



JUSTIS- OG POLITIDEPARTEMENTET

Veileder

Evaluering av lover

Med tilsvarende anvendelse på forskrifter og andre rettsregler



Innholdsfortegnelse

Forord	1
1 Innledning	2
1.1 Mål og målgruppe for veilederen	2
1.2 Leserveiledning	
2 Evaluering i forbindelse med regelverksutvikling og regelverksanvendelse	3
2.1 Hensikten med evaluering	3
2.2 Når og i hvilket omfang skal evalueringer gjennomføres	3
2.3 Evaluering sammenliknet med juridisk lovrevisjon og etterkontroll	4
2.4 Gjennomgangseksempel: evaluering av tvisteloven	5
2.5 Avgrensninger	5
3 Planlegging av evaluering	7
3.1 Formål, evalueringstema og berørte aktører	7
3.2 Evalueringsspørsmål	8
3.3 Databehov og indikatorer	8
4 Gjennomføring av evalueringen	10
4.1 Datainnsamling	10
4.1.1 Primær- og sekundærdata	10
4.1.2 Spørreundersøkelser	11
4.2 Analyse av innsamlet data	12
4.2.1 Bearbeiding av data og utledning av statistiske mål	12
4.2.2 Analyse av måloppnåelse og effekter	13
4.2.3 Analyse av produktivitet i forvaltningen	15
4.3 Fremstilling av resultatene	16
5 Oppfølging, evaluering og gevinstrealisering	17
5.1 Sammenheng mellom evaluering og gevinstrealisering	17
5.2 Innholdet i en gevinstrealiseringsplan	17
5.3 Oppfølging og anvendelse av evalueringen	18
6 Referanser og bakgrunnsdokumenter til eksemplene	19
Vedlegg A. Sjekkliste for planlegging, gjennomføring og oppfølging av en evaluering	20
Vedlegg B. Resultatkjeden	23
Vedlegg C. Eksempler på mulige evalueringsspørsmål	25
Vedlegg D. Nærmere om metodiske utfordringer ved gjennomføring av en evaluering	26
D.1 Kvantitativ eller kvalitativ evalueringsdesign	26
D.2 Innsamling av primærdata: utvelgelse av enheter og tidspunkt for målinger	26
D.3 Datas validitet og reliabilitet	27
D.4 Effektanalyse – metoder for å kontrollere for eksterne faktorer	27



Forord

Lov- og forskriftsreguleringer skal sikre visse samfunnsmessige og private verdier eller forhindre samfunnsmessige og private ulemper. Samtidig kan slike reguleringer i seg selv medføre forskjellige ulemper eller kostnader både i form av ressursbruk som går med til offentlige myndigheters håndhevelse og forvaltning, og i form av rapporteringskrav og andre plikter som pålegges private aktører. Lover og forskrifter legger også ofte begrensninger på folks frihet, samtidig som de bidrar til å støtte opp om og sikre denne friheten.

Alle offentlige tiltak og virksomheter, også lov- og forskriftsreguleringer og forvaltningen av disse, er underlagt evalueringsplikt. Evalueringer skal bidra til læring, styring og kontroll. Evaluering av lov- og forskriftsreguleringer innebærer å undersøke om reguleringen faktisk har effekter, og om den bidrar til måloppnåelse som forutsatt, både for det offentlige og for privat sektor. En god plan for evaluering av iverksatte lover og forskrifter bidrar både til å følge opp at prosesser blir gjennomført og til å etterprøve hvilke endringer de faktisk har medført. Økonomer vil ofte formulere dette som at evalueringer bidrar til å *realisere ønskede gevinster*, og det har i denne sammenheng derfor en annen betydning enn når man snakker om å realisere finansielle eller reale aktiva for å ta ut en økonomisk gevinst.

Veilederen gir et enkelt, men fullstendig rammeverk for hvordan man kan planlegge og foreta evalueringer i forbindelse med regelverksutvikling og regelverksanvendelse. Det

som står i denne veilederen, gjelder både lover og forskrifter, men for å forenkle fremstillingen er henvisning til forskrifter ofte utelatt. Veilederen er til støtte for både oppdragsgivere, planleggere og gjennomførere av evalueringer. Veilederen gir nyttig rettleiding uansett om man ønsker å sette ut selve gjennomføringen av evalueringen til eksterne miljøer eller om man vil foreta hele evalueringen selv.

H. Christopher Heyerdahl i analysestaben i Justisdepartementet har vært prosjektleder for arbeidet med veilederen, som er skrevet i samarbeid med Senter for statlig økonomistyring (SSØ). SSØ skal bidra til å heve kompetansen og gi råd til statlige virksomheter i forbindelse med hvordan samfunnsøkonomiske analyser og evalueringer bør gjennomføres, se www.sfsso.no for nærmere informasjon. Bidragsyttere fra SSØ har vært Espen Frøyland, Trude Gunnes og Gro Holst Volden. Vi har også fått gode innspill fra ekspedisjonssjef Knut Helge Reinskou i Justisdepartementets lovavdeling.

Veilederen legges ut på Internett og på hjemmesidene til både Justisdepartementet og SSØ, og vil derfor være tilgjengelig også for andre departementer og virksomheter som har ansvar for lovregulering og lovforvaltning. For å kunne evaluere nytten av veilederen og behovet for endringer vil Justisdepartementet gjerne ha tilbakemeldinger om erfaringer med bruken av veilederen. Tilbakemeldinger kan sendes på e-post til postmottak@jd.dep.no. Skriv "veileder i evalueringer" i emnefeltet.

1. januar 2009
Justisdepartementet



1 Innledning

1.1 Formål og målgruppe for veilederen

Formålet med denne veilederen er å vise hvordan man kan evaluere lov- og forskriftsreguleringer og anvendelsen av dem. Gjennom dette ønsker vi å bidra til at det blir gjennomført flere og bedre evalueringer både av lover innenfor Justisdepartementets eget ansvarsområde, og lover andre departementer har ansvar for. Veilederen gir en enkel og praktisk ramme for hvordan man kan planlegge, bestille, gjennomføre og følge opp evalueringer av lover og forskrifter. Finansdepartementet (2005a) har utgitt en overordnet veileder for gjennomføring av evalueringer. Dokumentet her er et supplement til Finansdepartementets veileder, og er en mer metodeorientert veileder som er spesielt tilpasset regelverksutvikling og regelverksanvendelse. For å lette forståelsen og illustrere fasene i evalueringsprosessen har vi brukt evaluering av tvisteloven som et gjennomgangseksempel.

Målgruppen for veilederen er ledere og saksbehandlere som forbereder og vurderer lovgivning i Justisdepartementet, og de som har forvaltningsansvar for lovverket. Ansatte som jobber med regelverksutvikling og forvaltning i andre departementer eller underliggende virksomheter vil også kunne ha nytte av veilederen.

1.2 Leserveiledning

Etter denne innledningen følger 4 kapitler som i stor grad følger fasene i en evaluering.

I kapittel 2 drøftes hva som menes med evaluering på lovområdet, og hvordan man kan vurdere behovet for å evaluere. Her introduseres også gjennomgangseksemplet.

I kapittel 3 nevnes de viktigste elementene i planleggingsfasen av en evaluering. Hvordan man stiller gode evalueringsspørsmål er sentralt her.

I kapittel 4 drøftes de viktigste faktorene i selve gjennomføringen av en evaluering. Viktige tema her er datainnsamling, analyse og fremstilling av resultatene. Vi konsentrerer oss særlig om evaluering av måloppnåelse/effekter av loven og om produktivitet.

I kapittel 5 gjennomgås oppfølgingen av en evaluering og sammenhengen mellom evalueringen og gevinstrealiseringsprosessen.

I vedleggene er det tatt inn en sjekkliste for fasene i en evaluering. I tillegg behandles tema som er mer faglig avanserte og særlig relevante ved store og omfattende evalueringer.



2 Evaluering i forbindelse med regelverksutvikling og regelverksanvendelse

2.1 Hensikten med evaluering

En av Justisdepartementets oppgaver er å ha det administrative ansvaret for regelverksutvikling på mange sentrale rettsområder. Med *regelverksutvikling* mener vi utforming av lover, forskrifter og andre rettsregler for å oppnå visse gevinster for samfunnet på kort eller lang sikt. Å regulere enkeltpersoners, bedrifters og andre aktørers atferd gjennom lover og andre reguleringer har alltid et formål. I noen tilfeller er dette eksplisitt, men med varierende presisjon, formulert som en egen formålsparagraf i loven. I andre tilfeller ligger det mer implisitt til grunn. Formålet er gjerne å fremme eller verne om viktige verdier, samfunns mål eller spesielle aktørers interesser. Det kan ligge ulike virkemidler i loven, og adferdsregulering og konflikthåndtering er to hovedkategorier. Innefor disse har man primærnormer som for eksempel kan gå ut på å gi personer rettigheter eller pålegge dem plikter, å påby eller forby bestemte handlinger eller en bestemt atferd, eller å gi ytelser. Videre har man sanksjonsnormer som kan gå ut på å ilegge administrative sanksjoner eller å straffe.

Man forventer altså visse gevinster for individer eller samfunnet i de situasjonene hvor det finnes lovregulering. Samtidig har lovregulering også ulemper og kostnader (ut over den tilsiktede ulempen for dem som rammes av sanksjoner dersom regelen brytes). Slike ulemper kan være kostnader i forbindelse med offentlige myndigheters håndhevelse og forvaltning, ressursbruk og bryderi i forbindelse med rapporteringskrav og andre plikter som stilles til private aktører, og ulempen ved å begrense folks frihet til å foreta handlinger slik de selv ønsker. Det er derfor viktig å spørre seg om innføring av en lovregulering faktisk fører til at samfunnsmålene nås og de tilsiktede målsetningene (gevinstene) realiseres. I tillegg ønsker man ofte å vite om loven er utformet slik at målsetningene kan oppnås (gitte gevinster kan tas ut) *samtidig* som ulempene ikke blir større enn nødvendig.

I forlengelsen av dette er det også viktig å undersøke om regelverksanvendelsen bidrar til å fremme målsetningene med rettsreglene. Med *regelverksanvendelse* mener vi både regelverksforvaltning og andre brukeres anvendelse av rettsreglene.

Evalueringer er et godt virkemiddel for å vurdere disse spørsmålene og dermed finne ut om lover og andre reguleringer virker etter hensikten. En evaluering av regelverksutvikling eller regelverksanvendelse er en systematisk datainnsamling, analyse og vurdering av et regelverk med tanke på utforming, målrealisering, effekter og/eller utvikling og anvendelse.

2.2 Når og i hvilket omfang skal evalueringer gjennomføres

Bruken av evalueringer er nedfelt i reglement for økonomistyring i staten § 16. Denne bestemmelsen gjelder all form for statlig virksomhet, også lover og reguleringer:

§ 16 Evalueringer

Alle virksomheter skal sørge for at det gjennomføres evalueringer for å få informasjon om effektivitet, måloppnåelse og resultater innenfor hele eller deler av virksomhetens ansvarsområde og aktiviteter. Evalueringene skal belyse hensiktsmessighet av eksempelvis eierskap, organisering og virkemidler, herunder tilskuddsordninger. Frekvens og omfang av evalueringene skal bestemmes ut fra virksomhetens egenart, risiko og vesentlighet.

Bestemmelsen er nært knyttet til ansvaret for å rapportere tilstrekkelig og relevant informasjon til Stortinget. Evalueringer skal supplere den løpende styrings- og kontrollinformasjonen, og de er på denne måten ment å bidra til effektive beslutningsprosesser, god risikostyring og generelt effektiv ressursbruk i staten. I noen tilfeller vil Stortinget eksplisitt be om en evaluering av et gitt lovverk. Når dette *ikke* skjer, er det virksomhetenes ansvar å vurdere behovet for en evaluering. Det følger av reglementet § 16 at både behovet for en evaluering og frekvens og omfang av denne vurderes ut fra faktorene egenart, risiko og vesentlighet. I tillegg må man vurdere hvor mye ressurser det er riktig å bruke på evalueringen. Nedenfor utdypes vi hva dette kan innebære ved evaluering av regelverk eller regelverksforvaltning.

- **Egenart.** Det kan være forhold ved selve loven, anvendelsen av den, brukerne eller området for øvrig som tilsier at det er behov for hyppigere eller mer omfattende evaluering på visse områder. Slike forhold kan være:
 - kompleksitet i loven
 - uklarheter og ufullstendigheter i loven
 - sårbar mål- eller brukergruppe (f.eks. barn) som ikke selv kan forventes å påpeke svakheter ved regelverket eller uten videre dra full nytte av regelverket
 - målkonflikter hos virksomheter som forvalter loven
- **Risiko.** Risiko er sannsynligheten for at det inntreffer hendelser som påvirker mulighetene for å oppnå de tilsiktede resultatene av lovreguleringen, eller som gjør at man oppnår utilsiktede effekter av lovgivningen og anvendelsen av den. Det kan i prinsippet være både positivt og negativt for resultatet. I staten er man ofte særskilt interessert i endringer som kan medføre *mindre gunstige* resultater. Noen aktuelle risikofaktorer kan være:
 - manglende erfaring (f.eks. ved en ny lov på et tidligere uregulert område)
 - usikkerhet når det gjelder om de valgte virkemidlene er egnet til å nå lovens formål (f.eks. valg av sanksjonsform, blant annet reaksjonsnivå)
 - usikkerhet om de tiltenkte målsetningene kan oppnås / gevinstene kan realiseres (f.eks. usikkerhet omkring hvordan målgruppen vil respondere)



- mulighet for misbruk av en ordning eller andre utilsiktede virkninger (f.eks. faren for at visse brukergrupper faller mellom to stoler)
- fare for manglende styring/kontroll med en ordning
- fare for at brukerne/målgruppen eller regelverksforvalteren har manglende kompetanse eller manglende kjennskap til regelverket

Vesentlighet. Med vesentlighet menes *hvor stor betydning* loven og dens effekter kan få. Eksempler på høy vesentlighet kan være følgende:

- det handler om særlig viktige samfunnsmessige verdier og/eller det er mange mennesker som berøres
- hver enkelt person berøres på en måte som i sterk grad påvirker vedkommendes velferd.
- loven og dens effekter er av stor politisk interesse
- ressursbruken i forvaltningen eller i den private sektoren er høy. Det er f.eks. et stort omfang av søknader/saker som behandles

Både departementene som har ansvaret for regelverksutviklingen av en lov eller forskrift, og virksomhetene som forvalter den, vil typisk ha både systematiske tilstandsmålinger og annen styringsinformasjon. I tillegg sitter de gjerne med mer usystematiske inntrykk av hvordan loven virker gjennom erfaring fra henvendelser fra brukere, medieoppslag og så videre. Slik informasjon kan være nyttig når man skal identifisere behovet for en evaluering.

En evaluering vil kreve ressurser. Hvor ressurskrevende det vil bli å finne svar på aktuelle evalueringsspørsmål, avhenger blant annet av eksisterende datagrunnlag og hvilke datainnsamlings- og analysemetoder som er aktuelle. En vurdering av behovet for evaluering, blant annet frekvens, omfang og metodevalg, bør alltid baseres på en nytte-kostnadvurdering. Ved store lovendringer og reformer bør man vurdere å inkludere planmessig registrering av data som er nødvendige for å evaluere loven.

2.3 Evaluering sammenliknet med juridisk lovrevisjon og etterkontroll

Det stilles visse krav til evalueringer som skiller dem fra andre typer utredninger, kartlegginger og vurderinger. Evalueringer skal

- vurdere verdien eller hensiktsmessigheten av noe i tillegg til å beskrive aktuelle forhold
- benytte systematikk og metode som er etterprøvbare
- utføres av en evaluator med en viss avstand til det som evalueres¹

¹ Dette innebærer imidlertid ikke at evalueringen må gjennomføres av et eksternt miljø. Evalueringer kan gjennomføres av både interne og eksterne aktører.

