for mer om kontrollfunksjonsmetodikk innenfor henholdsvis konvensjonell økonometri og frontestimering.

5.2 Alternativ metodikk

Statistiske metoder vil i utgangspunktet være mest egnet til å fastslå kriterievektene. I praksis kan det imidlertid være utfordrende å estimere slike ulemper på en pålitelig måte, for eksempel på grunn av komplekse sammenhenger, begrenset antall observasjoner eller manglende data. Kriterier som bygger på usikre estimater, vil derfor ofte være lite hensiktsmessige å innføre. I slike tilfeller kan alternative tilnærminger vurderes, som ressurskartlegging, normative vurderinger eller hybride metoder, hvilket vi vil redegjøre nærmere for i det følgende. Fellesnevneren ved disse metodene er at de kan være velegnede når datagrunnlaget for statistisk modellering ikke vurderes som tilfredsstillende, men man likevel ønsker et empirisk grunnlag. For eksempel ved få observasjoner eller at fylkene har innflytelse over utfallene.

5.2.1 Ressurskartlegging

I enkelte tilfeller beregnes kriteriene i kostnadsnøklene ut fra gjennomsnittlig ressursbruk eller kostnader knyttet til spesifikke brukergrupper eller tjenester. I dagens inntektssystem gjelder dette for eksempel tannhelsenøkkelen og enkelte kriterier innen videregående opplæring, som søkere til læreplass og søkere til høykostnadsutdanningsprogrammer. En variant er ressurskartlegginger, der man ønsker å måle faktisk ressursbruk per enhet, for eksempel basert på spørreundersøkelser, casestudier, effektstudier, ekspertvurderinger eller regnskapsanalyse for tjenestefunksjonene. Ved å bruke data om ressursinnsatsen for nøye definerte grupper, kan man få et relativt presist mål på gjennomsnittskostnadene knyttet til ulike tjenester og brukergrupper. Disse gjennomsnittskostnadene kan deretter brukes til å fastsette vektene i kostnadsnøklene, for eksempel basert på antall innbyggere i prioriterte brukergrupper innen tannhelsetjenesten.

En begrensning ved denne typen ressurskartlegginger er at de ikke fanger opp andre faktorer som kan bidra til variasjon mellom fylkeskommunene. Ofte samvarierer ulike kostnadsdrivere, for eksempel andelen innbyggere i ulike aldersgrupper og strukturelle forhold som reiseavstander eller bosettingsmønster. Ren gjennomsnittsberegning kan ha problemer med å fange opp slike samspill, noe som gjør det vanskelig å isolere effekten av en enkelt faktor, som andelen innbyggere i en bestemt brukergruppe, fra andre forhold som også påvirker kostnadene.

5.2.2 Normativ metodikk

En annen variant er normmodeller, der man istedenfor faktisk ressursbruk tar utgangspunkt i skjønnsbaserte eller erfaringsbaserte målsetninger om en referanseressursbruk som anses å være normal eller effektiv. En normativ modell bygger på fastsatte normer og standarder og beregner kostnaden ved å tilby samme tjenestenivå i alle fylkeskommuner. Dette forutsetter at det finnes normer eller standarder for tjenestetilbudet, slik at alle fylkeskommuner kan vurderes ut fra et felles referansegrunnlag. Der slike normer ikke foreligger i lov eller forskrift, kan det utarbeides egne standarder som kun benyttes til analyseformål.

I en normativ modell skal fylkeskommunene i prinsippet tilordnes midler etter de samme forutsetningene, uavhengig av hvordan de faktisk har organisert tjenestetilbudet. Dersom en fylkeskommune produserer dyrere enn det den fastsatte standarden tillater, må den dekke merkostnaden selv. Normative modeller kan på sitt beste bidra til å unngå belønning av ineffektivitet eller bortprioritering av en tjeneste, uten at man er avhengig av kvantitativ informasjon.

Kriterier basert på normative modeller brukes, på samme måte som statistiske kriterier, til å fordele en gitt ramme. Det er derfor de relative forskjellene mellom fylkeskommunene som er avgjørende, ikke det absolutte kostnadsnivået i modellen. Normative modeller er særlig relevante der det ikke finnes tilstrekkelige data for statistiske analyser, eller der statistiske metoder av andre grunner ikke kan benyttes. I dagens inntektssystem gjelder dette for kostnadsnøklene for fergeruter. Det ble brukt en normativ modell for båtruter fra 2022 til 2023 og en normativ modell for fylkesveier fram til 2021. Det er relativt få fylker som benytter ferger, samtidig som fergedrift involverer komplekse produksjonsforhold. Dette taler for bruk av normativ metode.

Utfordringen med normative modeller er at det ofte er vanskelig å utforme helt objektive kriterier. De normative modellene i dagens inntektssystem knyttes direkte til eksisterende tjenesteproduksjon. Kriteriet for ferger er for eksempel basert på forholdene ved de sambandene som faktisk er i drift. Dermed kan fylkeskommunene påvirke kriterieverdien og derved tildelingen av rammetilskudd ved å etablere nye samband eller endre eksisterende. Dette står i motsetning til prinsippet i inntektssystemet om at kriteriene i kostnadsnøklene skal være mest mulig objektive. Innslag av terskelverdier for å bli definert inn i en gruppe eller ikke kan også være problematisk for enheter som er i grenselandet for å oppfylle et kriterium.

