

© Concept-programmet 2007

Concept rapport nr. 19

Effektvurdering av store statlige investeringsprosjekter

Bjørn Andersen, Svein Bråthen, Tom Fagerhaug, Ola Nafstad, Petter Næss og Nils Olsson

ISSN: 0803-9763 (papirversjon)

ISSN: 0804-5585 (nettversjon)

ISBN: 978-82-92506-57-8 (papirversjon)

ISBN: 978-82-92506-58-5 (nettversjon)

Sammendrag: Denne rapporten beskriver eksisterende metoder og praksis for effektvurdering av offentlige prosjekter, med hovedvekt på metoder for evaluering i ettertid (ex post-effektvurderinger). Basert på gjennomgangen av eksisterende metoder og praksis presenteres et forslag til metodikk for ex post-evaluering av store offentlige investeringsprosjekter i Norge. Fokus er rettet mot prosjektenes samfunnsmessige effekter, dvs. det som betegnes som prosjektenes eksterne effektivitet. Dette i motsetning til en vurdering av hvorvidt resursene i selve prosjektet er brukt effektivt (intern effektivitet). Den anbefalte metoden illustreres med eksempler på mulig bruk i evaluering av ulike typer prosjekter.

Dato: 20.12.2007

Utgiver: Concept-programmet

Institutt for bygg, anlegg og transport

Norges teknisk- naturvitenskapelige universitet

Høgskoleringen 7A

7491 NTNU – Trondheim

Tel. 73594640

Fax. 73597021

<http://www.concept.ntnu.no>

Ansvaret for informasjonen i rapportene som produseres på oppdrag fra Concept-programmet ligger hos oppdragstaker. Synspunkter og konklusjoner står for forfatterens regning og er ikke nødvendigvis sammenfallende med Concept-programmets eller referansegruppens syn.

Forord

Det fremstår som noe av et paradoks at mange prosjekter, også store offentlige prosjekter bare unntaksvis evalueres etter de første årenes drift for å få klarlagt den langsiktige effekten av investeringen. Dette er uheldig. Ex-post vurderinger av effekt og konsekvens bidrar til den såkalt ”store læringsloopen” som kan gi oss kunnskap om nytten av investeringen. Eller med andre ord om en har gjort det rette ved å investere i prosjektet. Mer ressurser brukes vanligvis til å vurdere den ”lille læringsloopen” der en ser på sammenhengen mellom ressursinnsats og prosjektets leveranse, med andre ord bare hvordan prosjektet ble gjennomført.

Det kan være mange årsaker til at det så sjelden gjøres systematiske ex-post analyser. En av de viktigste er kanskje at den ansvarlige part i driftsfasen ikke alltid er den samme som finansierende eller gjennomførende part. Det er heller ikke vanlig å sette av midler i prosjektbudsjettet til slike undersøkelser.

Finansdepartementets ordning med ekstern kvalitetssikring i tidligfasen av store statlige investeringsprosjekter gir en unik mulighet for å gjøre systematisk informasjonsinnsamling og analyse av store statlige prosjekter gjennom hele livsløpet. Dette er da også informasjonsgrunnlaget for arbeidet i forskningsprogrammet Concept. Både kvalitetssikringsordningen og forskningsprogrammet skal bidra til å øke lønnsomheten og nytten av offentlige investeringer. Selv om fokuset er tidligfasen så er det å skaffe tilveie informasjon om effekten av prosjektene i driftsfasen helt avgjørende for å lykkes i arbeidet med å bedre tilrettelegging og utvikling av gode prosjektkonsepter.

Dette er bakgrunnen for delstudien ”Effektvurdering av store statlige investeringsprosjekter” som ble gjennomført i 2007 av institutt for produksjons- og kvalitetsteknikk ved NTNU. Oppgaven er å beskrive og vurdere forskning, kunnskap og praksis internasjonalt og i Norge på dette området og gi anbefalinger om hvordan dette bør gjennomføres for store statlige prosjekter.

Denne rapporten presenterer resultatene fra prosjektet. Den er skrevet av professor Bjørn Andersen, NTNU, som har ledet prosjektet, Svein Bråthen, Møreforskning, Tom Fagerhaug, NTNU, Ola Nafstad, ECON Pöyry, Petter Næss, Aalborg Universitet og Nils Olsson, SINTEF.

Referansegruppen for prosjektet bestod av Kjell Ottar Sandvik, Vegvesenet, Arvid Strand, Transportøkonomisk institutt, May Balkøy, Statsbygg, Frode Hjelde, Jernbaneverket og Kaare Falkenberg, Fornyings- og administrasjonsdepartementet.