- ta utgangspunkt i et faktagrunnlag som definerer hva som skal måles, og hva som er ønskede og uønskede effekter av et tiltak

Tradisjonelt har man ofte gjennomført såkalte lovrevisjoner² og etterkontroller³ av lovverk. Disse tilfredsstillende gjerne kravene over, og de kan derfor betraktes som én type evaluering. Men noen ganger opplever man imidlertid at denne tradisjonelle fremgangsmåten ikke i tilstrekkelig grad retter søkelyset mot de helhetlige samfunnsmessige virkningene av loven og vurderingene av samlet måloppnåelse. I mange tilfelle har man vært lite opptatt av forskjellige ulemper og kostnader som er knyttet til en lovregulering.

I tråd med kravene i reglementet for økonomistyring i staten, vil denne veilederen i særlig grad dreie seg om *samfunnsøkonomiske* virkninger av lovverk. Det handler om å inkludere alle slags fordeler og ulemper ved en regulering og se dem i sammenheng med hverandre. Dette skal i utgangspunktet representere en bredere og mer generell form for evaluering enn de lovrevisjonene og etterkontrollene som Justisdepartementet tradisjonelt gjennomfører. På den annen side vil i praksis særlig lovrevisjoner inkludere juridiske perspektiv og hensyn som ikke har vært vanlig å ta med i en evaluering. Dette er det viktig å være oppmerksom på og inkludere ved gjennomføring av evalueringer av lover. Spørsmål om ivaretagelse av normer og rettigheter, uklare juridiske tolkninger, rettstomme rom og så videre vil måtte inngå som ett eller flere sentrale evalueringsspørsmål. Med denne veilederen håper vi å vise hvordan de tradisjonelle metodene kan og bør inngå som en viktig del av en større evaluering.

Ved evaluering av en lov vil man vurdere om eventuelle virkninger er i samsvar med det man la til grunn ved utformingen av loven, og om de eventuelle virkningene er i den størrelsesorden man i utgangspunktet ønsket. En regulering kan også ha fått utilsiktede virkninger, og det kan føre til at man ønsker å evaluere nytten og kostnadene av disse. I noen tilfeller forventer man kanskje at en lovendring skal gi en konkret tallmessig

² Den tradisjonelle fremgangsmåten ved utforming eller revisjon av lover er å nedsette et utvalg som består av juridiske eksperter på problemfeltet, og iblant supplere utvalget med eksperter fra andre fagområder. Utvalget beskriver rettstilstanden i Norge og trekker frem svakheter ved det eksisterende regelverket (situasjoner som eventuelt ikke fanges opp av loven, muligheter til å omgå regelverket, uklare tolkninger og eventuelt behovet for ny lovgivning i rettstomme rom mv.). Deretter utredes mulige rettslige løsninger. Her vil man normalt også se på hvordan tilsvarende regulering er utformet i en del andre land og spesielt andre land i Norden. Utvalgets arbeid munnner ut i en lovutredning (NOU) som sendes ut til offentlig høring, og deretter starter departementets arbeid med å utforme et forslag til ny lov eller endringer i gjeldende lov (Ot.prp.).

³ Når en lov har virket en stund, vil departementet ofte vurdere å gjennomføre en etterkontroll. På grunnlag av tilbakemeldinger fra regelverksanvendere og andres erfaringer med loven og dessuten etablert rettspraksis vil departementet ofte lage et høringsnotat med forslag til endringer og presiseringer. Notatet sendes ut til offentlig høring før departementet utformer en proposisjon.



(kvantitativ) virkning, som for eksempel en reduksjon i antall lovbrudd på et gitt område. Mange, og kanskje de fleste, nytte-virkningene vil imidlertid være vanskelige å måle eller tallfeste direkte. Eksempler kan være verdier som rettssikkerhet, rettferdighet, etiske vurderinger og ulike politiske føringer. Man bør likevel søke etter gode og fortrinnsvis målbare indikatorer⁴ på dem og for øvrig beskrive slike kvalitative virkninger på en systematisk måte. Man kan for eksempel få en indikasjon på hvordan folk opplever eventuelle endringer i rettssikkerheten som følge av en ny lov ved å gjennomføre en spørreundersøkelse. Det er også viktig å kartlegge kostnadene i tilknytning til loven. Det gjelder både kostnadene i form av ressurser som går med til å administrere ordningen, og kostnadene som faller på andre grupper i samfunnet.

2.4 Gjennomgangseksempel: evaluering av tvisteloven

For å lette forståelsen av stoffet benyttes et gjennomgangseksempel i veilederen. Dette eksemplet er basert på en evaluering av tvisteloven, se boks 2.1. Det vil bli brukt tilsvarende "fargede bokser" gjennom hele veilederen når vi kommer inn på dette eksemplet. Eksemplet er basert på SSØ (2007a).

Boks. 2.1. Introduksjon til gjennomgangseksempel

Ny tvistelov trådte i kraft 1.1.2008. Denne erstatter tvistemålsloven fra 1915. Ifølge Ot.prp. nr. 74 (2005–2006) (s. 1) er tvistelovreformen ment å gi en mer effektiv sivil rettspleie som gir raskere, billigere og riktigere tvisteløsning og bidrar til rettsavklaring. I forarbeidene til loven, Ot.prp. nr. 51 (2004–2005), er dette beskrevet mer utfyllende (s. 1–2):

"(...) Det er et mål både at tvister om mulig kan bli løst utenfor domstolene, og å lette tilgangen til domstolene i saker der det er behov for det. I saker som føres for domstolene må reglene legges til rette for en effektiv og retts sikker behandling.

Lovforslaget legger grunnlaget for en kulturendring blant dommere, advokater og parter når det gjelder behandlingen av sivile tvister. Det legger opp til en aktiv saksstyring fra dommerens side, med større vekt på saksforberedelsen og kortere hovedforhandling enn i dag. Håndteringen av den enkelte sak skal gjennomføres av et proporsjonalitetsprinsipp, slik at saksbehandlingens omfang står i forhold til tvistens betydning (...).

For å bedre tilgangen til domstolene og fremme rettsavklaring foreslår departementet regler om småkravsprosess og om gruppesøksmål (...).

Departementet ser det som et viktig mål å begrense kostnaden ved domstolsbehandling av reelle retts tvister. En dyr rettspleie kan

medføre rettsfornektelse. Regler som fremmer raskere saksbehandling og konsentrasjon om tvistens kjerne vil kunne gjøre tvisteløsningen billigere. Lovforslagets regler om avklaring av tvisten før søksmål blir reist, om mekling og rettsmekling vil bidra til at en tvist kan bli løst uten å bli belastet med utgifter til domstolsbehandling. Departementet foreslår forskjellige regler med sikte på å begrense nivået i sakskostnader i sivile saker."

Tvistelovreformen er altså en omfattende og viktig reform som berører alle landets domstoler (jf. kriteriet om vesentlighet). Det hersker også en usikkerhet om hvorvidt målene med reformen vil bli realisert (jf. kriteriet om usikkerhet). En årsak til det siste er at de tilskattede effektene ikke kan forventes å inntreffe ved lovforslaget alene, men avhenger av komplementære tiltak som opplæring og endring av arbeidskultur i domstolene og investeringer i IKT-utstyr. Det er også usikkert om etterspørselen etter domstolenes tjenester, og dermed samlet ressursbruk, vil øke, og i hvilken grad domstolene vil øke sin produktivitet.

Med bakgrunn i dette har Justisdepartementet bestemt at det skal legges til rette for evaluering av om sentrale mål med loven blir oppnådd. Det skal foretas både en etterevaluering etter tre år (dvs. i 2011) og en underveisevaluering mellom 2008 og 2011. Gjennomføringen av evalueringen bør skje i nært samarbeid med Domstoladministrasjonen som vil få et sentralt ansvar for uttak av flere av gevinstene.

2.5 Avgrensninger

Evaluering er i utgangspunktet et vidt begrep og kan omfatte mange ulike tilnærminger og metoder. En viktig avgrensning som gjøres i denne veilederen, er at vi primært konsentrerer oss om såkalte underveis- og etterevalueringer. Utredningsinstruksen stiller også krav til gjennomføring av føreevalueringer, som gjerne skal være i form av en samfunnsøkonomisk analyse. I den forbindelsen bør man vurdere ulike måter å utforme regelverket på, og ideelt sett bør lovregulering som virkemiddel for å nå et mål, vurderes opp mot økonomiske, pedagogiske og andre alternative virkemidler. Samfunnsøkonomisk analyse er et viktig, men omfattende tema som behandles i en egen veileder, se Finansdepartementet (2005b).

Videre er veilederen avgrenset til regelverksutvikling og regelverksanvendelse som evalueringsobjekt.⁵ Regelverksutvikling og forvaltning er i mange tilfeller delegert til underliggende virksomheter, og evalueringen vil derfor kunne fokusere på hvordan disse utfører arbeidet sitt. Veilederen går imidlertid ikke inn på etatsstyring som sådan.

⁵ Andre objekter som i henhold til Reglement for økonomistyring i staten bør være gjenstand for evalueringer, er blant annet eierskap, organisering og administrative, økonomiske og pedagogiske virkemidler, som informasjonskampanjer og liknende.

⁴ Indikatorer blir forklart i punkt 3.3.



Det overordnede målet med en lovregulering og de tilhørende gevinstene som ønskes oppnådd, vil ofte være politisk begrunnede. Denne veilederen vil ikke drøfte hensiktsmessigheten ved ulike mål og verdier.

Veilederen går heller ikke inn på det spesielle ved evaluering av tilskudds-, stønads- og garantiordninger, selv om slike ordninger ofte er lovregulerte. Det vises her til en egen veileder i evaluering av tilskuddsordninger, se SSØ (2007b).

I mange tilfeller vil departementet ønske å overlate gjennomføringen av evalueringen til en ekstern evaluator, som for eksempel et konsultentselskap. Denne veilederen tar ikke for seg selve bestillingsprosessen, forholdet til anskaffelsesregelverket, utforming av konkurransegrunnlag eller liknende spørsmål som dukker opp i den forbindelse.



3 Planlegging og evaluering

I dette kapitlet tar vi for oss de mest sentrale punktene for planleggingsfasen til en evaluering. Sjekklisten i vedlegg A gir en fullstendig beskrivelse av alle trinnene i planleggingsfasen, mens i dette kapitlet ser vi kun på de viktigste av disse trinnene. I større evalueringer kan den såkalte resultatkjeden være et nyttig verktøy i planleggingsfasen. Resultatkjeden gjennomgås i vedlegg B.

3.1 Formål, evalueringstema og berørte aktører

Med formål menes her hvorfor evalueringen skal gjennomføres. Det handler om hva den skal brukes til, det vil si hvilke beslutninger den ferdige evalueringen skal understøtte. Vi skiller gjerne mellom motivene *styring* (justering av forvaltningen), *læring* og *kontroll*. Formålet kan blant annet legge føringer på hvilken type informasjon som bør fremskaffes (hvor strenge beviskrav som stilles). Eksempler på formål med en evaluering kan være:

- Å vurdere om en bestemt regulering skal videreføres, endres eller erstattes av en annen
- Å justere forvaltningen av en lov fortløpende
- Å utforme fremtidige reguleringer og politikk, f.eks. på bakgrunn av endringer i lovverket i sammenliknbare land
- Å lære om et forvaltningsorgans eller domstolens praksis

Videre må den som planlegger, peke ut ett eller flere *evalueringstema*. Det handler om å velge de sentrale synsvinklene på evalueringobjektet. I denne veilederen skal vi særlig ta for oss to evalueringstema som vi antar vil være de mest aktuelle i forbindelse med lovregulering og lovforvaltning. Disse er:

- Måloppnåelse og realisering av ønskede målsettinger og virkninger. Her setter man søkelyset på utformingen av lovverket og hvordan reguleringene virker på brukere og samfunn.
- Produktivitet i forvaltningen av reguleringen. Her er man mest opptatt av saksprosesser og øvrig aktivitet internt i forvaltningsorganet.

Iblant har ønskede målsettinger og virkninger endret seg i tiden etter at loven ble vedtatt, og da er det viktig å klargjøre dette. Et vesentlig formål med evalueringen vil da kunne være å undersøke hvordan loven virker i forhold til de nye målsettingene, og å la dette gå inn som del av arbeidet med å vurdere behovet for å endre loven.

Planleggingen bør inneholde en kartlegging av hvilke aktører som er berørt av tiltakene. Spesielt gjelder dette ved evalueringer som undersøker i hvor stor grad man har oppnådd tiltenkte målsettinger, og hvilke effekter eller virkninger loven har hatt på individ og samfunn. Lovverk kan innebære både tilsiktede og utilsiktede effekter for mange ulike aktører. Aktørene kan være regelverksutviklere, regelverksforvaltere og andre brukere som privatpersoner, næringsliv og andre som direkte eller indirekte drar nytte av eller får ulemper i tilknytning til loven. I

praksis vil man bli nødt til å avgrense det antallet berørte grupper man skal konsentrere seg om i evalueringen. Men generelt bør alle effekter av et visst omfang og en viss betydning inkluderes. Man bør også ha en strategi for hvordan berørte aktører eventuelt skal involveres i evalueringsprosessen, og hvordan man skal håndtere dem som eventuelt ekskluderes.

I boks 3.1. nedenfor følger et eksempel på formål, evalueringstema og berørte aktører:

Boks 3.1. Formål, evalueringstema og berørte aktører i evaluering av tvisteloven

Formål

Formålet med evalueringen er å svare på om målene med tvistelovreformen nås. På grunnlag av dette kan man si at formålet er nært knyttet opp mot kontroll. Disse målene er om redusert behov for søksmål, lavere kostnader, raskere prosess, senket domstolsterskel, større gjennomslag for den materielle rett og en rettferdig og tillitskapende domstolsbehandling.

På noen områder er det satt opp konkrete måltall, og evalueringen må sørge for å samle inn data som kan gi svar på om disse nås:

- Partenes sakskostnader bør i snitt reduseres med 30 prosent, sett bort fra saker etter småkravprosessen.
- Hovedforhandling og ankeforhandling skal i snitt ta 30 prosent kortere tid enn i dag, sett bort fra saker etter småkravprosessen.
- Et dommerårsverk bør gjennomføre 15 prosent flere saker enn i dag, sett bort fra saker etter småkravprosessen.
- Et dommerårsverk bør gjennomføre 2,5 ganger flere saker etter småkravprosessen enn i dag etter tvistemålsloven.

Underveisevalueringen vil også gi viktige holdepunkter for å vurdere behovet for tiltak som kan rette opp skjevheter med ordningen. Det er meningen at evalueringen skal føre til læring.

Evalueringstema

Som det går frem ovenfor, vil evalueringen altså dreie seg om to hovedtema:

- Måloppnåelse og virkninger
- Produktivitet

Aktørene

Aktører som blir berørt av tvistelovreformen, og som det derfor kan være viktig å involvere i evalueringen:

- Domstolene og Domstoladministrasjonen
- Den enkelte dommer
- Advokatbransjen
- Partene i sivile saker
- Befolkningen for øvrig
- Justisdepartementet



3.2 Evalueringsspørsmål

Etter at formål og evalueringstema er avklart, må det formuleres et begrenset antall evalueringsspørsmål. Disse spørsmålene er ytterligere klargjøringer og operasjonaliseringer av hva som skal undersøkes og hvordan, og alle evalueringsspørsmålene skal til sammen dekke hele formålet. Det er ofte en fordel om den som skal gjennomføre evalueringen, kan delta i arbeidet med å formulere evalueringsspørsmålene.