Etter 2022 har det blitt brukt en modell for fylkesvei som egentlig er basert på regresjonsanalyser, men der den avhengige variabelen er de normerte kostnadene i 2019 i stedet for faktiske kostnader. De normerte kostnadene har sitt opphav i MOTIV-modellen, som brukes til å beregne normerte kostnader for ferger og senere også båtruter i inntektssystemet for fylkeskommunene. Agder-modellen er en beregningsmodell som tidligere ble brukt i utformingen av kostnadsnøkler i inntektssystemet, særlig knyttet til grunnskole og senere også andre kommunale tjenester. Modellen er basert på detaljerte beregninger av ressursbruk per elev, der man forsøker å fange opp hvordan ulike behovsdrivere påvirker kostnadene. Den er et eksempel på en slik modell som simulerer ulike skolestrukturer for å minimere reiseavstander eller kostnader under gitte geografiske og demografiske rammer. Samtidig har modellen vært gjenstand for betydelig faglig og forvaltningsmessig diskusjon, blant annet knyttet til valg av spesifikasjon, klassifisering av elever og sensitivitet for endringer i datagrunnlaget. Den har derfor blitt oppfattet som mer metodisk kompleks og mindre stabil enn mer aggregatbaserte modeller som MOTIV-tilnærmingen. Denne kompleksiteten har bidratt til at Agder-modellen i større grad har vært gjenstand for revisjoner og faglig kritikk, særlig når det gjelder robusthet og forutsigbarhet i kostnadsberegningene over tid (se for eksempel NOU 2005:18 og NOU 2022:10).

5.2.3 Hybridmodeller

Hybridmodeller innebærer at man kombinerer egenskapene til de øvrige modellene. For eksempel kan normative tillegg inkluderes for enkelte fylkeskommuner, dersom mål basert på statistiske analyser og ressurskartlegging ikke fullt ut fanger ut variasjonene i bruttodriftsutgiftene. Ulempen er at de er gjerne noe mindre transparente. Når det gjelder erfaringsbaserte estimater kan det også være en glidende overgang fra normative metoder til ressurskartlegging. For eksempel er fergestandarden som ligger til grunn for beregningen av fergekriteriet, basert på en kartlegging av faktisk antall avganger per fergesamband i ulike kategorier tilbake i 2018.

6 Systemevaluering

I dette kapittelet drøftes hvordan systemet kan evalueres, herunder utformingen av datasettet i delkapittel 6.1 og evaluering av modellvalget i delkapittel 6.2.

6.1 Utforming av datasettet

Statistisk metodikk er å foretrekke i inntektssystemet, da det gir mest objektive kriterier. Det er imidlertid ikke tilrådelig å bruke vekter som ikke er troverdige og robuste. Transparens er også ønskelig for å oppnå legitimitet. I slike tilfeller kan ressurskartlegginger, normative metoder eller hybridmodeller vurderes, jamfør delkapittel 5.2. I vår gjennomgang av dataimplementeringen vil vi like fullt fokusere på statistiske metoder.

6.1.1 Datasettets dimensjoner

Ved bruk av et lite datasett med få observasjoner er det begrenset hvor mange variabler man kan bruke. Dette vil spesielt være tilfelle ved regresjoner med tverrsnittsdata (uten flere tidsperioder), men også ved bruk av paneldata (med flere tidsperioder) eller alternativ metodikk vil det være grenser for hvor mye man får ut av tidsdimensjonen. Mange av de relevante kriteriekandidatene i inntektssystemet vil være nokså persistente over tid, slik at det er begrenset hva man klarer å få ut av tidsvariasjonen. Vi viser til seksjon 5.1.2 for en nærmere diskusjon av dette.

Regresjonsanalyse basert på paneldata kompliseres ytterligere av at inndelingen i fylkeskommuner endrer seg over tid. På kort sikt kompliseres estimeringen av paneldatasett ytterligere av at inndelingen i fylkeskommunene endres over tid. En nærliggende løsning er å bare ta for seg årene etter den siste revideringen av inntektssystemet, altså 2024 og senere. Ved en slik løsning mister man mange år, men det blir et stadig mindre problem ettersom årene går. Kanskje trenger man heller ikke så mange år i regresjonene. En annen løsning er å behandle fylkene før og etter sammenslåinger og splittinger som separate enheter, slik at man får et såkalt ubalansert datasett. Det medfører seleksjonsskjevhet ved at sammenslått fylke telles for én enhet og de oppsplittede fylkene telles som flere enheter, hvilket illustrerer at den administrative inndelingen betyr noe.