Prosjektet er finansiert gjennom forskningsprogrammet Concept, som har som mandat å utvikle kunnskap som sikrer bedre ressursutnytting og effekt av store statlige investeringer. Programmet utføres ved NTNU, Institutt for bygg, anlegg og transport, og finansieres av Finansdepartementet.

Trondheim, januar 2008

Knut Samset, professor, NTNU

Programansvarlig, Concept-programmet, NTNU.

Innhold

Forord	2
Innhold	4
Figuroversikt	6
Tabelloversikt	7
Sammendrag	8
Summary	16
1 Innledning	24
1.1 Bakgrunn	24
1.2 Målsetninger	26
1.3 Metode	27
2 Eksisterende teori om ex ante og ex post-evaluering og effektvurdering	28
2.1 Nytte-kostnadsanalyse	29
2.2 Samfunnseffekt-analyse	39
2.3 Prestasjonsvurdering	48
2.4 Benchmarking som virkemiddel i prestasjonsvurdering	56
3 Eksisterende praksis for ex post-vurderinger	62
3.1 Transportetater i amerikanske delstater	62
3.2 SIKAs oppfølging av transportpolitiske mål i Sverige	65
3.3 Sound Transit	65
3.4 Tennessee Rail Systems	66
3.5 Karakterboken i den danske byggebransjen	69
3.6 Independent Project Assessment: Evaluering og benchmarking av prosjekter	70
4 Eksempler på ex post-evalueringer i Norge	72
4.1 Prosjekt Nytt Rikshospital	72
4.2 Nytteeffekter av IT-prosjekter i Trygdeetaten	74
4.3 Troll Olje-prosjektet	75
4.4 Nitten materiellinvesteringsprosjekter i Forsvaret	76
4.5 Rehabilitering av det kongelige slott, Regjeringens representasjonsbolig og Stiftsgården	77
4.6 Planlegging og oppfølging av store veginvesteringer	78
4.7 Etterprøving av nytte-kostnadsanalyser av samferdselsprosjekter	79
4.8 Feltutbyggingene Åsgard, Visund og Jotun	81
4.9 Tetting av Romeriksporten	81
4.10 Anskaffelse av F-16 kampfly i perioden 1975-98	82
4.11 Winix	83
4.12 Tress-90	84

4.13	Effektvurdering av alternative utbyggingsprosjekter for Nasjonalmuseet for kunst, arkitektur og design	85
4.14	Samlet Plan for vassdrag	86
4.15	Hvilke erfaringer kan vi høste fra tidligere evalueringer?.....	87
5	Foreslått prosedyre og metode for ex post-evaluering	88
5.1	Hensikten med ex post-evaluering	88
5.2	Referansepunkt og sammenlikningsgrunnlag for ex post-evaluering	91
5.3	Tidspunkt for ex post-evaluering	92
5.4	Metode for ex post-evaluering	94
5.5	Ansvar for og roller i gjennomføring av ex post-evaluering.....	98
5.6	Konsekvenser av den foreslåtte tilnærmingen og forutsetninger for gode ex post-evalueringer.....	100
6	Eksempler på gjennomføring av ex post-evaluering med ny metode	102
6.1	Evaluering av Nytt Rikshospital.....	102
6.1.1	Foreliggende evalueringer	102
6.1.2	Anvendelse av vår foreslåtte metode på Rikshospital-prosjektet.....	105
6.1.3	Konsekvenser	107
6.2	Effektvurdering av fire jernbaneprosjekter	111
6.2.1	Vestfoldbanen og Sandeparsellene.....	111
6.2.2	Gardermobanen	115
6.2.3	Østfoldbanen, Ski-Sandbukta.....	118
6.2.4	Bergensbanen, Finsetunellen og Finse-Gråskallen.....	119
6.2.5	Sammenstilling.....	122
6.3	Evaluering av prosjektet for bygging av Frøyatunnelen	123
6.3.1	Innledning.....	123
6.3.2	Stortingsvedtak om bygging, prosjektomtale i Stortingsproposisjoner....	125
6.3.3	Planer, historikk og omfang	126
6.3.4	Trafikkgrunnlag og nyskapt trafikk.....	127
6.3.5	Nyttkostnadsanalyser.....	133
6.3.6	Resultater og konklusjon.....	140
	Konklusjoner.....	146
7	Referanser	148