Spørsmålene må være presise, entydige og mulige å besvare med tilgjengelige data. Ved utforming av evalueringsspørsmål kan det være hensiktsmessig å angi mulige årsaker eller hypoteser som evalueringen kan avkrefte eller bekrefte. Antallet spørsmål bør ikke være for stort, man bør unngå overlappinger, og man bør gjøre en nøye prioritering. Det kan gjerne gå frem av evalueringsspørsmålene hvilke variabler som skal undersøkes. Dersom det er viktig for bruken av resultatene, bør det også gå frem *hvilken form* som ønskes på resultatene (kvalitative eller kvantitative). Videre kan det være viktig å si noe om hvilken målestokk eller hvilke standarder resultatene vil bli holdt opp mot, og dessuten hvor strenge krav som stilles til datakvalitet. Evalueringsspørsmålene bør være **SMARTE**:

- **Spesifikke:** Det må være klart hva det spørres etter.
- **Målbare:** Spør etter størrelser som direkte eller indirekte kan måles.
- **Aksepterte:** Spørsmålene må oppleves som riktige, relevante og interessante.
- **Realistiske:** Spørsmålene må rettes mot mål som er mulige å innfri, og samtidig ambisiøse.
- **Tidfestede:** Det må være klart når målet som spørsmålet gjelder, skal være oppfylt.
- **Enkle:** Spørsmålene må være enkle å formidle og å forstå.

Ved å gjøre et grundig arbeid med utformingen av evalueringsspørsmålene, oppnås følgende:

- Man får en mer bevisst holdning til hva som er kjernen i problemstillingen, og dermed hva man skal spørre etter.
- Man tvinges til å ta stilling til om evalueringsspørsmålene faktisk lar seg besvare, og man unngår dermed problemer i gjennomføringsfasen.
- Man øker sannsynligheten for at evalueringen gir styrings- og kontrollrelevant informasjon siden evaluators rom for tolkninger av evalueringsspørsmålene blir mindre.
- Man blir bedre i stand til å rettlede arbeidet til evaluator.

I boks 3.2. gis noen eksempler på evalueringsspørsmål i tilknytning til evalueringen av tvisteloven. Se vedlegg C for en oversikt over aktuelle evalueringsspørsmål ved lovevalueringer.

Boks 3.2. Evalueringsspørsmål ved evaluering av tvisteloven

Som antydnet i boks 3.1. skal det gjøres en vurdering av både produktivitet og effektoppnåelse.

Evalueringsspørsmålene som er knyttet til produktivitet, er følgende:

1. Har dommernes arbeidsproduktivitet økt?
 - Man må måle antall saker per dommerårsverk etter hhv. allmennprosess og småkravprosess for å kunne svare på om konkrete måltall er oppnådd.
2. Tar saksprosessen kortere tid?
 - Man må måle tid fra hovedforhandlingen starter til ankeforhandlingen slutter, for å kunne svare på om konkrete måltall er oppnådd.
3. Er proporsjonalitetsprinsippet gjennomført i økt grad?
4. Hvilken rolle har IKT spilt for å øke produktiviteten i domstolsbehandlingen?
5. Hvordan har den totale ressursbruken i domstolene endret seg?

Evalueringsspørsmålene som er knyttet til effekter, er følgende:

1. Er partenes ventekostnader redusert?
2. Er sakskostnadene for partene redusert?
 - Man må måle gjennomsnittlige sakskostnader etter allmennprosessen for å kunne svare på om konkrete måltall er oppnådd.
3. Er behovet for søksmål redusert?
4. Er domstolsterskelen senket for de som har behov?
5. Er kvaliteten på saksbehandlingen minst like god som tidligere?
6. Slår den materielle retten mer gjennom?
7. Opplever domstolsbehandlingen som rettferdig og tillitskapende?

3.3 Databehov og indikatorer

Neste trinn vil være å kartlegge hvilken informasjon som er nødvendig for å kunne besvare evalueringsspørsmålene. Man må få oversikt over hvilke data som allerede finnes, og hvilke data man eventuelt vil bli nødt til å samle inn selv, og man må avgjøre hvilke metoder som skal benyttes ved innsamlingen. Det er hensiktsmessig at den som planlegger evalueringen, gjør en overordnet vurdering av dette. Det er imidlertid evaluators ansvar å detaljplanlegge og gjennomføre datainnsamlingen (se avsnitt 4.1).

Ved å vurdere databehov i planleggingsfasen oppnår man følgende:

- Eksisterende data blir gjort tilgjengelig for evaluator. Det kan for eksempel gjelde tilsvarende studier utført for annen type virksomhet, for andre virkemidler eller i andre land.

- Evaluatør foretar ikke unødvendig datainnsamling. Det er ofte planlegger (og oppdragsgiver) som kjenner best til hva som finnes av eksisterende data.
- Planlegger blir bedre i stand til å vurdere budsjettbehov for evalueringen og kompetansekrav til den som skal velges som evaluatør. Datainnsamlingen er ofte både ressurs- og tidkrevende, og noen fremgangsmåter (f.eks. spørreundersøkelser) kan kreve særskilt kompetanse og erfaring.
- Man kan legge til rette for fleksibilitet i struktureringen av datagrunnlaget som gjør det mulig å foreta eventuelle endringer om det oppstår behov for det.

Evalueringsspørsmålene kan ofte være vanskelige å formulere på en måte som kan besvares direkte. Et eksempel er spørsmål 1 i boks 3.2: "Har dommernes arbeidsproduktivitet økt?" Man må da finne frem til *indikatorer* som kan tjene til å belyse evalueringsspørsmålene, og finne data for disse. En indikator gir ikke nødvendigvis direkte svar på evalueringsspørsmålet, og det er viktig å finne indikatorer som belyser det sentrale i evalueringsspørsmålet. Er man uheldig med valg av indikator kan det medføre at evalueringen og oppfølgingen får feil fokus. Ofte bruker man flere indikatorer for hvert evalueringsspørsmål for å få et sikrere svar på spørsmålet.

I boks 3.3. gis et eksempel på overordnet vurdering av databehov. Se også boks 4.4 avsnitt 1 og 2.

Boks 3.3. Overordnet vurdering av databehov ved evaluering av tvisteloven

Det gjøres allerede en omfattende datafangst, og evalueringen bør i stor grad kunne gjøre bruk av denne. Domstoladministrasjonen, via saksbehandlingsregisteret LOVISA, samler løpende inn data om blant annet:

- Sakstilfang etter type prosess og type sak
- Tid mellom ulike milepæler i saksprosessen
- Størrelse på betalt rettsgebyr

For å svare på om dommerne spiller en mer aktiv rolle, og om proporsjonalitetsprinsippet gjennomføres, bør vi også ha mer nøyaktig informasjon om hva dommerne bruker tiden sin på. Domstolsadministrasjonen har tidligere gjennomført en tidsbruksstudie. Studien gjaldt omfattende tidsregistreringer i 27 domstoler. En mulighet kan være å følge opp med en tilsvarende studie etter at tvisteloven har trådt i kraft. Da kan vi benytte den tidligere studien som sammenlikningsgrunnlag.

De fleste av evalueringsspørsmålene som er knyttet til effekter, vil ikke kunne besvares godt uten innslag av subjektive vurderinger fra partene, deres advokater, dommerne og kanskje også forliksråd, utenrettslige instanser og forbrukerorganisasjoner. Det bør tas sikte på å gjennomføre spørreundersøkelser allerede høsten 2007 for å få et godt sammenlikningsgrunnlag for å måle effekter.

Det anses som mulig å fremskaffe dette datamaterialet innenfor prosjektets tids- og budsjettramme.



4 Gjennomføring av evalueringen

Dette kapitlet tar for seg den tredje og mest omfattende fasen i en evalueringsprosess, nemlig selve gjennomføringen. Denne fasen kan deles i tre deler: datainnsamlingen, dataanalysen og til slutt fremstillingen av resultatene. Sjekklisten i vedlegg A gir en fullstendig beskrivelse av alle trinnene som kan være relevante i gjennomføringsfasen. Her skal vi kun ta for oss de mest sentrale temaene.

Det er evaluator som har ansvaret for gjennomføringsfasen, selv om oppdragsgiver (planlegger) kan ha gitt mer eller mindre detaljerte føringer for hvilke data som skal samles inn, hvilke analysemetoder som skal benyttes, og hvordan man ønsker resultatene presentert. Oppdragsgiver for evalueringen har ansvar for å følge opp og kvalitetssikre arbeidet underveis, for eksempel i form av jevnlig møter, vurdering av foreløpige resultater og gjerne bruk av en referansegruppe. Vi går ikke inn på dette her.

Det går et hovedskille mellom såkalte kvantitative og kvalitative datainnsamlings- og analysedesign.⁶ For en utdyping av dette og enkelte andre tema som er knyttet til data- og analysemetoder, se vedlegg D.

4.1 Datainnsamling

Det er evalueringsspørsmålene (jf. avsnitt 3.2) som er utgangspunktet for å identifisere behov for indikatorer og databehov for øvrig, innsamlingsform og innsamlingsmetode. I noen tilfeller er evalueringsspørsmålene stilt relativt bredt og kan følgelig besvares på ulike måter. Et eksempel er dette spørsmålet: "Har produktiviteten i domstolene økt?" Produktivitet kan defineres på ulike måter (arbeidsproduktivitet, tidsproduktivitet mv.) Andre ganger kan planlegger ha etablert helt konkrete måleindikatorer som skal følges opp. Et spørsmål kan for eksempel lyde slik: "Har antall sivile saker per dommerårsverk økt?" Tilsvarende kan det være mer eller mindre gitt hvilke datakilder som skal benyttes, hvilke og hvor mange enheter som skal undersøkes, og med hvilken frekvens målingene skal foretas.

4.1.1 Primær- og sekundærdata

Noe av det første man må ta stilling til, er om man kan benytte ulike typer data som allerede er innsamlet for andre formål (sekundærdata), eller om man skal samle inn egne data for å svare på evalueringsspørsmålene (primærdata). Det første er vanligvis enklere og billigere, men spørsmålet er om dataene er relevante og treffsikre for evalueringen. Fordelen med

primærdata er at evaluator selv kan skreddersy datafangsten etter evalueringsspørsmålene. På den annen side er innsamling av primærdata ikke alltid mulig innenfor realistiske tids- og budsjettammer.

Uansett bør man alltid starte med å kartlegge hva som allerede eksisterer av sekundærdata. Det kan være samlet inn og bearbeidet data i forbindelse med andre liknende problemstillinger, og det kan være utarbeidet offentlige utredninger, forskningsrapporter og andre dokumenter om temaet. I mange tilfeller kan vi også få tilgang til data i ubearbeidet form, for eksempel fra ulike offentlige registre eller aktuelle virksomheters saksbehandlingssystemer.

Primærdata kan samles inn på ulike måter. De mest aktuelle teknikkene når det gjelder evaluering av regelverksutvikling og regelverksforvaltning, vil trolig være intervjuer eller spørreskjemaundersøkelser, jf. punkt 4.1.2. I enkelte tilfeller kan også observasjon eller fysiske målinger være aktuelt.⁷

Det er ofte hensiktsmessig å benytte kombinasjoner av primær- og sekundærdata. Det identifiseres gjerne visse "huller" i eksisterende datamateriale. Det gjør at det er behov for en egen innsamling. Primærdata vil dessuten være et komplement til eksisterende data, og denne typen data besvarer evalueringsspørsmålet på en annen måte.

Ved innsamling av primærdata må man gjøre et valg med hensyn til hvilke og hvor mange observasjoner man skal inkludere i undersøkelsen, og dessuten hvor ofte og på hvilke tidspunkter undersøkelsen skal gjennomføres. Videre må man, både ved primær- og sekundærdata, gjøre en vurdering av kvaliteten på datagrunnlaget som benyttes i undersøkelsen. Begge disse forholdene utdypes nærmere i vedlegg D.

I boks 4.1. illustrerer vi drøfting rundt datavalg for ett evalueringsspørsmål ved evalueringen av tvisteloven.

Boks 4.1. Datavalg for evalueringsspørsmålet "Er domstolsterskelen senket for dem som har behov?"

Ett av evalueringsspørsmålene som er knyttet til effekter, er dette: "Er domstolsterskelen senket for dem som har behov?" Det handler om at personer, også de som ikke har store økonomiske ressurser, bør få økt tilgang til å prøve saken sin, og at rettsforhold som det er uklarhet eller uenighet om, faktisk skal kunne prøves for retten.

Et annet evalueringsspørsmål er dette: "Er sakskostnadene for partene redusert?" Svaret på det spørsmålet vil indirekte

⁶ Med kvantitativt design menes at man samler inn mange observasjoner, og analysen foregår ved en statistisk bearbeiding av observasjonene med tanke på å kunne generalisere ut fra resultatene. Dersom man har en ferdig formulert hypotese om årsak og virkning på forhånd, velger man gjerne et slikt design. Med kvalitativt design menes bruk av innsamlingsteknikker som observasjon og åpne intervjuer med et knippe sentrale personer, og analysen foregår ved at man leter etter mønstre og forklaringer. Når det man søker er åpent og ukjent velger man gjerne en slik design.

⁷ Se for eksempel Ringdal (2000) for en grundigere gjennomgang av datainnsamlingsteknikker.

også si noe om domstolsterskelen og kan derfor også brukes som en indikator på om domstolsterskelen er redusert. Man tenker å måle dette blant annet ved å gjennomføre en bred spørreundersøkelse som er rettet mot et utvalg parter. Man vil i forbindelse med dette også stille spørsmål om størrelsen på betalt salær til advokat eller annen rettshjelper. Videre er de to nye tiltakene småkravprosess og gruppesøksmål ment som konkrete tiltak for å redusere barrierene mot å fremme sak. Nedenfor ser vi nærmere på hvordan effekten av disse kan måles.

I LOVISA vil man fra 1.1.2008 finne informasjon om utviklingen i antall saker etter henholdsvis småkravprosess og gruppesøksmål. Dette er fra evaluators ståsted å anse som sekundærdata. Omfanget av slike saker vil være en indikator på hvor viktige disse nye og billigere mulighetene til å reise sak er.

Det omfatter likevel ikke alle sider ved domstolsterskelen for dem som har et reelt behov for å prøve saken sin for retten. For å få flere indikatorer kan man derfor i tillegg vurdere å sende ut et spørreskjema til personer som har vært part i slike saker. Her vil man spørre om hva de ville gjort med saken i fravær av muligheten for henholdsvis småkravprosess og gruppesøksmål. Ett av svaralternativene må være "ville ikke reist sak". Utvalget av parter i undersøkelsen bør være tilstrekkelig stort og representativt til at man kan trekke generelle konklusjoner ut fra resultatene. Når det gjelder gruppesøksmål, forventes det å bli så få slike at det blir mulig å velge ut en eller flere enkeltparter fra alle sakene. Når det gjelder saker etter småkravprosess, vil man først foreta en tilfeldig trekning av 15 domstoler og deretter sørge for at ulike sakstyper er representert innenfor hver av disse. Det kan man gjøre ved først å sortere sakene etter sakstype og deretter trekke ut et tilfeldig utvalg fra hver sakstype hvor størrelsen på utvalgene avhenger av hvor mange saker det er av hver type (vektet utvalg).

Til slutt vil man i den planlagte spørreundersøkelsen som er rettet mot dommere og advokater, inkludere spørsmål om hvordan disse vurderer domstolsterskelen for privatpersoner som ikke får fri rettshjelp. Dette spørsmålet kan eventuelt også stilles til forbrukerorganisasjonene siden de kjenner saker som aldri når domstolene.