Det er også mulig å slå sammen separate fylkeskommuner som tidligere eller senere har vært én fylkeskommune. Eller man kan noe sjablongmessig eller med mer detaljerte analyser forsøke å splitte fylkeskommuner som tidligere og senere har vært flere fylkeskommuner. I det første tilfellet mister man en del observasjoner, mens det andre tilfellet vil være vanskelig å gjøre presist. Begge tilfellene vil dessuten lide av at det fundamentalt er forskjellig å drive én fylkeskommune og å drive flere fylkeskommuner, da både summen av antall lovpålagte og frivillige oppgaver vil avvike fra hverandre. En siste løsning er å ta ut fylkeskommunene som ikke er stabile over tid. I en slik løsning vil man miste såpass mange enheter, og det er antakelig uansett ikke tilrådelig annet enn i eventuelle robusthetssjekker.

Oslo kommune er et spesielt tilfelle, i og med at kommunen både er en primærkommune og en fylkeskommune. Dette er ikke noe problem for de sektorielle kriterieregresjonene, i og med at tjenestefunksjonene i Oslo enten er klart definerte med en tydelig kilde mellom primærkommunens tjenestefunksjoner og fylkeskommunens tjenestefunksjoner. For eventuelle systemkriterier

som bygger på aggregerte tall for kommunen, kan det imidlertid være hensiktsmessig å gjøre en slik oppsplitting basert på størrelsen på aktiviteten i hver del av kommunen (Vista Analyse 2024).

Dessuten er det mulig å foreta beregninger av mål og effekter på lavere aggregeringsnivåer enn fylker, for eksempel kommuner, tjenestevirksomheter eller individdata. Dermed kan man enklere ta hensyn til kjennetegn på et disaggregert nivå, samtidig som man fremdeles kan kontrollere for aggregerte kjennetegn. Slike beregningers relevans er ikke begrenset til statistiske metoder, men gjelder også ressurskartlegging, normativ metodikk og hybridmodeller.

6.1.2 Utvalget av variabler

Hvilke data som benyttes, avhenger av hvilken empirisk forklaringskraft variablene gir, og hvor robuste resultatene er. Indikatorene bør basere seg på offentlig statistikk og jevnlig oppdateres. Endringer i strukturelle forhold og tilgjengelige data over tid innebærer at man bør evaluere og eventuelt revidere hvilke kriterier som er med i inntektssystemet ved jevne mellomrom. I noen tilfeller kan mulige indikatorer komme ut som insignifikante eller med uforholdsmessig stor påvirkning på resultatene. Et viktig hensyn er at de underliggende dataene er målbare med en akseptabel presisjon og tilgjengelige, slik at beregningene kan etterprøves. For anvendelser knyttet til fylkeskommunenes inntektssystem vil mange av de nødvendige dataene være tilgjengelige eller kunne tilgjengeliggjøres. Som ellers i inntektssystemet, bør de også være utenfor kommunens kontroll for ikke å støte på kausalitetsutfordringer.

Forklaringsvariablene kan omfatte sosioøkonomi, geografi og næringsøkonomi. I tillegg kommer sektorielle kriteriekandidater knyttet til produksjonen. I appendiks A gir vi noen eksempler på utvalg av variabler som eventuelt kan utforskes i videre empiriske undersøkelser av potensielle kriterier i inntektssystemet. Mer utforskende utprøving av kriteriekandidater og kontrollvariabler i inntektsregresjonene kan bidra til å identifisere hvilke forhold i omgivelsene og øvrige rammebetingelser som bør hensyntas. Valg av variabler bør vurderes i lys av hvilke faktorer som har vist seg å ha stabil forklaringskraft i de innledende, utforskende kriterieregresjonene. I tillegg må man ta hensyn til at mange av kandidatvariablene som er meningsfulle på lavere aggregeringsnivå, ofte er tett korrelerte når de anvendes på fylkesnivå, noe som gjør det vanskelig å inkludere dem samtidig i modellen.

Som vi er inne på i seksjon 5.1.6, innebærer estimeringen i det kommunale inntektssystemet strengt tatt en spesifikasjonsfeil ved at man beregner omgivelsenes innvirkning på utgiftene uten å ta hensyn til nivået og kvaliteten på produksjonen av en gitt kommunal tjeneste. Produksjonsvariabler er innenfor kommunenes kontroll og egner seg dermed ikke som kriterier, men det kan likevel være ønskelig å ta hensyn til produksjonsteknologiens egenskaper når man beregner kriterienes innvirkning på utgiftene for en gitt tjeneste. Både for videregående skole og tannhelse er gode produksjonsmål tilgjengelige.

For videregående skole er antall elever gode volummål, mens karakterer og skolemiljøindikatorer er gode kvalitetsmål. For tannhelse er antall pasienter innenfor ulike alders- og prioriteringsgrupper med og uten tilsyn gode volummål, mens mulige kvalitetsmål inkluderer Significant Caries Index (SIC) og Decayed, Missing, Filled and Teeth Index (DMFT). I tillegg kan de nevnte kontekstuelle variablene – som fanger opp omgivelsene – inkluderes.

6.1.3 Bruk av komplekse variabler

I noen tilfeller kan finnes kriteriekandidater som fanger opp de samme fenomenene som dagens indikatorer, men som har et sterkere teoretisk og empirisk fundament. Det er ikke gitt at et skiftet fra dagens indikatorer til disse indikatorene vil tilføre praktisk verdi. Det avgjørende er hvor godt de aktuelle alternativene fungerer som kriterier i fylkeskommunenes utgiftsutjevning, med hensyn til forklaringskraft, stabilitet, kausalitet, målenøyaktighet og lav multikollinearitet, slik at man klarer å skille styrken på forskjellige mekanismer fra hverandre. I tillegg må man ta hensyn til praktiske aspekter, som at kriteriene er lett forståelige for brukerne og at målingene oppfattes som legitime.