Figuroversikt

Figur 1	Resultatmåling med vekt på intern og ekstern effektivitet	25
Figur 2	Behandling av ulike konsekvenskategorier i nytte-kostnadsanalyser av transportprosjekter (der slike analyser benyttes), fordelt på ulike land, kilde Odgaard et al., 2005	38
Figur 3	Verdifastsettelse av effekter	42
Figur 4	Sammenlikningsmatrise for benchmarkingpartnere	58
Figur 5	Radardiagram for benchmarkinganalyse (der beste prestasjon ligger lengst fra sentrum av diagrammet)	58
Figur 6	Datainnhyllingsanalyse (DEA)	59
Figur 7	Struktur for måling i Minnesota, kilde: http://www.dot.state.mn.us/dashboards/plan-n-prog1.html	63
Figur 8	Dashboard fra Virginia http://dashboard.virginiadot.org/	64
Figur 9	Effektutvikling Rikshospitalet (Rikshospitalet 2004, side 10)	73
Figur 10	Referansegrunnlag for ex post-evaluering	92
Figur 11	Utviklingen i antall reisende, antall avganger, punktlighet og reisetid. Situasjonen i 1987 er satt til indeks 100 for alle disse forholdene. Antall avganger er telt som ankomster til Oslo S fra Vestfoldbanen, med varierende utgangspunkt (Skien, Larvik, Sandefjord eller Tønsberg)	113
Figur 12	Utviklingen i antall reisende på Vestfoldbanen. Kilder: 1986-2004: På Sporet (1987-2005). Faktisk 1978-1988: Bruer (1989:12)	115
Figur 13	Utviklingen i antall reisende, punktlighet, og reisetid fra 1991 til 2004 på strekningen der Gardermobanen går. Situasjonen i 1991 er indeksert til 100	116
Figur 14	Utviklingen i antall reisende, punktlighet, og antall avganger fra 1987 til 2004 på Østfoldbanen. Utviklingen gjelder InterCity/mellomdistanse, med unntak av antall avganger, som er for alle typer tog. Situasjonen i 1987 er indeksert til 100	118
Figur 15	Utviklingen i antall reisende, punktlighet, og antall avganger fra 1987 til 2004 på Bergensbanen. Situasjonen i 1987 er indeksert til 100	121
Figur 16	Utviklingen i reisetid på Bergensbanens persontog 1987 – 2004, fra Oslo til Bergen	122
Figur 17	Oversiktskart	125
Figur 18	Trafikkutviklingen i Frøyatunnelen og på bru til Fjellværøy og i Hitratunnelen	128
Figur 19	Strekninger som er benyttet som referansestrekninger for trafikkutvikling	129
Figur 20	Trafikkutvikling på Frøyaforbindelsen, Hitratunnelen og to andre ferjeforbindelser i samme område (referanseår 1999 = 1,0)	130
Figur 21	Trafikkutviklingen 1999 – 2006 som skal legges inn i nytteberegningene. Figuren viser avrundede trafikk tall	132
Figur 22	Etterspørselskurve for personbiltrafikk (antall personturer). Skravert område er nytte av 140 nyskapte personturer, gitt en vekst fra 400 til 540 personturer pr. døgn	136

Tabelloversikt

Tabell 1	Ti trinn i utarbeiding av en vurdering av sosiale effekter (Social Impact Assessment – SIA)	46
Tabell 2	Tennessee-modellen for evaluering av prosjekter	67
Tabell 3	Vurderingsparametre for ulike typer prosjekter	98
Tabell 4	Vurdering av aktører for gjennomføring av ex post-evalueringer	99
Tabell 5	Utbygde nye parseller på Vestfoldbanen, med åpningsdato og kostnad ..	112
Tabell 6	Sammenligning mellom utgangspunktet for utbygging, utvalgte forventede effekter av utbyggingen og situasjonen i 2004 for Vestfoldbanen	114
Tabell 7	Utviklingen i kostnad for Gardermobanen (basert på NOU 1999:28, side 97-99)	116
Tabell 8	Nøkkeltall for Gardermobanen, før tiltaket, estimert og faktisk utfall	117
Tabell 9	Nøkkeltall for Østfoldbanen, før utbygging av dobbeltspor Ski-Sandbukta, estimert etter utbygging og faktisk utfall	119
Tabell 10 ..	Oversikt over de analyserte byggetrinnene bygde. Kostnadene viser til henholdsvis første og siste gang prosjektene er omtalt enten i statsbudsjett	120
Tabell 11 ...	Nøkkeltall for Bergensbanen, før utbygging av Finsetunellen og Finse-Gråskallen, estimert etter utbygging og faktisk utfall	120
Tabell 12 ..	Forskjell mellom faktisk utvikling og forventet situasjon, målt i prosent. Forskjellen = (faktisk – forventet) / forventet. * = Forventet antall er ikke oppgitt i den tilgjengelige dokumentasjonen	122
Tabell 13 ..	Forskjell mellom faktisk utvikling før og etter investeringene, målt i prosent. Forskjellen = (etter - før) / før	123
Tabell 14 ..	Trafikkutviklingen for tunge og lette kjøretøy og for nyskapt/ ikke-nyskapt trafikk	132
Tabell 15 ..	Trafikkgrunnlag for beregning ved bevilgning, dvs. trafikk på ferjestrekning og nyskapt trafikk fordelt på reiserelasjon og på tunge/lette kjøretøy. ÅDT 2001	135
Tabell 16 ..	Nyskapt trafikk for strømmene 1-4 og 2-4	140
Tabell 17 ..	Nøkkeldata for nyttekostnadsanalysene	141
Tabell 18 ..	Resultater fra nytteberegninger	142