4.1.2 Spørreundersøkelser

Spørreundersøkelse er en mye brukt metode for å samle inn primærdata. Det er imidlertid enkelte vurderinger som bør gjøres rundt egnethet og utforming av spørreundersøkelser. For det første må man vurdere om målgruppen for spørreundersøkelsen faktisk har kompetanse til å svare på de spørsmålene man ønsker å stille. Når det gjelder lovreguleringer, bør man være særlig forsiktig med å gå ut til alminnelige borgere og ta for gitt at disse kjenner det aktuelle regelverket og effektene

av dette. Generelt vil brukere av en tjeneste kunne svare på hvor tilfredse de er med informasjon, saksbehandlingstid og andre serviceaspekter, men de vil ikke nødvendigvis være i stand til å vurdere det faglige innholdet i tjenesten. En spørreundersøkelse rettet mot advokater, dommere og andre som er tilknyttet rettssektoren, kan derimot inneholde mer faglig relaterte spørsmål.

Videre er det flere måter å utforme spørreundersøkelser på, fra helt standardiserte spørreskjemaer hvor respondenten krysser av for valgt svaralternativ, til mer åpne spørsmål hvor man tillater respondenten å utdype svarene. Standardiserte spørreskjemaer kan være en svært effektiv metode. Det er også enkelt og billig å analysere svarene i ettertid. Metoden er imidlertid kun egnet dersom man på forhånd har tilstrekkelig informasjon om problemstillingen og har utviklet hypoteser om årsak og virkning. Dersom man primært ønsker innspill om mulige årsaksforklaringer, er standardiserte spørreskjemaer i liten grad egnet i datainnsamlingen.

En spørreundersøkelse kan gjennomføres på ulike måter, blant annet ved personlig korrespondanse (intervju), på telefon, per post eller via internett. Ved spørsmål om sensitive forhold kan post og internett⁸ ha en fordel fremfor mer personlige innsamlingsmetoder. For øvrig avhenger valget av praktiske forhold, det vil si hvordan man enklest og billigst kan nå målgruppen og oppnå så høy svarprosent som mulig. Dersom man ikke henvender seg til personer som har en egeninteresse av å svare på undersøkelsen (f.eks. personer som berøres sterkt av en lovregulering), er det viktig at spørreskjemaet er så kort og lite tidkrevende som mulig, og at det er enkelt for respondenten å returnere det.

Spørsmålsformuleringen og oppsettet av svaralternativer er svært viktig. Det gjelder både med tanke på å vekke og opprettholde interessen (øke svarprosenten) og med tanke på å unngå å lede eller forvirre respondenten. Nedenfor følger noen enkle (ikke uttømmende) tips til utforming av spørreskjemaer:

- Åpne med lette og "ufarlige" spørsmål – eventuelle "provoserende" eller vanskelige tema introduseres til slutt.
- Hvis det er mange spørsmål, del inn i tema med egne overskrifter og kort innledning, og finn en naturlig kronologi på temaene.
- Bruk "trakt-teknikken" innenfor hvert tema – start på et generelt nivå og still oppfølgingsspørsmål kun til dem som det er aktuelt for basert på avgitte svar.
- Sikre at måleskalaen og svaralternativene er godt forklart og ikke kan misforstås.
- Ikke still ledende spørsmål.

⁸ Justisdepartementet har lisenser på dataprogram for gjennomføring av spørreundersøkelser via internett.



- Ikke still for generelle spørsmål – knytt dem til de helt konkrete tjenestene eller lovbestemmelsene som skal evalueres.
- Bruk enkle og klare begreper, ikke faguttrykk som kan være vanskelige å forstå for målgruppen.
- Svarkategoriene må være logiske og relevante, og de må dekke alle mulige svar.
- Ikke tillat kun ja/nei-svar (tilfreds / ikke tilfreds) ved spørsmål om holdninger, men bruk en mer differensiert skala, gjerne fem- eller sjudelt.
- Det er viktig å prøve spørsmålene på en gruppe piloter for å undersøke om de blir oppfattet slik det er tenkt.

I boks 4.2. gir vi et eksempel på deler av et spørreskjema i forbindelse med evalueringen av tvisteloven.

Boks 4.2. Spørreskjema rettet mot advokater. Evaluering av tvisteloven

Flere av evalueringsspørsmålene i evalueringen av tvisteloven kan besvares helt eller delvis gjennom et standardisert spørreskjema rettet mot et utvalg av advokater med tilknytning til de utvalgte tingrettene. Det vil være mest effektivt å samle disse spørsmålene i et felles spørreskjema. Skjemaet sendes advokatene per e-post.

Nedenfor gjengis deler av spørreskjemaet:

For alle spørsmålene er det fem svaralternativer i tillegg til kategorien "vet ikke". 1 betyr "i svært liten grad" og 5 betyr "i svært stor grad".

1. I hvilken grad gjennomføres proporsjonalitetsprinsippet, dvs. at man i retten bruker mest tid og ressurser på "viktige" saker?
2. Er det ditt inntrykk at det er de "mindre viktige" sakene som avvises eller heves?
3. I hvilken grad oppfatter du at dommeren styrer saken aktivt med sikte på å konsentrere behandlingen rundt sentrale punkter som det er strid om?
4. Blir sakene tilstrekkelig opplyst i løpet av prosessen, etter din oppfatning?

4.2 Analyse av innsamlet data

Når data er samlet inn, må materialet bearbejdes før vi kan bruke det til å besvare evalueringsspørsmålene. Analyse handler om å studere sammenhenger og mønstre for å kunne trekke konklusjoner om virkeligheten.

I dette avsnittet gjør vi først noen generelle betraktninger

rundt bearbejding av datamaterialet og utledning av de vanligste statistiske målene for enkeltvariabler. Deretter ser vi på analyse av de evalueringstemaene som vi har utpekt som mest sentrale i forbindelse med lovregulering og lovforvaltning:

- Analyse av måloppnåelse for loven og lovens virkninger på individ og samfunn
- Analyse av produktivitet i praktiseringen av loven

Vi presenterer kun en nokså overordnet tilnærming. I vedlegg D utdypes to mer avanserte analysemetoder for effektvurdering.

4.2.1 Bearbejding av data og utledning av statistiske mål

Man bør alltid starte analysefasen med å skaffe seg en oversikt over datamaterialet som er samlet inn. Vi er typisk interessert i å få vite hvor hovedtyngden av verdiene ligger, for eksempel hvor tilfreds flertallet av brukerne synes å være med saksbehandlingen. I tillegg er det interessant å observere spredningen i materialet. Det kan være av interesse å se om man har enkelte ekstreme verdier, og vurdere hva disse eventuelt kan skyldes. I noen tilfeller kan ekstreme observasjoner være av særlig interesse. Det kan for eksempel være at én domstol viser seg å behandle langt flere saker enn de øvrige som har like mye ressurser. I andre tilfeller kan man få mistanke om at det dreier seg om feilregistrering, misforståelser fra intervjuobjektets side eller liknende. I slike tilfeller bør man vurdere å renske denne typen observasjoner ut før dataene benyttes til analyse. Da unngår man skjevheter i evalueringresultatene.

Gruppering av observasjonene og frekvensfordelinger

Det er ofte hensiktsmessig å vise fordelingen av observasjonene, den såkalte frekvensfordelingen, grafisk. Til dette benytter vi gjerne stolpe- eller sektordiagram. Slik kan vi for eksempel enkelt vise hvor stor andel av et utvalg som er "Helt enig" i at kvaliteten på saksbehandlingen har vært god, og hvor stor andel som er "Helt uenig" i dette.

For å lage en frekvensfordeling vil vi ha behov for å gruppere observasjonene. Ved kvantitative design har man ofte etablert grupper eller kategorier på forhånd som dataene registreres på. Når man ser hvilke verdier som observeres, kan man eventuelt vurdere å slå sammen enkelte kategorier. Også ved kvalitative design kan dataene svært ofte grupperes og i noen tilfeller kvantifiseres i ettertid.

Statistiske mål på hovedtyngden av observasjonsverdier

Dersom de observerte verdiene er, eller lett kan omformuleres til, tallverdier, vil man også ønske å beregne et tallmessig mål på "midtverdi" i utvalget. Det mest brukte målet er gjennomsnittsverdien. Et annet alternativ er den såkalte medianverdien. Dette er den verdien som observeres hos enheten som befinner seg i midten når alle enhetene ordnes etter hverandre fra høyeste til laveste verdi.

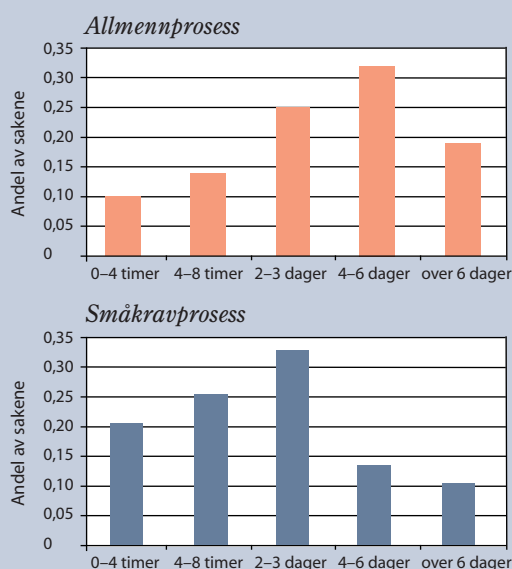
Statistiske mål på spredningen av observasjonene

Videre er det hensiktsmessig å få frem spredningen i utvalget. Et svært enkelt spredningsmål er variasjonsbredden. Dette er differansen mellom høyeste og laveste observerte verdi. Et mer avansert spredningsmål er det såkalte standardavviket.⁹ Standardavviket viser hvor mye en serie med verdier avviker fra seriens gjennomsnitt. Standardavviket varierer mellom null og uendelig, og det avhenger blant annet av variabelens målestokk. Det gir derfor størst mening i situasjoner hvor man har flere utvalg som sammenliknes med hverandre, eller hvor man har verdier for de samme enhetene på ulike tidspunkter.

I boks 4.3. nedenfor vises et konstruert eksempel på frekvensfordeling og dessuten mål på tyngdepunkt og spredning i fordelingen av tidsbruk for ulike tvistesaker.

Boks 4.3. Frekvensfordeling for tidsbruk under hovedforhandling

Fra 1.1.2008 vil mindre tvistesaker behandles etter den nye småkravsprosessen. Som del av ny tvistelov iverksettes videre en rekke tiltak som skal redusere tidsbruken i retten også i saker etter allmennprosess. Anta at vi har registrert tidsbruk under hovedforhandling i saker etter henholdsvis allmenn- og småkravprosess i et utvalg saker over ett år. Vi velger å dele dataene inn i følgende grupper: 0–4 timer, 4–8 timer, 2–3 dager, 4–6 dager og over 6 dager. En helt hypotetisk frekvensfordeling for året 2008 gjengis nedenfor som illustrasjon.



Fra figurene ser man at for denne tenkte frekvensfordelingen er den vanligste kategorien for saker etter småkravprosessen 2–3 dager, mens den vanligste kategorien for saker etter allmennprosessen er 4–6 dager.

⁹ Standardavviket er kvadratroten av variansen, som i sin tur er summen av kvadrerte avvik fra gjennomsnittet dividert på antall observasjoner i utvalget.

4.2.2 Analyse av måloppnåelse og effekter

Innføring eller endring av lovreguleringer, eventuelt endring av arbeidsform eller prosesser hos regelverksforvalter, har alltid et mål om å gi visse effekter. Med effekter mener vi primært virkninger for brukere og samfunn, det vil si aktører som vi ønsker å påvirke, men som er utenfor regelverksutviklers og regleverksforvalters direkte kontroll (se også bruker- og samfunnseffekter i vedlegg B om resultatkjeden). Ofte vil de overordnede målene med en lovregulering være identiske med de tiltenkte virkningene for brukere og samfunn, og da er effektanalyse og måloppfyllelsesanalyse to sider av samme sak. I sjeldne tilfeller kan det imidlertid være et mål *i seg selv* for eksempel å øke eller redusere en gitt aktivitet hos regelverksforvalter, uavhengig av hvilke effekter det gir. Vi ser bort fra dette her.

En analyse eller vurdering av lovens virkninger er sentral i svært mange evalueringer. Vi ønsker å vite om de tilskuede effektene av lovreguleringen faktisk ble oppnådd (realisert). I motsatt fall kan det bety at noen av virkemidlene er innrettet feil, at de har for liten styrke, eller at våre antakelser om hvordan brukere og samfunn ville endre atferd var feil.

Formålet med effektanalysen er å finne sammenhengen mellom lovendringen og observert tilstand på de aktuelle områdene. Tilstand kan for eksempel være tilfredshet, trygghetsfølelse eller helsetilstand hos en gitt målgruppe, eller tidsbruk, kostnader og andre forhold ved saksbehandlingen som antas å være viktig for brukere og samfunn. Tilstanden kan i noen tilfeller måles direkte (f.eks. kan tidsbruk måles i antall dager før man får svar på en søknad), men vanligvis må vi bruke mer eller mindre gode indikatorer (f.eks. kan trygghetsfølelse måles ved spørreundersøkelser eller salgsvolum for alarmsystemer).

En fullverdig effektanalyse må oppfylle følgende to krav:

1. Vi må påvise at en tilstandsendring har forekommet, og at dette har skjedd i tidsrommet etter at lovendringen har funnet sted.¹⁰ Det innebærer at det må foretas målinger både før og etter lovendringen. For å kunne avgjøre når ettermålingen bør gjøres, må vi blant annet ha kunnskap om hvor lang tid det antas å ta før virkningene oppstår.
2. Vi må kunne utelukke at tilstandsendringen har andre årsaker enn lovendringen. For å få til dette må vi kontrollere for alle eksterne faktorer som kan tenkes å ha en indirekte eller tilfeldig effekt på tilstanden.

Det siste kan være krevende fordi det ofte er mange andre faktorer som har endret seg i perioden, og det er vanskelig å avgjøre i hvilken grad dette har bidratt til tilstandsendrings-

¹⁰ Dersom lovendringen er varslet i god tid, bør man også åpne for muligheten for at aktørene har tilpasset seg, og at effekten kan være realisert allerede før lovendringen inntreffer. Det gjelder særlig muligheter for tilpassing i tidsrommet fra lovendringen blir vedtatt til den tretr i kraft.



gen. Det gjelder særlig dersom de andre faktorene påvirker tilstanden i samme retning som lovendringen er ment å gjøre. To vitenskapelig anerkjente måter å påvise en effekt på er eksperimentelle opplegg og regresjonsanalyse. Se vedlegg D for eksempler på bruk av de to metodene i tilknytning til evalueringen av tvisteloven.

Svært ofte er det imidlertid vanskelig eller umulig å isolere den effekten man er interessert i fra andre faktorer som også kan ha endret seg i samme periode. Det er likevel viktig å kartlegge hvilke andre eksterne faktorer som kan ha påvirket tilstanden, og skaffe informasjon om hvordan disse utvikler seg. Dersom man kan argumentere for at ingen andre faktorer kan bidra til å forklare en observert tilstandsendring, er dette en *indikasjon* på at det er lovendringen som har forårsaket endringene.

En annen enkel, men ikke veldig sikker metode er å foreta en spørreundersøkelse blant aktører som kjenner det aktuelle området godt. Aktørene blir spurt om hvorvidt de har registrert en endring i tilstanden etter lovendringen, og i hvilken grad de tror endringen skyldes lovendringen. Metoden er enkel fordi man ikke behøver å foreta førmålinger, og man slipper å bruke avanserte teknikker for å kontrollere for eksterne faktorer. På den annen side er aktørers subjektive vurderinger aldri å anse som "bevis", og man må være klar over at aktørene på området ofte vil ha en egeninteresse overfor resultatene. Som supplement til andre målinger kan dette imidlertid være nyttig for å belyse effektoppnåelse.