På et så høyt aggregeringsnivå som fylker vil mange variabler som fanger opp ulike fenomener, være tett korrelerte. For eksempel er kvinneandelen, innvandringsandelen og inntektsnivået relativt høye i sentrale strøk. Derfor er det trolig ikke hensiktsmessig å inkludere alle variabler som har et teoretisk fundament og framviser høy forklaringskraft i empiriske målinger på et lavere aggregeringsnivå.

Maskinlæringsmetoder – som omtales i seksjon 5.1.4 – kan brukes eksplorativt i forbindelse med variabelseleksjon i de lineære kriterieregresjonene som ligger til grunn for inntektssystemet. De kan for eksempel bidra til å identifisere ikke-lineære sammenhenger, terskeleffekter eller relevante interaksjoner mellom kriterievariabler, som deretter kan brukes til å informere mer strukturerte regresjonsspesifikasjoner. Slike metoder kan også gi innsikt i hvilke variabler som har størst prediktiv betydning, og dermed hjelpe til med å velge hvilke kriterievariabler som bør inkluderes i de lineære modellene. Samtidig forutsetter bruk av maskinlæring forsiktighet, spesielt i små datasett, fordi både prediktiv ytelse og modellvalg kan være følsomme for enkeltobservasjoner, støy i dataene og valg av regulariseringsparametere. Dette innebærer at modellene lett kan overtilpasses små datasett dersom man ikke tar hensyn til kryssvalidering, parameterjustering og andre stabiliseringstiltak.

I tilfeller der det er mistanke om ikke-lineære sammenhenger, kan potensfunksjoner, polynomer, interaksjonseffekter og teoretisk funderte uttrykk være relevante, uten at man bør benytte seg av ren data mining (det vil si utforskende analyse av store datamengder for å finne mønstre uten a priori spesifisert modell). Komplekse sammenhenger vil være vanskelig å estimere på små datasett med bare fylkeskommunene. Det er imidlertid mulig å foreta estimeringen på forhånd på et rikere datasett og deretter bruke den estimerte variabelen inn i kriterieregresjonene. For mer sammensatte og komplekse variabler kan et alternativ være å undersøke betydningen på et lavere nivå, for eksempel for den enkelte videregående skole, tannhelseklinikk eller andre underliggende enheter, i stedet for på fylkesnivå. Det kan for eksempel dreie seg om sentralitetsmål som reflekterer avstandsulemper. Geografisk oppløsning og datatilgjengelighet spiller også inn på hvilket aggregeringsnivå analysene bør gjøres på (se Vista Analyse 2026 for detaljer).

6.2 Evaluering av modellvalget

Som vi har redegjort for i kapittel 5, fins det mange alternative metoder som man kan utlede kriteriene i fylkeskommunenes inntektssystem på. Dette vil henge sammen med de alternative metodenes treffsikkerhet og robusthet.

6.2.1 Overordnet om evaluering ved statistiske modeller

Utgangspunktet er at statistiske modeller gir de mest presise svarene på spørsmål om hvordan omgivelsene og produksjonsforhold virker inn på (de korrigerte) bruttodriftsutgiftene, som utgjør utfallsvariablene i kriteriefunksjonene. Hvor vidt man bør velge en statistisk modell, og hvilken det i så fall bør være, vil avhenge av de statistiske modellenes forklaringskraft og robusthet. Modellenes forklaringskraft gir en indikasjon på hvor godt modellen beskriver variasjonen i dataene, mens modellens robusthet handler om hvor sensitive parameterestimatene er for alternative spesifikasjoner, uteliggere eller endringer i datasettet.

Betydningen av målefeil kan undersøkes på flere måter. Enkle robusthetssjekker innebærer å analysere effekten av å introdusere målefeil i utvalgte forklaringsvariabler, samt å vurdere sensitiviteten for utelatelse av enkeltobservasjoner, observasjonsenheter eller år. Det kan også være at enkeltfylker får uforholdsmessig stor påvirkning på resultatene.

Videre kan en analysere hvordan resultatene påvirkes av alternative antakelser om datagrunnlaget, herunder ulike scenarier for sammensetningen av forklaringsvariabler eller vektingen av enkelte bestanddeler. Slike analyser gir innsikt i hvilke deler av modellen og datagrunnlaget som er mest utslagsgivende for estimatene, og bidrar til å vurdere resultatenes stabilitet. Robusthetssjekkene kan også utformes slik at en skiller mellom tilfeldige og systematiske målefeil, for eksempel ved å analysere konsekvensene av skjevheter som varierer systematisk mellom enheter eller over tid. Metodetriangulering – med uttesting av betydningen av valg av metode – vil også være av relevans.

Komplekse estimeringsmetoder kan ha problemer å konvergere på små datasett, hvilket representerer en vesentlig ulempe. Resultatene bør også være troverdige og robuste. Dette taler i utgangspunktet mot å bruke dem. Komplekse variabler med stor betydning for fylkeskommunenes tjenesteproduksjon kan eventuelt estimeres for fylkeskommunene med paneldata på forhånd eller med på et lavere aggregeringsnivå, som underliggende tjenesteenheter eller kommuner for tilgrensende tjenester. Vi viser til Vista Analyse (2026) for en mer detaljert diskusjon av denne problematikken i forbindelse med bruk av markedstilgangsmål som en måte å håndtere avstandsulemper på.

Det vil være relevante å vurdere teststatistikker for modellenes forklaringsevne som R^2 (R i annen) og $\check{R}^2$ (R hat i annen) for lineære modeller, og pseudo R^2 (pseudo R i annen), χ^2 (chi i annen), log-likelihood-baserte indikatorer og Root Mean Squared Error (RMSE) for ikke-lineære modeller. I evalueringen av modelltipasningen til ikke-lineære modeller, særlig i konteksten av både lineære og ikke-lineære modeller, gir Akaike Information Criterion (AIC) eller Bayesian Information Criterion (BIC) mål på modellkvaliteten med straff for kompleksitet.

I tillegg til forklaringsevne bør modellene vurderes med hensyn til prediksjonsrobusthet og overtilpasning. I praksis kan man gjøre dette ved bruk av kryssvalidering eller testdatasett, der man for eksempel utnytter data for andre årganger . Robusthet og sensitivitet kan undersøkes gjennom bootstrap-analyser eller sensitivitetsanalyser. For lineære modeller er det viktig å vurdere multikollinearitet og å teste for antakelser om normalitet og uavhengighet av residualene. I maskinlæringsmodeller bør forklarbarhet også vurderes, for eksempel ved å identifisere hvilke variabler som har størst innvirkning på prediksjonene. Endelig kan kryssmodellsammenlikning bidra til å finne modeller som balanserer forklaringskraft og kompleksitet optimalt.

6.2.2 Evaluering av statistiske modeller ved Monte Carlo-simuleringer

Monte Carlo-simuleringer gir et systematisk rammeverk for å etterprøve regresjonsresultater, som gjør det mulig å gå utover tradisjonelle diagnostiske tester og eksplisitt analysere hvordan usikkerhet, utvalgsstørrelse og modellantakelser påvirker inferensen. Simuleringene kan brukes til å teste og evaluere regresjonsmodeller ved å generere kunstige datasett under kontrollerte forutsetninger, slik at man kjenner den «sanne» datagenererende prosessen. Dette gir et verdifullt supplement til asymptotiske analyser og vanlige diagnostiske tester, særlig når modellen er kompleks eller utvalget er lite, som er tilfellet for kriterieregresjonene for fylkeskommunene. Fremgangsmåten innebærer typisk å definere en antatt datagenererende prosess, i dette tilfellet i form av en lineær modell med spesifikke koeffisientverdier, variansstruktur og mulige korrelasjoner i feilleddet. Deretter trekkes gjentatte tilfeldige utvalg fra denne prosessen, regresjonen estimeres for hvert utvalg, og fordelingen av koeffisientanslag, standardfeil og teststatistikker analyseres. Dette gjør det mulig å undersøke om estimatoren er tilnærmet forventningsrett, om variansene stemmer med hva man kan forvente av størrelsesorden og fortegn med basis i teorien, og om konfidensintervallene har korrekt dekningsgrad.

Monte Carlo-tilnærmingen er også nyttig for å teste modellens robusthet mot brudd på antakelser. Man kan simulere situasjoner med heteroskedastisitet, autokorrelasjon, ikke-normalitet i feilleddet eller multikollinearitet, og studere hvordan slike forhold påvirker koeffisientanslag, hypotesetester og modellvalg. Dette gir en konkret måte å vurdere hvor sensitive resultatene er for avvik fra standardantakelsene. Metoden kan videre brukes til å sammenlikne alternative estimeringsmetoder eller modellspesifikasjoner. Ved å holde den sanne datagenererende prosessen konstant, kan man vurdere forskjeller i skjevhet, effektivitet og teststyrke mellom ulike tilnærminger. I praksis kalibreres simuleringene ofte med observerte data, slik at resultatene speiler den faktiske situasjonen som analyseres.

Monte Carlo-simuleringer gir også et egnet rammeverk for å vurdere modellens prediktive egenskaper utenfor estimeringsutvalget, særlig når tilgjengelige data er begrensede. Ved å generere gjentatte datasett fra den samme datagenererende prosessen kan en splitte hvert simulert utvalg i et estimeringsutvalg og et prediksjonsutvalg, og deretter evaluere hvor godt modellen predikerer observasjoner som ikke er brukt i estimeringen. Dette gjør det mulig å analysere systematiske forskjeller mellom «in-sample»-tilpasning og «out-of-sample»-treff, samt å identifisere overtilpasning, som er en særlig utfordring i små utvalg. I denne sammenhengen kan Monte Carlo-simuleringer brukes til å sammenlikne prediksjonsfeil på tvers av alternative spesifikasjoner og estimeringsmetoder, og dermed gi et mer informativt beslutningsgrunnlag enn rene tilpasningsmål basert på estimeringsutvalget alene.