Til slutt må man alltid være oppmerksom på muligheten for at de aktivitetene og tiltakene som iverksettes, kan medføre andre og utilsiktede virkninger (positive eller negative) for ulike grupper – i tillegg til de tiltenkte gevinstene man ønsker å oppnå. Det er viktig å ha oversikt over utilsiktede virkninger. De vil nemlig virke inn på den samlede vurderingen av måloppnåelse og tiltakets samfunnsmessige verdi for øvrig.

I boks 4.4 nedenfor gis et eksempel på vurdering av en effekt i forbindelse med tvisteloven.

Boks 4.4. Effektmåling ved innføring av tvisteloven

Ett av evalueringsspørsmålene ved evaluering av tvisteloven er følgende: "Er domstolsterskelen senket for dem som har behov?" Tanken er at redusert tidsbruk og lavere kostnader ved å føre sak vil redusere terskelen.

Det er ikke mulig å finne en parameter som måler domstolsterskelen direkte, og vi må derfor ty til mer eller mindre treffsikre indikatorer. Som redegjort for i boks 4.1. kan man følge flere ulike parametere:

1. utviklingen i betalt salær, målt ved spørreundersøkelse mot partene

2. utviklingen i antall saker etter småkravprosess samt antall gruppesøksmål
3. spørreskjema til parter i ovennevnte saker for å avdekke om de ville ha reist sak dersom allmennprosess fortsatt var eneste mulighet
4. spørreundersøkelse mot dommere, advokater og eventuelt forbrukerorganisasjoner for å avdekke deres subjektive vurdering av domstolsterskelen

Når det gjelder punktene 1 og 4, er det viktig å starte målingene før den nye tvisteloven trer i kraft, for å kunne vurdere om eventuelle endringer kan skyldes ny lov, jf. krav nr. 1 til en effektmåling.

Videre må vi ta hensyn til andre faktorer som også kan ha påvirket verdien til de nevnte indikatorene i perioden, jf. punkt 2. Anta for eksempel at utviklingen i betalt salær er uendret etter ikrafttreddelsen av tvisteloven, til tross for aktiv saksstyring, småkravprosess og andre tiltak for å redusere tidsbruken i retten. To andre forhold som kan påvirke betalt salær, er sakenes kompleksitet og advokatenes timepriser. Det første kan for eksempel ha sammenheng med andelen av saker som har internasjonale forgreininger, mens det andre avhenger blant annet av konkurransetrykket i advokatbransjen. Det er viktig å vurdere hvordan disse forholdene har utviklet seg i perioden.

Tilsvarende kan man tenke seg at lav utbredelse av småkravsaker og gruppesøksmål, mange som svarer at de ville reist sak uansett, og en økende andel av dommere og advokater som vurderer domstolsterskelen til fortsatt å være "høy", kan forklares ved endringer i andre faktorer enn ny tvistelov. Dersom man uten å ha tatt hensyn til eksterne faktorer konkluderer med at tvisteloven ikke har hatt effekt, kan det være en feilvurdering. Sannheten kan være at tvisteloven isolert sett har hatt stor effekt, men at manglende utslag på indikatorene skyldes andre faktorer som trekker i motsatt retning.

Hvis effektmålingen viser at de tilsiktede virkningene uteblir (også når vi tar hensyn til eksterne faktorer), eller at negative utilsiktede effekter er betydelige og overskygger de positive effektene, må vi gå videre og spørre hva som er årsakene til dette, og hvordan lovreguleringen eventuelt kan endres.

Mulige årsaker til manglende effekter kan være:

- At andre nødvendige støttetiltak mangler, som opplæring eller tilføring av ressurser
- At forvaltningen av regelverket er mangelfull, f.eks. at søknader fra den tiltenkte målgruppen blir avslått, ikke blir behandlet til rett tid, følges opp i for liten grad e.l.
- At man ikke har innsett hvor viktig det er å innarbeide nødvendige incentiver hos de partene som man har et mål om å endre atferden for

- At brukere, advokater eller andre aktører har manglende kompetanse, manglende kjennskap til eller manglende informasjon om reguleringen
- At styrken på reguleringen og/eller oppfølgingen av den ikke er korrekt tilpasset, f.eks. kan svak straff og/eller lav risiko for å bli avslørt ved brudd på reguleringen medføre at flere velger ikke å følge loven.

4.2.3 Analyse av produktivitet i forvaltningen

I tillegg til å analysere effekter for brukere og samfunn vil man i en evaluering ofte være opptatt av å se på produktiviteten i *forvaltningen* av regelverket. Mange regelverksendringer kan være begrunnet nettopp med et ønske om kortere saksbehandlingstider og/eller redusert ressursbruk per sak.

Produktivitet kan defineres som forholdet mellom produksjonen og medgått ressursbruk i forvaltningen. Produksjon måles ved antall behandlede saker eller søknader eller eventuelt ved andre tjenester eller aktiviteter som er aktuelle. Ressursbruk kan enten måles samlet i kroner, eller man kan ta for seg de mest sentrale innsatsfaktorene, for eksempel arbeidstid. Avhengig av hva man er mest interessert i å studere, kan ulike produktivetsmål defineres. Noen av de vanligste målene er:

- Arbeidsproduktivitet: antall saker per års-/månedsværk
- Kostnadseffektivitet: ressursbruk i kroner per sak

For at slike målinger skal være meningsfulle, er det helt avgjørende å ha en konkret definisjon av innholdet av en aktivitet eller egenskapene til et produkt, for eksempel hva som utgjør en "sak". Det er vanligvis hensiktsmessig å dele inn i undergrupper dersom det er stor variasjon mellom sakene. Skal man for eksempel evaluere produktiviteten ved domstolene, kan det være misvisende bare å se på samlet antall domsavsigelser. Endringen i ressursbruk per dom kan for eksempel skyldes en endring i sammensetningen av sakene. Derfor kan det være hensiktsmessig å skille mellom ulike sakstyper og ulike prosessformer.

Produktiviteten vil direkte eller indirekte bli vurdert opp mot en norm for best mulig produktivitet for et gitt kvalitetsnivå. Man bør søke å finne frem til denne produksjonsnormen. For eksempel kan produktivetsmålingene brukes som utgangspunkt for å identifisere underenheter eller produksjonsområder som er særlig produktive, og forsøke å finne årsakene til variasjonen. Slik kan alle lære av de beste. Mer generelt gir produktivetsmålinger grunnlag for å vurdere hvordan prosesser kan forenkles, hvordan flaskehalsen kan fjernes, og så videre. Målet er å ta ut potensialet for produktivetsforbedringer.

En norm for best mulig produktivitet må nødvendigvis ta for seg de begrensningene en forvaltningsvirksomhet driver innenfor. Slike begrensninger kan være juridiske så vel som geografiske og organisatoriske. Ved å måle produktiviteten i forbindelse med endringer i nettopp regelverk, geografisk

struktur, organisasjon eller annet kan vi svare på om selve potensialet for produktiviteten har økt.

I boks 4.5. gis det et eksempel på vurdering av produktivitet i tilknytning til tvisteloven.

Boks 4.5. Produktivetsmåling ved innføring av tvisteloven

Ved evaluering av tvisteloven anvendes to komplementære mål på produktivitet:

1. tidsbruk per sak
2. antall sivilsaker per dommerårsverk

Vi forventer at de to målene samvarierer i stor grad. Det første målet bør reduseres og det andre øke når ny tvistelov trer i kraft. Sentrale tiltak i loven som skal bidra til økt produktivitet, er blant annet innføring av småkravprosess, regler om aktiv saksstyring og regler om plikter før sak reises (for å klargjøre saken bedre før stevning).

For å måle tidsbruk per sak benytter vi Domstoladministrasjonens statistikk over tid mellom ulike milepæler: (i) tidsintervallet fra stevning til endt saksforberedelse, (ii) tidsintervallet fra start på hovedforhandling til avsluttet forhandling og (iii) tiden fra avsluttet hovedforhandling til dom foreligger. I noen tilfeller vil det også bli nødvendig å måle tidsperioden fra ny saksforberedelse etter anke til ankeforhandling og dom.

Antall sivilsaker per dommerårsverk kan måles ved bruk av informasjon om antall sivile saker som er behandlet, og antall dommerårsverk.

Vi bør skille mellom ulike prosessformer (allmennprosess og småkravprosess) for begge målene. På denne måten kan vi følge utviklingen i produktivitet innenfor hver prosessform i tillegg til gjennomsnittet. Tilsvarende kan det være hensiktsmessig å skille mellom ulike sakstyper som krever henholdsvis store og små ressurser. I tillegg bør målingene skje på domstolsnivå, ikke kun som et overordnet gjennomsnitt, slik at produktive domstoler kan identifiseres og deres praksis kan kartlegges.

Til slutt nevner vi to relevante problemstillinger knyttet til måling av produktivitet:

- Det er ofte slik at flere faktorer spiller inn når det gjelder å forbedre produktiviteten. For eksempel er flere effektivitetsfremmende virkemidler iverksatt som del av den nye tvisteloven. Det kan da være vanskelig å isolere de ulike faktorene.
- Når vi snakker om produktivitet, legger vi hele tiden implisitt til grunn at kvaliteten holdes uendret. Tiltak for å effektivisere produksjonen kan imidlertid i en del tilfeller medføre incentiver som er med på å redusere kvaliteten. Derfor er det viktig at vi også måler kvaliteten og følger utviklingen til denne.



4.3 Fremstilling av resultatene

I dette avsnittet gir vi noen retningslinjer for fremstilling av resultatene av evalueringen. De fleste evalueringer vil innebære at man skriver en evalueringsrapport. Resultatene kan også fremstilles som lysbildeserie, muntlig presentasjon eller liknende. Oppdragsgiver bør ha spesifisert dette i oppdragsbestillingen dersom det er bestemte ønsker knyttet til fremstillingsform eller konkret utforming av evalueringsrapporten. Evaluatør må på sin side sørge for at rapporten gjenspeiler innholdet i bestillingen, blant annet at alle evalueringsspørsmål er besvart.

Generelt er det viktig at rapporten har en enkel og logisk struktur, at de viktigste resultatene kommer klart frem, og at resultatene er pedagogisk fremstilt. Med andre ord skal rapporten være ryddig redigert og ha klare konklusjoner. Det bør redegjøres for metodevalg og fremgangsmåte. Rapporten skal videre gi leseren tilstrekkelig bakgrunn til å vurdere om konklusjonene er holdbare. Alle datakilder må være dokumenterte. Det bør også redegjøres for hvordan de ulike resultatene skal tolkes,

og med hvilken sikkerhet man eventuelt trekker konklusjoner og generaliserer resultatene. I en rapport er det essensielt også å påpeke eventuelle uenigheter og begrensninger.

En mulig struktur for en evalueringsrapport kan være slik:

- Sammendrag
- Introduksjon
 - bakgrunn og formål med evalueringen
 - oppdragsgiver
 - viktige rammer eller føringer
 - kort om evalueringsprosessen, herunder bruk av referansegruppe eller andre til kvalitetssikring
- Evalueringstema og evalueringsspørsmål
- Metode og datagrunnlag
 - valg av analysemetode
 - valg av datakilder og omtale og dokumentasjon av disse
- Hovedfunn
- Drøfting og vurdering av resultatene, påpeking av begrensninger
- Konklusjoner og eventuelle anbefalinger



5 Oppfølging, evaluering og gevinstrealisering

Gevinstrealisering kan ha flere betydninger, og i denne sammenheng brukes det som nevnt i innledningen til veilederen, om oppfølging av måloppnåelse. En gevinstrealiseringsplan er et verktøy for å følge opp gjennomføringen av tiltaket og at de planlagte målsetninger nås (realiseres). Både gevinstrealisering og evaluering skjer som hovedregel etter at prosjektet er igangsatt. I dette kapitlet tar vi for oss sammenhengen mellom evaluering og gevinstrealisering og peker på sentrale momenter i forbindelse med oppdragsgivers oppfølging av evalueringen i etterkant, jf. sjekklisten i vedlegg A.

5.1 Sammenheng mellom evaluering og gevinstrealisering

Innføring eller endring av lover er ofte begrunnet med at dette skal føre til at man når fastsatte målsettinger, og at man oppnår nyttevirksomheter for individ eller samfunnet. I fortsettelsen blir dette kalt gevinster. Ideelt sett bør det være gjennomført en samfunnsøkonomisk analyse i forkant av regelverksendringene som identifiserer og konkretiserer gevinstene og kostnader, ulemper mv.¹¹ En samfunnsøkonomisk analyse skal i prinsippet inkludere alle virkninger av tiltaket. Det er imidlertid en utfordring å identifisere alle virkninger, og mange juridiske virkninger vil kreve at jurister er med på å utforme analysen. Med utgangspunkt i en slik analyse vil det være enklere å etablere en gevinstrealiseringsplan. En gevinstrealiseringsplan er en plan for oppfølgingen av de endringene som er gjort i regelverk eller forvaltningspraksis for å nå forventede gevinster (målsetninger), og den kan være svært viktig når det gjelder hvor store gevinster man vil klare å oppnå og dermed påvise i en evaluering.

I en evaluering vil man typisk se på om man har klart å realisere (oppnå) de forventede gevinstene (målsetningene) av lovreguleringen, og man vil vurdere årsaker til eventuelle avvik (f.eks. i form av at resultatene blir bedre eller dårligere enn forventet).

Gitt den nære sammenhengen mellom en gevinstrealiseringsplan og en evaluering er det en fordel om man kan legge planer for disse samtidig. Slik kan man oppnå synergieffekter i planleggingsfasen og i gjennomføringsfasen av både gevinstrealisering og evaluering.

5.2 Innholdet i en gevinstrealiseringsplan

En gevinstrealiseringsplan skal være en handlingsplan til bruk i oppfølgingen av regelverksendringene, og den bør inneholde følgende momenter:

- Hvilke gevinster som skal oppnås (for eksempel bedre kvalitet i saksbehandlingen)
- Hvor gevinstene vil oppstå (f.eks. i hvilken del av saksbehandlingsprosessen/lovens virketid, hos hvilken målgruppe o.l.)

- Når gevinstene forventes realisert (på kort eller lang sikt)
- Hvilke (resultat)indikatorer som er valgt for å si noe om i hvilken grad gevinstene er realisert
- Datakilde for resultatindikatorerne måling/evaluering av den aktuelle gevinsten, medregnet beskrivelse av hvor ofte måling/evaluering bør gjennomføres, og hvem som skal foreta dette – ofte er det også nødvendig å beskrive sammenlikningsgrunnlaget for målingen
- Hvem som er ansvarlig for at gevinstene realiseres forutsetninger og nødvendig tilrettelegging for å oppnå gevinstene, f.eks.:
 - endret organisering
 - opplæring av ansatte eller andre
 - innføring av økonomiske eller ikke-økonomiske incentiver, som belønning av gode resultater
 - informasjon om regelverksendringene til brukere, advokater mv.
- Sentrale risikofaktorer som kan påvirke mulighetene for å oppnå gevinstene med loven (gevinstuttak), som kan tilsi behov for eventuelle risikoreduserende tiltak (for eksempel mer opplæring)

Grovt sett kan vi si at alle de fem første punktene er sammenfallende i en evaluering og i en gevinstrealiseringsplan. Potensielle gevinster realiseres ofte ikke av seg selv straks loven trer i kraft. Man må derfor også utpeke konkrete personer eller enheter som er ansvarlige for at de ulike gevinstene (målsetningene) realiseres. De må sørge for at aktuelle tilretteleggingstiltak og risikoreduserende tiltak iverksettes i tide, og at det blir rapportert på resultatindikatorene innen gitte frister.