6.2.3 Evaluering av statistiske modeller ved maskinlæring

Maskinlæringsmetoder kan brukes som et rammeverk for å evaluere modellvalg, prediksjon og variabelseleksjon. Ved å anvende teknikker som kryssvalidering og bootstrap kan man systematisk teste hvordan ulike spesifikasjoner, inkludert ikke-lineære transformasjoner og interaksjoner, påvirker prediktiv ytelse og stabilitet i koeffisientanslagene. Maskinlæring gjør det mulig å vurdere hvilke kriterievariabler som har størst prediktiv betydning og dermed informere valg av variabler i de lineære kriterieregresjonene. Samtidig forutsetter slike metoder forsiktig bruk, særlig i små datasett, ettersom resultatene kan være følsomme for enkeltobservasjoner, støy og valg av regulariseringsparametere. Dette gir likevel et verdifullt supplement til Monte Carlo-simuleringer, ved

at man kan kombinere prediksjonsfokus og variabelseleksjon med innsikt i modellens robusthet og generaliserbarhet.

6.2.4 Evaluering av alternativ metodikk

Normative modeller, ressurskartlegginger og hybride modeller i inntektssystemet kan evalueres ved å vurdere hvor godt de reflekterer det faktiske ressursbehovet i fylkeskommunene. Dette innebærer å sammenlikne modellens anslag med observerte kostnader og oppgavebelastning for å undersøke samsvar mellom forutsetninger i modellene og empiriske data. Samtidig bør man teste robustheten i resultatene ved å variere forutsetninger og kriterier, slik at man ser hvor følsomt systemet er for endringer. Et annet viktig hensyn er at beregningene kan gjenskapes og at datagrunnlaget er tilgjengelig og etterprøvbart. Til slutt må man også vurdere den praktiske relevansen av modellene, det vil si om resultatene gir kriterier som er forståelige og legitime for brukerne av inntektssystemet. Det er også mulig å kjøre robusthetssjekker med kriterieregresjoner i de tilfellene det er mulig, men statistisk metodikk ikke brukes av inntektssystemet.

Referanser

Basuchoudhary, A., Bang, J. T., & Sen, T. (2017). Machine-Learning Techniques in Economics: New Tools for Predicting Economic Growth. Springer.

Borge, L.-E., Naz G., & Tovmo, P. (2003). Kostnads- og etterspørselsforhold i videregående opplæring. Allforsk.

Commonwealth Grants Commission (2021). Horizontal Fiscal Equalisation in the Australian Federation.

Commonwealth Grants Commission (2025). Commission's Assessment Methodology. 14. March 2025.

Dougherty, S., Montes Nebreda, A., & Urrutia, N. (2025). Intergovernmental Fiscal Transfers and Fiscal Equalisation in a Time of Consolidation. OECD Working Papers on Fiscal Federalism No. 50. OECD Publishing.

Eidgenössische Finanzverwaltung (2025). Cost Compensation 2026 – Excel Spreadsheet. Data – Fiscal Equalization. Published on 30. June 2025. Hentet fra https://www.efv.admin.ch/en/data-fiscal-equalization.

Federal Finance Administration (2025). National Fiscal Equalization. Swiss Confederation. 15. July 2025.

Finansdepartementet, & Justis- og beredskapsdepartementet (2024). Kort og spisset områdegjennomgang av økonomistyringen og budsjettsituasjonen i politiet. Arbeidsgrupperapport.

Greene, W. H. (2018). Econometric Analysis. Global Edition. Eight Edition. Pearson.

Kommunal- og distriktsdepartementet (2022). Forslag til nytt inntektssystem for fylkeskommunene fra 2024. Rapport fra ekspertutvalg.

Kommunal- og distriktsdepartementet (2023). Kommuneproposisjonen 2024. Prop. 112 S (2022 – 2023). Tilråding fra Kommunal- og distriktsdepartementet 11. mai 2023, godkjent i statsråd samme dag. (Regjeringen Støre)

Kommunal- og distriktsdepartementet (2024). Lov om kommuner og fylkeskommuner (kommuneloven). Dato: LOV-2018-06-22-83. Sist endret: LOV-2023-06-16-62 fra 01.05.2024.

Kommunal- og distriktsdepartementet (2025). Inntektssystemet for kommunar og fylkeskommunar 2026. Berekningsteknisk dokumentasjon til Prop. 1 S (2025–2026). Grønt hefte.

NOU (1996:1). Et enklere og mer rettferdig inntektssystem for kommuner og fylkeskommuner. Delutredning I fra Inntektssystemutvalget ved kongelig resolusjon av 24. februar 1995. Avgitt til Kommunal- og arbeidsdepartementet 11. januar 1996.

NOU (2005:18). Fordeling, forenkling, forbedring – Inntektssystemet for kommuner og fylkeskommuner. Utredning fra utvalg oppnevnt ved kongelig resolusjon 3. oktober 2003. Avgitt til Kommunal- og regionaldepartementet 10. oktober 2005.

NOU (2019:24). Inntektsfordeling mellom regionale helseforetak. Utredning fra et utvalg oppnevnt ved kongelig resolusjon 23. mars 2018. Avgitt til Helse- og omsorgsdepartementet 16. desember 2019.

NOU (2022:10). Inntektssystemet for kommunene. Utredning fra et utvalg oppnevnt ved kongelig resolusjon 12. mai 2020. Avgitt til Kommunal- og distriktsdepartementet 29. august 2022.

Rødseth, K. L, Førsund, F., Holmen, R. B. & Kittelsen, S. A. C. (2022a). Forbedringspotensial ved måling av effektivitet i kommunal sektor. TØI rapport 1879/2022.

Rødseth, K. L., Dong, B., Jakobsen, S., & Danebergs, J. (2022b). Revidert modell til beregning av fergekriteriet i inntektssystemet for fylkeskommunene. TØI rapport 1914/2022.

Rødseth, K. L., Kuosmanen, T., & Holmen, R. B. (2025). Mitigating Simultaneity Bias in Seaport Efficiency Measurement. Transportation Research Part A: Policy and Practice, 192, 104333.

SCB (2025). Kommunalekonomisk utjämning. Utjämningsåret 2026.

Simar, L., & Wilson, P. W. (2007). Estimation and Inference in Two-Stage, Semi-Parametric Models of Production Processes. Journal of Econometrics, 136(1), 31–64.

SOU (2024:50a). Nätt och jämnt. Likvärdighet och effektivitet i kommunsektorn. Del 1. Betänkande av Utjämningskommittén 2022. Stockholm 2024.

SOU (2024:50b). Nätt och jämnt. Likvärdighet och effektivitet i kommunsektorn. Del 2. Bilagor. Betänkande av Utjämningskommittén 2022. Stockholm 2024.

Stock, J. H., & Watson, M. W. (2020). Introduction to econometrics. Fourth Edition. Pearson.

Svendsen, H., F. Müller, H. Thune-Larsen, H., & Bråthen, S. (2017) Utgiftsbehov til ferjer og hurtigbåter. Møreforsking. Rapport nr. 1701.

Van Beveren, I. (2012). Total factor productivity estimation: A Practical Review. Journal of Economic Surveys, 26(1), 98-128.

Vista Analyse (2024). Bruk av kapitalkostnader i utgiftsutjevningen i inntektssystemet for kommunal sektor. Prinsipielle vurderinger og empiriske undersøkelser. Holmen, R. B., Eidsmo, E., Moum, K., Vennemo, H., Valseth, Å. S., & Ringdal. H. Rapport 2024/20.

Vista Analyse (2026). Avstandsulemper i utgiftsutjevningen for fylkeskommunene. Håndtering av geografi i fylkeskommunenes inntektssystem. Holmen, R. B., Ringdal, H., Svartsund, H., & Rasmussen, I. Rapport 2026/2.

Vista Analyse, Frischsenteret, OsloMet og Transportøkonomisk institutt (2026). Politiets tjenesteproduksjon. En undersøkelse av effektiviteten i politidistriktenes oppgavening. Holmen, R. B., Hoel-Holt, A. S., Kittelsen, S. A. C., Rødseth, K. L., Allvin, A., & Ellingsen, D. Rapport 2026/3.

Appendiks

A Mulige variabler i kriterieregresjonene

I dette appendikset gir vi eksempler på noen utvalg av variabler som eventuelt kan utforskes i videre empiriske undersøkelser av potensielle kriterier i inntektssystemet. Det dreier seg om forhold forbundet med sosioøkonomi, næringsøkonomi, geografi og produksjonsprosessen. I tillegg kommer sektorielle kriteriekandidater knyttet til produksjonen innenfor videregående skole og tannhelse. Appendikset utfyller med dette diskusjonen om dataimplementering i delkapittel 6.1.

Hvilke variabler som til slutt utvelges bør – som tidligere – ses i sammenheng med hvilke variabler som viser robustforklaringsstyrke i utforskende kriterieregresjoner, jamfør delkapittel 6.2. Det spiller også inn at mange av de kriteriekandidatene som er relevante på et lavere aggregeringsnivå blir vanskelig å utnytte samtidig på fylkesnivå på grunn av høy korrelasjon dem imellom.

A.1 Tverrsektorielle kontekstuelle variabler

I Tabell A.1 gir vi en oversikt over tverrsektorielle kontekstuelle variabler benyttet i Vista Analyse (2024) sin utredning om eventuelt innpassing av kapitalkostnader i det kommunale inntektssystemer, gjennomført på oppdrag for Kommunal- og distriktsdepartementet. Disse variablene har vesentlig innflytelse både kommuneøkonomi og regionaløkonomi.