For å sikre en god planlegging og oppfølging av en gevinstrealiseringsplan er det generelt viktig at alle berørte aktører involveres tidlig. Da vil man øke forståelsen for og eierskapet til de tiltakene som blir gjennomført. Særlig viktig er det at ledelse og ansatte i forvaltningsorganet involveres når det er snakk om endringer i måten disse forvalter regelverket på, når målet er å oppnå produktivitetsgevinster og liknende. Et bidrag til dette kan være at linjeledelsen er med på å utrede lovendringen fra starten av ved for eksempel:

- Å bli rådspurt i prosessen med å beskrive gevinstpotensialet og de forutsetningene som ligger til grunn for dette – gjennom å være med på å sikre realistiske tall og vurderinger får ledelsen tiltro til at de beregnede gevinstene kan og bør realiseres
- Å delta i en styringsgruppe for utredning av lovendringen
- Å delta i arbeidet med gevinstrealiseringsplanen og gevinstrealiseringsprosessen

I boks 5.1. viser vi et eksempel på utdrag av en gevinstrealiseringsplan for tvisteloven.

¹¹ Se også SSØ (2006a) for en ytterligere innføring i gevinstrealisering.



Boks 5.1. Gevinstrealiseringsplan for tvisteloven

Nedenfor har vi satt opp tre tabeller som oppsummerer gevinstrealiseringssplanen for én av de planlagte gevinstene i

forbindelse med innføringen av ny tvistelov, nemlig ”raskere saksprosess”.

Mål / ønsket gevinst	Indikatorer på måloppnåelse	Hvor gevinst vil oppstå	Når i tid gevinsten kommer	Datakilde/ målemetode	Frekvens på måling	
	Tid mellom ulike milepæler, bl.a. fra: <ul style="list-style-type: none"> • stevning til start hovedforhandling (saksforberedelse) • start til avslutning hovedforhandling • avsluttet hovedforhandling til dom 	Domstolene og partene	2009–2012	Registeret til Domstols-administrasjonen	Følge opp hver 6. mnd.	Domstol-administrasjonen
Raskere saksprosess	Hvor godt forberedt er saken ved stevning (jf. nye plikter for partene før sak reises)	Domstolene og partene	Fra 2008	Spørreundersøkelse blant dommerne	Høst 2007 (før-måling) og etter 1,5 år og etter 3 år	Justis-departementet er ansvarlig for å gjennomføre spørreundersøkelse
	Økende bruk av skriftlige redegjørelser	Domstolene og partene	Fra 2008	Finnes ikke i registeret til Domstol-administrasjonen i dag, bør innføres	Følge opp hver 6. mnd.	Domstol-administrasjonen

Nødvendige tilretteleggende tiltak	Ansvarlig tilretteleggende tiltak	Frist for gjennomføring
Opplæring dommere	Domstoladministrasjonen	1.1.2008
Informasjon til parter og advokater om ny lov, bl.a. partenes plikter	Den enkelte domstol	1.3.2008
Investering i fjernmøteteknologi for å dra nytte av utvidet adgang til fjernmøter og -avhør	Domstoladministrasjonen	1.5.2008
Nedarbeide restanser	Den enkelte domstol	1.1.2008

Risikoreduserende tiltak	Ansvarlig risikoreduserende tiltak	Frist for gjennomføring
Manglende motivasjon for kulturendring hos dommerne		

5.3 Oppfølging og anvendelse av evalueringen

Når evalueringen er ferdigstilt og resultatene foreligger, er det viktig at man har en plan for hvordan resultatene skal inngå i styrings- og kontrollgrunnlaget til departementet eller andre som har ansvar for regelverket. Evalueringen kan ha pekt på behov for justeringer eller endringer i lover og forskrifters innhold, utforming, sanksjonenes styrke, måter regelverket forvaltes på, eller andre forhold som bør bearbeides på en systematisk måte. Funnene bør gjerne omtales både i dokumenter til Stortinget, i tildelingsbrev til regelverksforvalter (dersom dette er en underliggende virksomhet) og i andre relevante sammenhenger.

Et formål med evaluering kan også være læring og oppbygging av ny kunnskap. Da er det viktig at resultatene formidles til alle aktuelle aktører, både internt og eksternt. Det kan for eksempel være aktuelt å drøfte resultatene på et avsluttende seminar hvor ulike berørte aktører inviteres. Bare når det er helt spesielle begrunnelser for det, bør rapporten unntas offentlighet, selv om offentlighetsloven formelt skulle gi adgang til det.

I tillegg bør man sammen med evaluator vurdere av erfaringene med evalueringssprosessen. Var bestillingen, medregnet evalueringsspørsmålene, klare og entydige? Var det realistisk å kunne besvare spørsmålene med det angitte ambisjonsnivået innen den gitte budsjettammen? Var datamaterialet tilstrekkelig presist og relevant? Kunne resultatene vært fremstilt på en annen måte som var mer tilpasset bestillerens behov?



6 Referanser og bakgrunnsdokumenter til eksemplene

Finansdepartementet (2005a): *Veileder til gjennomføring av evalueringer*

<http://www.regjeringen.no/>

En kort og generell oversikt over gjennomføringen av en evaluering.

Finansdepartementet (2005b): *Veileder i samfunnsøkonomisk analyse*

Dette gir en generell og overordnet innføring i samfunnsøkonomiske analyser.

Ot.prp. nr. 51 (2004–2005): *Om lov om mekling og rettergang i sivile tvister (tvisteloven)*

Bakgrunnsdokument for eksemplet i denne veilederen.

Ot.prp. nr. 74 (2005–2006): *Om lov om endringer i tvisteloven (endringer i straffeprosessloven og andre lover)*

Bakgrunnsdokument for eksemplet i denne veilederen.

Ringdal, Kristen (2000): *Enhet og mangfold – samfunnsvitenskapelig forskning og kvantitativ metode*. Bergen: Fagbokforlaget.

SSØ (2006a): *Veileder i samfunnsøkonomisk analyse og gevinstrealisering av IKT-prosjekter*

<http://www.sfsso.no>

SSØ (2006b): *Mål og resultatstyring i staten – En veileder i resultatmåling*

<http://www.sfsso.no>

SSØ (2007a): *Plan for evaluering av tvisteloven* (foreløpig upublisert notat)

SSØ (2007b): *Evaluering av statlige tilskuddsordninger. Veileder*

<http://www.sfsso.no>



Vedlegg A Sjekkliste for planlegging, gjennomføring og oppfølging av en evaluering

Oversikt over evalueringsprosessen

En evalueringsprosess består i grove trekk av fire faser regnet fra det tidspunktet det tas en beslutning om å iverksette evaluering:

- Fase 1: utarbeide oppdragsbeskrivelse
- Fase 2: planlegge evaluering
- Fase 3: gjennomføre evaluering og kvalitetssikre underveis

Man bør følge opp evalueringen med en fjerde fase etter de tre første fasene:

- Fase 4: følge opp i etterkant av evalueringen og realisere gevinster

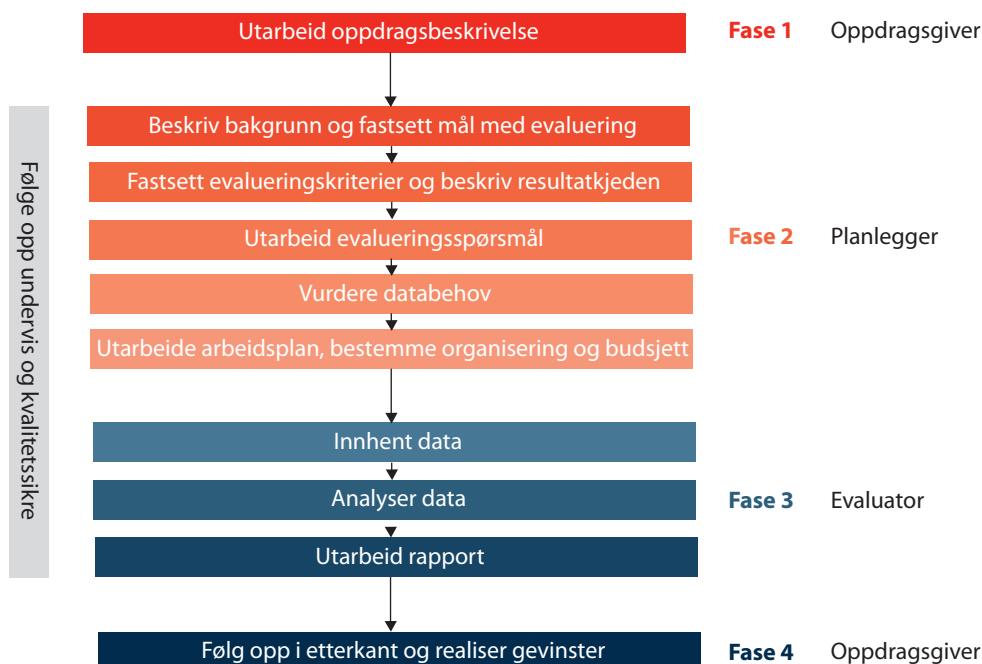
Aktørene som er involvert i evalueringsprosessen, kan ha tre ulike roller:

- Oppdragsgiver: Dette er den personen som utarbeider oppdragsbeskrivelsen og følger opp i etterkant av evalueringen og realiserer gevinster.
- Planlegger: Dette er den personen som på vegne av oppdragsgiver sørger for at evalueringen planlegges og iverksettes, og som skaffer en evaluator.
- Evaluator: Dette er den personen som gjennomfører evalueringen.

I en del tilfeller vil det være de samme personene som har de ulike rollene.

Figur A.1 viser trinnene i de fire fasene av en evaluering og hvilke roller som er aktuelle i hver fase.

Figur A.1. Faser, trinn og aktører i en evaluering



Fase 1: utarbeide oppdragsbeskrivelse

Bakgrunnen for evalueringen er i noen tilfeller et eksternt pålegg fra lovgiver og i andre tilfeller et resultat av interne strategiske vurderinger i departementet av behovet for å gjennomføre en evaluering. Viktige elementer i slike strategiske vurderinger er blant annet virksomhetens eller tiltakets egenart, risiko og vesentlighet.

Oppdragsbeskrivelsen (det overordnede mandatet) bør inneholde minimum følgende elementer:

- evalueringens formål og eventuelt avgrensning (ambisjonsnivå)
- tidsramme og eventuelt ressursramme for evalueringen
- ansvarsforhold og roller og dessuten eventuelt juridiske eller organisatoriske føringer som den planleggingsansvarlige må forholde seg til

Det kan i noen tilfeller være aktuelt at oppdragsbeskrivelsen legger utfyllende føringer på innholdet i evalueringen. Det er imidlertid uheldig om en slik beskrivelse blir for detaljert eller skaper sterke bindinger for den som skal planlegge gjennomføringen. Oppdragsgiver er som regel ikke den som har de beste forutsetningene for å legge de operative rammene for evalueringen.



Fase 2: planlegge evaluering

Beskriv bakgrunn, mål og berørte aktører

- Beskriv bakgrunnen for evalueringen med utgangspunkt i mandatet.
- Avgrens formålet ytterligere hvis det er nødvendig.
- Identifiser berørte aktører.
-

Fastsett evalueringstema og beskriv resultatkjeden

- Bestem hvilke hovedtema evalueringen skal konsentrere seg om, dvs. hvilke aspekter ved området evalueringen skal søke å belyse. Slike kan for eksempel være
 - effekter/samfunnsverdier
 - måloppnåelse
 - bærekraftighet
 - kvalitet/produktivitet

Når evalueringstema er bestemt, kan det benyttes ulike analysetyper i evalueringen, for eksempel måloppfyllelsesanalyse, effektanalyse og prosessanalyse (se Finansdepartementet, 2005a).

- Vurder å lage en såkalt resultatkjede¹²
 - Beskriv hvilke aktiviteter og produkter (tjenester) loven gir opphav til hos forvaltningsenheten og/eller hos målgrupper på utsiden.
 - Drøft hvilke effekter dette har for brukere og samfunn.
 - Beskriv hvordan (gjennom hvilke sammenhenger) effektene oppnås.

Utarbeid evalueringsspørsmål

- Formuler konkrete evalueringsspørsmål ut fra målene, evalueringstemaet, resultatkjeden og de valgte analysetypene.
- Sjekk at spørsmålene har et tilstrekkelig presisjonsnivå, og at de er mulige å besvare.
- Prioriter spørsmålene og vurder om det er overlappinger.

Vurder behov for eventuelle indikatorer og databehov

- Gjør en overordnet kartlegging av hvilken type informasjon som trengs, og hvilke datakilder (registre, tidligere studier mv.) som er tilgjengelige.
- Hvilke egne innsamlinger (f.eks. spørreundersøkelser) kan det være aktuelt å gjennomføre?
- Dersom evalueringsspørsmålene ikke kan besvares direkte, må man finne indikatorer som man kan skaffe data om og som kan brukes til å belyse spørsmålet.

Utarbeid arbeidsplan og bestem organisering og budsjett

- Fastsett et omfang av evalueringen som er faglig forsvarlig ut fra hva som synes å være ressursmessig realistisk.
- Anslå ressursbehovet og utarbeid et budsjett.
- Fastsett de viktigste milepælene.
- Ta stilling til hvilken kompetanse som kreves for å gjennomføre evalueringen:

- metodefaglig kompetanse (evalueringsmetodikk)
- kunnskap om saksområdet/regelverket
- Gjør et valg mellom intern og ekstern evaluator. Sterk vektlegging av sakkunnskap vil normalt trekke i retning av intern evaluator, mens sterk vektlegging av kompetanse på evalueringsmetodikk vil oftest trekke i retning av ekstern evaluator. Andre aspekter som kan legges til grunn, er betydningen av distanse til det som evalueres, og behovet for å bygge opp intern evalueringskompetanse.
- Lag oppdragsbestillingen/konkurransegrunnlaget.

Fase 3: gjennomføre evaluering og kvalitetssikre underveis

Ofte er det ved oppstart av selve gjennomføringsarbeidet fornuftig å definere en organisatorisk struktur rundt evalueringsprosjektet.

Innhent data

- Få oversikt over tilgjengelige data. Dette gjelder i hovedsak data som
 - er fremskaffet av oppdragsgiver eller innhentet i planleggingsfasen
 - kan lastes ned fra internett, er tilgjengelige i publikasjoner, mv.
 - kan bestilles fra en dataleverandør (f.eks. fra offentlige registre)
 - kan innhentes gjennom spørreundersøkelser, intervjuer, o.l.
- Gjør en prioritering. Hvilke data *må* vi ha for å belyse evalueringsspørsmålene? I noen tilfeller vil det være tilstrekkelig å benytte seg av allerede eksisterende data (sekundærdata). I andre tilfeller vil det være behov for å innhente data spesielt til analysens formål (primærdata).
- Velg metode(r) for datainnsamling.
 - Metoden(e) må tilpasses informasjonsbehov, krav til kvalitet og presisjonsnivå, tids- og budsjetttrammer, evaluators kompetanse mv.
 - Er det behov for kvantitative data (målbare i form av tall) eller kvalitative data (tekstlige beskrivelser og formuleringer)?
 - Gjør en vurdering av ressursbehov/kostnader for de ulike metodene.
- Sørg for en entydig og detaljert beskrivelse av de aktuelle data (definisjoner, avgrensninger, inndeling i underkategorier mv.).
- Ved bestilling av data fra dataleverandør (registre mv.) må man finne ut hvilke datatekniske utfordringer man står overfor (bl.a. krav til programvare), og man må bli enig med leverandøren om datatekniske format, utskriftskvalitet o.l.
- Lag en oppsummering av svakhetene ved data (skjevhet, unøyaktighet mv.).

¹²Se vedlegg B.



Analyser data

- Kvalitetssikre og bearbeid data, eliminer feil i datamaterialet, og tolk dataenes informasjonsverdi.
- Utled statistiske mål som gjennomsnitt og standardavvik for enkeltvariabler der dette er aktuelt.
- Søk etter mønstre og sammenhenger i datamaterialet.
- Sett data inn i en sammenheng – vanligvis en teori, en utvikling (historikk) eller en tilsvarende størrelse for en sammenliknbar aktivitet (benchmarking mv.).
- Vurder om resultatene gir grunnlag for klare konklusjoner, og om de kan generaliseres.
- Hvor ligger den største usikkerheten?
- Har resultatene vesentlige mangler eller svakheter (grunnet datamaterialet, teori- eller metodevalg mv.)?

Utarbeid rapport

- Utarbeid en foreløpig struktur på sluttrapport på et tidlig tidspunkt.
- Kontroller at sluttrapporten gjenspeiler innholdet i oppdragsbestillingen og arbeidsplanen, bl.a. at den svarer på evalueringsspørsmålene.

Følg opp underveis og kvalitetssikre

Oppdragsgiver/planlegger må følge opp evaluator underveis:

- Kontroller fremdrift i prosjektet og gi løpende tilbakemelding til evaluator.
- Vurder behovet for referansegruppe eller annen form for ekstern kvalitetssikring.

Fase 4: følge opp i etterkant av evalueringen og realisere gevinster

Følg opp i etterkant av evalueringen

- Gi tilbakemelding til evaluator på det ferdige produktet.
- Spre informasjon om evalueringen internt og eksternt dersom det ikke er særlige grunner til at rapporten ikke bør offentliggjøres.

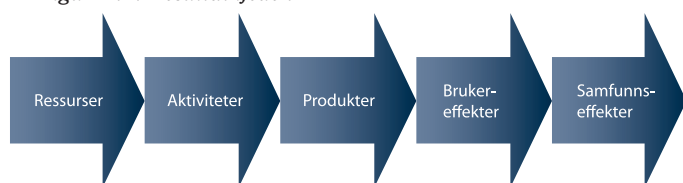
Anvend evalueringen ut fra formålet med den

- Anvend evalueringen til kontroll.
- Anvend evalueringen til læring for å styrke oppbyggingen av kompetanse og kunnskap om regelverket og dets virkninger (og anvend den eventuelt innen evalueringsmetodikk).
- Anvend evalueringen i styringsprosesser for
 - å vurdere å endre regelverket dersom en annen utforming kan gi gevinster for samfunnet
 - å vurdere å bruke andre virkemidler hvis disse fremstår som mer effektive eller rettferdige

Vedlegg B Resultatkjeden

Resultatkjeden er et nyttig verktøy for å beskrive hvordan en virksomhet via ulike aktiviteter omformer ressurser (innsatsfaktorer) til sluttprodukter eller tjenester som er rettet mot gitte brukergrupper og i siste instans samfunnet som helhet. Produktene er ment å gi visse bruker- og samfunnseffekter. Resultatkjeden er altså en årsakskjede, og pilene fra venstre mot høyre viser årsakssammenhengene. Den generelle modellen fremgår av figur B.1 nedenfor.

Figur B.1. Resultatkjeden



Det er kun de to (til dels tre) første trinnene i kjeden som man har direkte kontroll over. Bruker- og samfunnseffektene vil i tillegg til kontrollerbare faktorer også avhenge av en rekke eksterne faktorer, som målgruppens holdninger, andre aktørers respons på lovreguleringen og samfunnsutviklingen for øvrig.

Med utgangspunkt i resultatkjeden kan vi vurdere hvilke trinn eller sammenhenger vi er særlig interessert i, og deretter hvilke evalueringstema og evalueringsspørsmål som er relevante. Å sette opp resultatkjeden kan også ha den fordelen at vi tvinges til å gjøre eksplisitte antakelser om hvilke aktører som påvirkes, i hvilke trinn de påvirkes, og dessuten om responsen fra, eller utviklingen i, eksterne faktorer. Kanskje kommer vi til at noen av disse er helt kritiske for at effekter skal oppnås, noe som taler for at de bør måles og følges opp.

Ved analyse av lover kan modellen være et nyttig hjelpemiddel til å få frem de økonomiske egenskapene og virkninger. Ofte vil de overordnede målene med en lovregulering eller med saksbehandlingsproduksjonen hos en offentlig forvalter være identiske med de tiltenkte samfunnseffektene. I så fall vil en effektanalyse kunne være identisk med en måloppfyllelsesanalyse, og den bør rettes mot effektleddene i resultatkjeden. De overordnede målene kan for eksempel reflekteres i lovens formålsparagraf eller andre rettskilder. Ofte vil imidlertid målene på samfunnseffektnivå være nokså vagt og abstrakt formulerte, og de vil være realiserbare bare på lang sikt. Det gjør at de er vanskelige å måle. Vi kan derfor også knytte mål til elementene lenger til venstre i kjeden og på den måten snakke om en måloppfyllelsesanalyse her òg.

I tillegg til å være opptatt av måloppfyllelse og effekter er man ofte interessert i forvaltningens produktivitet eller kostnadseffektivitet. Det handler om sammenhengen mellom ressursbruken (jf. første ledd i kjeden) og produktene/tjenestene (jf. tredje ledd). Det er viktig å legge merke til at det forutsettes at tjenestene holder en gitt kvalitet – det betraktes altså *ikke* som

en forbedring i produktiviteten dersom ressursbruken per tjeneste går ned som følge av at kvaliteten forringes. For å bedre produktiviteten må man gå inn i saksprosessene og analysere årsakene til lav produktivitet.

I boks B.1 nedenfor forklarer vi oppsettet av resultatkjeden i tilknytning til evalueringen av tvisteloven.

Boks B.1. Resultatkjeden for domstolenes arbeid etter den nye tvisteloven

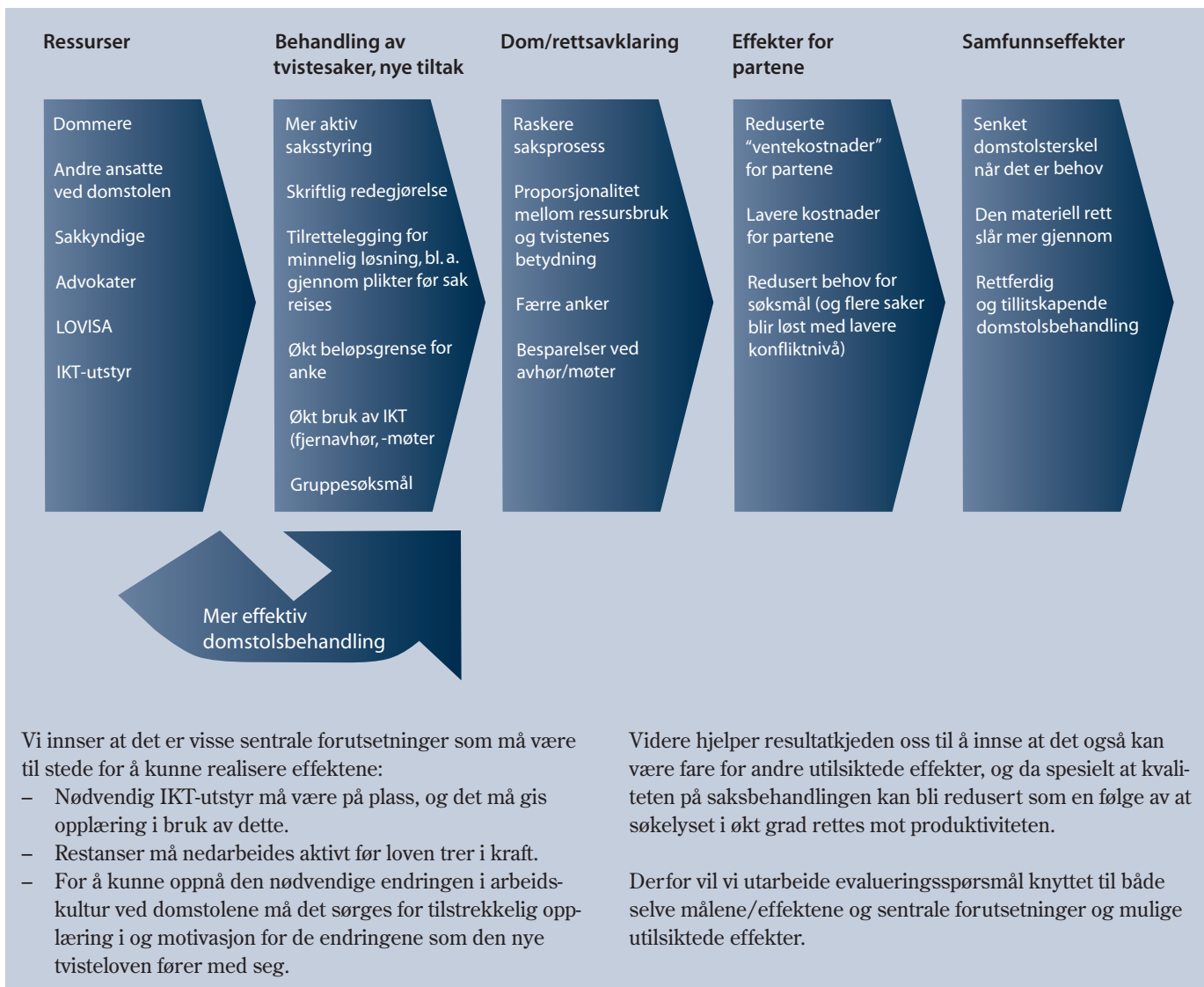
Figuren nedenfor viser resultatkjeden slik at den beskriver hvilke ressurser og konkrete aktiviteter/tiltak som settes inn med ny tvistelov, og hvilke tiltenkte effekter dette har. Ressurser her er dommere, advokater, sakkyndige, saksbehandlingsutstyr, medhjelpere og så videre. Disse bidrar inn i den ”produksjonen” som foregår i domstolsapparatet. Aktivitetsleddet inneholder de nye tiltakene som innføres som del av tvistelovsreformen, blant annet mer aktiv saksstyring, innføring av småkravsprosess, tilrettelegging for minnelig løsning og så videre.

Produktleddet er i dette tilfellet en beslutning eller avgjørelse som gir en avklaring av saken (dom, forlik, at saken trekkes e.l.). I figuren har vi på produktleddet plassert ulike egenskaper ved avklaring av saken som følge av tiltakene som ble iverksatt på forrige ledd. Det dreier seg blant annet om raskere prosess, økt proporsjonalitet mellom ressursbruk og betydningen av tvisten og færre anker.

Med brukereffekter menes de tiltenkte virkningene av ny lov på partene i tvisten. De mest sentrale brukereffektene er reduserte ”ventekostnader” for partene (i form av kortere tid til saksforberedelse og deretter raskere gjennomføring av hovedforhandling), lavere sakskostnader for partene og flere saker som kan avsluttes før sak reises (slik reduseres behovet for søksmål, og tvistene kan avklares med et langt lavere konfliktnivå enn tidligere).

Med samfunnseffekter menes de mer indirekte og langsiktige virkningene av ny lov for alle samfunnsborgere som i teorien vil kunne ha behov for rettssystemet i fremtiden, og som dermed er avhengige av tillit til dette. De viktigste samfunnseffektene er senket domstolsterskel når det er behov for det (følger bl.a. av reduserte ”ventekostnader” og lavere sakskostnader), større gjennomslag for den materielle retten og dessuten at en god og rettferdig domstolsbehandling skaper tillit til hele systemet og til rettssikkerheten i Norge.

I tillegg er det knyttet mål til sammenhengen mellom ressursbruk og tjenesteproduksjon, nemlig at domstolsbehandlingen skal bli mer produktiv. Det vises i figuren som en pil mellom disse to leddene.



Se også SSØ (2006b) for en ytterligere innføring i bruk av resultatkjeden.



Vedlegg C Eksempler på mulige evalueringsspørsmål

Det er viktig å merke seg at man bør forsøke å begrense antallet evalueringsspørsmål til det som er nødvendig for å besvare det overordnede målet. Oppgaven for evaluator skal være overkommelig gitt tilgjengelig tid og ressurser, og man må derfor prioritere det som er mest sentralt.

Nedenfor har vi listet opp en rekke eksempler på evalueringsspørsmål som kan være mer eller mindre aktuelle i forbindelse med evaluering av lovregulering og -forvaltning.

- Er behov og formål med lovreguleringen klart definert?
 - Er problemene og behovsbeskrivelsen reell? Er det mulighet for at problemene kan løses av seg selv over tid?
 - Er målene operasjonelle?
 - Er målene realistiske?
- Er lovreguleringen nødvendig?
 - Unødvendig duplisering av annet regelverk mv.
- I hvilken grad etterleves loven?
- Hvilke gevinster/effekter har loven for ulike brukergrupper?
 - Hvilke mekanismer stimulerer til effektene, og hvilke incentiver skapes? Har brukerne endret atferd etter at loven trådte i kraft?
 - Hvor lang tid tar det før effektene oppstår? Hvor lenge varer de?
 - Har effektene hatt den tiltenkte styrken?
 - Hva er årsaken til at effektene eventuelt uteblir?
 - Har loven hatt utilsiktede effekter (positive eller negative)?
- I hvilken grad bidrar effektene til oppnåelse av samfunns-målet?
- Vurder om loven er i tråd med EU-direktiv eller andre overordnede føringer.
- Hvordan opplever hhv. de profesjonelle brukerne og de ikke-profesjonelle brukerne at regelverket fungerer?
 - Hvordan er regelverkets brukervennlighet, brukernes kjennskap til loven mv.?
 - Har økt kompleksitet i regelverket medført større behov for advokatbistand?
- Hvilke andre sektorområder påvirkes av den aktuelle lovbestemmelsen?
- Hvilke kostnader har lovreguleringen for brukere, forvaltere og andre?
 - spesielt om de statsfinansielle kostnadene
 - spesielt om de administrative byrdene (rapporteringskrav mv.) som reguleringen påfører næringsliv og privatpersoner
- Hva er nytte-kostnad-forholdet ved reguleringen? Gjennomfør en samfunnsøkonomisk analyse hvor alle virkninger tallfestes og verdsettes så langt det er mulig.
- Hvilke andre alternative virkemidler finnes, og hva er gevinster og kostnader av disse?
 - økonomiske: skatter, avgifter, tilskudd, investeringer og stønadsordninger
 - ”myke” reguleringer som anbefalinger, standarder, holdningskampanjer, informasjon
- I hvilken grad ser man et samspill med andre virkemidler (effekter som forsterker hverandre e.l.)?
- Hvilke fordelingsvirkninger har lovreguleringen?
 - Har vi identifisert alle grupper som berøres?
 - Er interessekonflikter kartlagt?
 - Man må vurdere om uønskede fordelingsvirkninger kan motvirkes gjennom alternativ utforming eller særlige kompensasjons-/avhjelpingstiltak rettet mot de som rammes, f.eks. unntak for små bedrifter.
- Hvordan er kostnadseffektiviteten i saksbehandlingsprosessene og myndighetenes gjennomføring?
- Følges interne rutiner i forvaltningen i forbindelse med saksbehandlingen?
- Hvordan er produktiviteten/kostnadseffektiviteten i forvaltningen av loven?
 - ressursutnyttelse
- Hvordan er organiseringen og rollefordelingen?
 - Hvilket forvaltningsnivå er best egnet til å forvalte det aktuelle regelverket? Man bør vurdere bl.a. synergier i forbindelse med forvaltning av annet regelverk.
- Hvordan er kvaliteten i forvaltningen?
 - likebehandling, avgjørelsens riktighet, bl.a. vurdering av klagesaker
 - betydningen av saksbehandlernes kompetanse
 - hvilke kvalitetssikringsprosesser som er iverksatt
 - brukernes tilfredshet med kvalitet, svartider mv.