Tabell A.1 Oversikt over tverrsektorielle kontekstuelle variabler i utredningen om kapitalkostnaders innpass i det kommunale inntektssystemet

Kontekstuelle faktorer
Befolkning, årsgjennomsnitt
Intern reiseavstandsproxy (summen av radien av en sirkel tilsvarende tettstedsarealet multiplisert med gjennomsnittlig befolkningsandel i tettsted og radien av en sirkel tilsvarende resten av kommunearealet multiplisert med gjennomsnittlig befolkningsandel for spredtbygde strøk, jamfør Holmen 2022)
Årlig befolkningsvekst, de siste ti år
Årlig sysselsettingsvekst etter bosted, de siste ti år
Sysselsetting per innbygger, årsgjennomsnitt
Gjennomsnittsalder for hele befolkningen
Gjennomsnittlig personinntekt for befolkningen over 17 år
Andel med høyere utdanning blant sysselsatte
Lokal arbeidsplassintensitet i form av sysselsetting etter arbeidssted over sysselsetting etter bosted
Herfindahl-indeks, næringskonsentrasjon i privat sektor (basert på to-siffer NACE utenom NACE 84 til 88, jamfør Kommunal- og moderniseringsdepartementet 2020)
Sysselsettingsandel i vareproduksjon (NACE 1 til 43)
Sysselsettingsandel i ikke-markedsrettede tjenester (NACE 84 til 99)

Kilde: Vista Analyse (2024)

I Tabell A.2 gir vi en oversikt over dataomfanget for tverrsektorielle omgivelser i forskningsstudien om effektivitet i de tolv politidistriktene i Norge, utført av Vista Analyse med flere (2026) på oppdrag for Justis- og beredskapsdepartementet.

Tabell A.2 Oversikt over tverrsektorielle kontekstuelle variabler i forskningsstudien om effektivitet i politisektoren

Kontekstuelle faktorer
Andelen av befolkningen over 16 år som er høyere utdannede, der høyt utdannede personer er personer som har fullført en universitets- og høgskoleutdanning (*Statistisk sentralbyrå*)
Andelen av befolkningen som bor i tettsted, justert for ukjent bosted (*Statistisk sentralbyrå*)
Andelen av befolkningen som er første- eller andregenerasjonsinnvandrere, utenom norskfødte med minst én norsk forelder og utenlandsfødte med minst én norsk forelder (*Statistisk sentralbyrå*)
Andelen av befolkningen som er kvinner i alle aldre (*Statistisk sentralbyrå*)
Andelen av befolkningen som er uføretrygdede, helt eller delvis (*Statistisk sentralbyrå*)
Andelen av befolkningen som har skilt eller separert som sivil status, hvilket ekskluderer personer som senere har inngått nytt formelt partnerskap (*Statistisk sentralbyrå*)
Andelen av befolkningen som tilhører ulike aldergrupper, 15 til 17 år, 18 til 20, 21 til 24 år, 25 til 29 år, 30 til 39 år, 40 til 49 år, 50 til 59 år, 60 til 69 år og over 70 år (*Statistisk sentralbyrå*)
Andelen av husholdningene som er trangbodde. Husholdninger regnes som trangbodd dersom antall rom i boligen er mindre enn antall personer eller én person bor på ett rom og antall kvadratmeterer under 25 kvm per person, eller det ene kriteriet er oppfylt med manglende data for det andre kriteriet. (*Statistisk sentralbyrå*)
Barneverntiltak per innbygger, der barnevernssaker inkluderer omsorgstiltak, hjelpetiltak, akuttiltak og atferdstiltak (*Statistisk sentralbyrå*)
Befolkning per landareal, der totalarealet tilsvarer totalarealet fratrukket ferskvannsarealet (*Statistisk sentralbyrå*)
Gjennomsnittlig inntekt i form av bruttoinntekt og personinntekt (*Statistisk sentralbyrå*)
Registrert arbeidsledighetsrate, målt som antall registrerte ledige over arbeidsstyrken (*NAV Arbeids- og velferdsforvaltningen*)
Sysselsettingsgrad, målt som sysselsetting på bosted over befolkningen (*Statistisk sentralbyrå*)
Ulikhetsmålene til EU og OECD, befolkningsandelen med mindre enn 50 og 60 prosent av medianinntekt, tilsvarende andelen personer med ekvivalert disponibel inntekt (bred forstand i EU-målet og snever forstand i OECD-målet) under 50 og 60 prosent av medianinntekten (*Statistisk sentralbyrå*)

Kilde: Vista Analyse med flere (2026)

A.2 Sektorielle kvalitetsindikatorer

I Tabell A.3 lister vi opp sektorielle kvalitetsindikatorer for fylkeskommunenes tjenester. Disse kan hensyntas i kriterieregresjonene, dersom man håndterer mulige simultanitetsutfordringer.

Tabell A.3 Sektorielle kvalitetsindikatorer for fylkeskommunenes tjenesteproduksjon

Tjenestefunksjon	Potensielle indikator	Kilde for grunndata
Administrasjon	Befolkningens aldersfordeling Tettheten av tjenestekontorer	Statistisk sentralbyrå
Samferdsel	Andeler bro og tunnel Gjennomsnittlig veibredde	Statens vegvesen
Tannhelse	Significant Caries Index (SIC) Decayed, Missing, Filled and Teeth Index (DMFT)	Statistisk sentralbyrå
Videregående skole	Skolebidrag Karakterer	Utdanningsdirektoratet

Kilde: Vista Analyse

Vista Analyse AS
Meltzers gate 4
0257 Oslo

rbh@vista-analyse.no
www.vista-analyse.no