Vedlegg D Nærmere om metodiske utfordringer ved gjennomføring av en evaluering

D.1 Kvantitativ eller kvalitativ evalueringsdesign

Det går et hovedskille mellom såkalte kvantitative og kvalitative evalueringsdesign. Dette valget har i stor grad sammenheng med hva som er formålet med evalueringen, og hva resultatene skal brukes til.

I et kvantitativ design vil man typisk ha en ferdig formulert hypotese om årsak og virkning som man ønsker å teste. Til dette trenger man mange observasjoner/respondenter for å finne omfang, frekvenser og så videre og for å kunne generalisere ut fra resultatene. Dette tilsier ofte bruk av statistikk eller store, standardiserte spørreundersøkelser (surveyer). Evaluatør vil være nøye på at observasjonene er representative, og at resultatene som presenteres, skal være "objektive fakta". Ved spørreundersøkelsene er det for eksempel viktig å stille nøyaktige de samme spørsmålene til alle respondentene og å stille dem på samme måte, slik at faren for påvirkning fra utenforliggende forhold minimeres.

Kvalitative design er rettet mot å gå i dybden på den enkelte observasjonen for å kunne trekke ut informasjon og (ved intervjuobjekter) også vurderinger og innspill til mulige årsaker eller virkninger som evaluatør eventuelt ikke har sett. Det ideelle er selvsagt flest mulig observasjoner, men ofte vil tids- og ressursituasjonen tilsa en avveining mellom dybde og bredde, og da velges dybde. Det er gjerne ikke utformet noen ferdig hypotese på forhånd. Målet er primært å forstå og tolke en problemstilling. Datainnsamlingsteknikker vil typisk være observasjon, dokumentstudier (f.eks. ulike sakstiltfeller eller klagesaker), dybdeintervju og fokusintervju. Fokusintervju, eller gruppeintervju, er en metode som gjennom gruppedynamikk kan stimulere til flere innspill enn det intervjuer av én og én person kan. Det skyldes at det gjerne er lettere å bli seg bevisst egne meninger og å spinne videre på hverandres argumenter når det er flere intervjuobjekter til stede.

Valget mellom kvantitativt og kvalitativt design avhenger i stor grad av hvor god kjennskap man har til problemstillingen som skal undersøkes. Ved for eksempel vurdering av mulige utilsiktede effekter av en regulering starter evaluatør gjerne med nokså liten kunnskap og oversikt. Da er det viktig å skaffe innsikt og forståelse, noe man ofte oppnår best ved å studere eksisterende litteratur, observere eller snakke med aktuelle informanter på området (kvalitativt design). Dersom man allerede har innført et nytt tiltak med sikte på å rette opp gitte skjevheter, vil det derimot være mer interessant å gjøre en breddeundersøkelse for å verifisere om omfanget av problemet faktisk er redusert som følge av tiltakene (kvantitativt design).

Det er selvsagt ikke slik at man må velge strengt mellom *enten* kvantitativt *eller* kvalitativt design. Ofte vil det tvert imot være nyttig å kombinere de to typene datainnsamlingsmetoder.

Sammen kan de utfylle hverandre og bidra til å belyse saken på ulike måter.

D.2 Innsamling av primærdata: utvelgelse av enheter og tidspunkt for målinger

Ved innsamling av primærdata må man gjøre et valg med hensyn til hvilke og hvor mange observasjoner man skal inkludere i undersøkelsen, og dessuten hvilket tidspunkt undersøkelsen skal gjennomføres på.

Først må målgruppen (den såkalte populasjonen) defineres, for eksempel alle førsteinstansdomstoler i Norge eller alle saker i en bestemt kategori. Deretter bør man avklare forholdet til eventuelle bestemmelser om taushetsplikt etc. og stille spørsmål ved om det er behov for å inkludere *hele* populasjonen i undersøkelsen. Dette har selvsagt også en kostnadsside, og innsamling av store mengder data tar tid. I noen tilfeller er imidlertid populasjonen så liten at det er naturlig å inkludere alle, for eksempel dersom man er interessert i saker som omfatter en gitt straffebestemmelse som svært sjelden er i bruk.

Som nevnt ovenfor skiller kvantitative og kvalitative design seg fra hverandre ved at man er opptatt av henholdsvis bredde/representativitet og dybde. Det tilsier at man i en kvantitativ undersøkelse er mest opptatt av å inkludere mange enheter, mens man i en kvalitativ undersøkelse er mest opptatt av å finne noen få, men riktige enheter som kan bidra til å belyse problemstillingen. I en casestudie vil man ofte bare ha med én enhet.

Vi kan skille mellom ulike typer utvalg. De to hovedkategoriene er sannsynlighetsutvalg og ikke-sannsynlighetsutvalg. Disse omtales nedenfor.

Sannsynlighetsutvalg

Dersom man er opptatt av representativitet, bør man tilstrebe et sannsynlighetsutvalg.

En type sannsynlighetsutvalg er et såkalt tilfeldig utvalg. Det krever at vi kan foreta en helt tilfeldig trekning blant alle enhetene i populasjonen. Dersom viktige faktorer som påvirker de forholdene vi er interessert i (f.eks. tidsbruk, brukertilfredshet, sammensetning av saker e.l.), varierer gjennom året, bør man i tillegg spre trekningene over tid.

Ofte er vi imidlertid også opptatt av å kunne belyse forskjellen mellom ulike undergrupper, for eksempel friske personer versus personer med en bestemt funksjonshemming, eller "tunge" saker versus "lette" saker. Da er det viktig å tilstrekkelig mange enheter innenfor hver av undergruppene er representert i utvalget, også for relativt sjeldne undergrupper. For å sikre dette kan man foreta et såkalt stratifisert utvalg. Det innebærer å forhåndsdefinere et antall undergrupper (strata) som man deretter foretar en tilfeldig trekning fra. Utvalget av enheter fra



sjeldne undergrupper gjøres gjerne større enn denne gruppens forekomst skulle tilsi. Det skyldes at man ønsker å trekke generelle konklusjoner om den aktuelle gruppen.¹³

En tommelfingerregel er at utvalget må være større jo flere egenskaper ved populasjonen man senere ønsker å trekke inn i analysen, uavhengig av om man stratifiserer eller ikke.

En annen og ofte mindre ressurskrevende form for sannsynlighetsutvalg er såkalte klyngeutvalg. Man foretar da tilfeldige trekninger i to trinn: først velges et utvalg av klynger (f.eks. domstoler, kommuner e.l.), og deretter trekker man ut et utvalg av enheter innenfor disse klyngene. Dersom klyngene skiller seg fra hverandre i betydelig grad, vil denne metoden være noe mindre god med tanke på representativitet.

Ikke-sannsynlighetsutvalg

I mange tilfeller vil man også benytte såkalte ikke-sannsynlighetsutvalg, noe som innebærer at observasjonshetene velges ut fra hva som er praktisk og økonomisk mest gunstig. Eksempler her er å velge intervjuobjekter ut fra hvilke som frivillig melder sin interesse for å delta, eller for eksempel å velge blant studenter i en gitt forelesningstime. Vi vil da ikke kunne generalisere resultatene med samme statistiske sikkerhet, men ofte kan vi likevel få mye nyttig informasjon gjennom slike undersøkelser.

D.3 Datas validitet og reliabilitet

Vi bør alltid gjøre en vurdering av kvaliteten på de dataene som benyttes i undersøkelsen. Drøftingen av dette bør dreie seg om datas validitet og datas reliabilitet.

Validitet handler om resultatenes gyldighet eller relevans, det vil si om vi kan trekke konklusjoner fra undersøkelsen og tilbake til teorien. Dette avhenger i sin tur av om måleparametrene som er brukt, virkelig måler det vi var interessert i å si noe om. Dersom man for eksempel ønsker å si noe om brukernes tilfredshet med domstolene, vil "antall brukere som reiser ny sak for samme domstol innen X år" ikke nødvendigvis være et godt mål på dette. Brukerne har jo i liten grad mulighet til å velge andre domstoler.

Reliabilitet dreier seg om selve målingens nøyaktighet eller pålitelighet, og den kan uttrykkes ved i hvilken grad man får det

samme resultatet dersom forsøket gjentas flere ganger. For å sikre reliabiliteten bør man alltid vurdere kritisk både kildene, svarprosenten, selve datainnsamlingssystemet og måten data måles og registreres på. Om mulig bør dataene kvalitetssikres, for eksempel ved kildetriangulering¹⁴, ved stikkprøver eller ved å be respondenten om å verifisere referat fra intervju.

Validitet handler om "hvor godt vi måler det vi ønsker å måle". Reliabilitet handler om "hvor godt vi måler det vi måler".

Spørsmålet om datainnsamlingsmetoder vil til syvende og sist alltid handle om en avveining mellom behov for valid og reliabel dokumentasjon på den ene siden og forhold som kostnad, tid og tilgjengelighet på den andre.

D.4 Effektanalyse – metoder for å kontrollere for eksterne faktorer

Eksperimentelle opplegg

Denne metoden krever at man kan etablere henholdsvis én eksperimentgruppe og én kontrollgruppe fra samme populasjon ved hjelp av tilfeldig uttrekk. Deretter utsettes kun eksperimentgruppen for den aktuelle reguleringen eller det aktuelle tiltaket. På den måten kan man sammenlikne utviklingen over tid hos de to gruppene. Siden disse ikke er systematisk forskjellige, kan vi legge til grunn at alt annet enn tiltaket påvirker begge gruppene likt. En utvikling som kun observeres hos eksperimentgruppen, må derfor skyldes tiltaket. Vi kan også snakke om kvasi-eksperimentelle opplegg hvor gruppene ikke er helt tilfeldig trukket, men hvor det likevel kan være nyttig å sammenlikne dem.

Eksperimentelle opplegg er vanligvis ikke mulig når det gjelder lovreguleringer. Det er fordi en lov nødvendigvis gjelder alle personer og geografiske områder og så videre når den først trer i kraft. Man har derfor ingen kontrollgruppe. Eksperimentelle opplegg kan derimot være aktuelt når det gjelder nye arbeidsmetoder eller konkrete virkemidler fra lovforvalters side. Slike kan for eksempel innføres kun i et utvalg av regioner i første omgang, typisk som et pilotforsøk, før man vurderer å la ordningen gjelde alle.

I boks D.1 nedenfor viser vi et hypotetisk eksempel på bruk av eksperimentelle opplegg ved evaluering av tvisteloven.

¹³ Dette kalles disproporsjonal stratifisering og korrigeres senere i analysen ved veiing.

¹⁴ Dette betyr å kombinere ulike typer data for å belyse samme fenomen på ulike måter, for eksempel med både kvalitative og kvantitative metoder. To metoder kan være likeverdige, eller den ene kan fungere bare som en kvalitetssikring av den andre.



Boks D.1. Eksperimentelt opplegg

Anta at investering i nytt IKT-utstyr gjennomføres kun ved $\frac{1}{4}$ av domstolene i første omgang. Vi forsøker å få det til slik at utvalget av domstoler som får IKT-utstyr, består av både store og små domstoler, domstoler fra ulike landsdeler og så videre. På den måten blir de som gruppe mest mulig lik gruppen som ikke får utstyr.

Etter 1,5 år ønsker vi å se om investeringen har bidratt til tidsbesparelser ved domstolene. Det krever at vi måler gjennomsnittlig tidsbruk under hovedforhandling i gruppen med og uten IKT-utstyr både før investeringen og etter 1,5 år.

Før investeringen er gjennomsnittlig tidsbruk omtrent lik i de to gruppene av domstoler, cirka 30 dager. Etter 1,5 år er gjennomsnittlig tidsbruk fortsatt omtrent lik i begge gruppene, nå cirka 22 dager. Reduksjonen skyldes trolig at innføringen av ny tvistelov har skjedd i samme periode. Siden reduksjonen ikke er større for domstolene som har fått IKT-utstyr, kan vi imidlertid konkludere med at denne investeringen ikke har hatt noen effekt på tidsbruken

Regresjonsanalyse

Denne teknikken innebærer at vi eksplisitt modellerer tilstanden vi er interessert i, som en funksjon av alle mulige forklaringsvariabler. En av variablene er lovreguleringen, mens de andre er eksterne faktorer som vi antar også vil kunne påvirke tilstanden. Ved å modellere alle mulige forklaringsfaktorer eksplisitt kan vi kontrollere for disse og dermed skille ut effekten av lovreguleringen. Bruk av denne metoden krever at alle eksterne faktorer kan operasjonaliseres og måles kvantitativt, og at vi faktisk finner, og inkluderer i modellen, alle de sentrale eksterne faktorene.

I boks D.2 nedenfor viser vi et hypotetisk eksempel på bruk av regresjonsanalyse ved evaluering av tvisteloven.

Boks D.2. Regresjonsanalyse – tvisteloven

Vi ønsker å måle om partene i ulike typer saker har fått lavere saks kostnader. Vi legger til grunn at faktorene som bidrar til endring i saks kostnader er ny tvistelov og sakens kompleksitet. Betalt salær finner vi ved spørreundersøkelse blant et representativt utvalg av parter, og dette måles både før innføring av ny tvistelov, etter 1,5 år og etter 3 år. Anta at gjennomsnittlige saks kostnader viser seg å være uendret etter 1,5 år, og at disse er redusert med 1 prosent etter 3 år (korrigert for inflasjon). Tilsynelatende har tiltakene i ny tvistelov altså ikke hatt særlig stor effekt. I samme periode har imidlertid sakens kompleksitet økt, og det vil isolert sett bidra til å øke saks kostnadene.

Vi lar X^A være forklaringsvariabelen "ny tvistelov" og X^B være "sakens kompleksitet". X^A vil ha verdi 0 for observasjonene ved første måling, deretter verdi 1 i de to neste målingene. Vi lar X^B representeres ved en 10-delt skala (0 er en svært enkel sak, og 10 er en svært kompleks sak), og vi antar at verdien registreres etter en vurdering fra dommerens side. Y er vekst i saks kostnader per år.

En enkel lineær regresjonsmodell vil kunne se slik ut:1)

$$Y = \beta^0 + \beta^A * X^A + \beta^B * X^B$$

Estimering av modellen gir parametrene $\beta^A = -3$ og $\beta^B = 2,5$, og begge estimatene er signifikante. Tolkningen av β^A er at innføring av ny tvistelov har ført til en klar nedgang i saks kostnadene. Tolkningen av β^B er at en økning i sakens kompleksitet med 1 prosent vil øke saks kostnadene med 2,5 prosent.

Vi har dermed fått en enkel modell til å forklare utviklingen i saks kostnadene. Innføring av tvisteloven har isolert sett ført til lavere saks kostnader, men det at kompleksiteten i sakene har økt, har også hatt en signifikant effekt. I fravær av ny tvistelov ville man imidlertid ha sett en betydelig økning i saks kostnadene.

Utarbeidet av Justis- og politidepartementet
og Senter for statlig økonomistyring

Offentlige institusjoner kan bestille flere
eksemplarer fra:
Departementenes servicesenter
Post og distribusjon
E-post: publikasjonsbestilling@dss.dep.no
Faks: 22 24 27 86

Publikasjonskode: G-0415 B
Trykk: Grøset 01/2009 - opplag 2000