Sammendrag

Denne rapporten beskriver eksisterende metoder og praksis for effektvurdering av offentlige prosjekter, med hovedvekt på metoder for evaluering i ettertid (ex post-effektvurderinger). Basert på gjennomgangen av eksisterende metoder og praksis presenteres et forslag til metodikk for ex post-evaluering av store offentlige investeringsprosjekter i Norge. Fokus er rettet mot prosjektenes samfunnsmessige effekter, dvs. det som betegnes som prosjektenes eksterne effektivitet. Dette i motsetning til en vurdering av hvorvidt resursene i selve prosjektet er brukt effektivt (intern effektivitet). Den anbefalte metoden illustreres med eksempler på mulig bruk i evaluering av ulike typer prosjekter.

Eksisterende teori om evaluering og effektvurdering

Det meste av litteraturen om metoder og praksis for effektvurderinger gjelder vurdering av ulike løsningsalternativers antatte effekter før det tas beslutning om hvilket alternativ som skal gjennomføres, dvs. ex ante-effektvurderinger i stedet for ex post. Det er imidlertid betydelig metodisk sammenfall mellom ex ante- og ex post-effektvurderinger, både når det gjelder relevante konsekvenskategorier, indikatorer for å måle effektene i forhold til disse kategoriene og mht. vurdering av hvor nyttige eller skadelige effektene er. I kapittel 2 gjennomgås fire hovedgrupper av metoder: Nytte-kostnadsanalyse m.m., samfunnseffekt-analyse, prestasjonsvurdering og benchmarking. De to første av disse kategoriene er utviklet for og benyttes primært i ex ante-vurderinger, men vil også kunne benyttes i forbindelse med ex post-etterprøving. Prestasjonsvurdering og benchmarking er metoder som i større grad er utviklet med tanke på evaluering i ettertid.

Nytte-kostnadsanalyse innebærer at nyttevirkinger og kostnader verdisettes i kroner så langt det er faglig holdbart. Dersom en legger sammen den beregnede verdien av alle effektene ved et tiltak og summen av disse er større enn kostnadene, sier en at tiltaket er samfunnsøkonomisk lønnsomt. Realøkonomiske virkninger omfatter både virkninger som er kvantifiserbare og prissatte, virkninger som er kvantifiserbare i fysiske størrelser, men som ikke lar seg verdisette i kroner, og ikke-kvantifiserbare virkninger. Det blir arbeidet kontinuerlig for å prissette virkninger som det hittil ikke har vært praksis for å verdisette, bl.a. gjennom undersøkelser av folks betalingsvillighet. Det teoretiske grunnlaget for denne type analyse er imidlertid blitt kritisert fra flere vinkler. Tidligere arbeider i Concept-programmet peker på at

samfunnsøkonomisk lønnsomhet gir en bedre oppsummering av et tiltaks konsekvenser jo mindre betydningsfulle fordelingsvirkningene er, jo flere konsekvenser vi finner det faglig forsvarlig å verdisette i kroner, og jo mindre innslag det er av etisk vanskelige spørsmål.

Siden særlig de sosiale, men også mange av de miljømessige effektene som typisk kan opptre som følge av store offentlige investeringsprosjekter, er vanskelige å fastslå med høy grad av presisjon, og enda vanskeligere å prissette, vil nytte-kostnadsanalyse av prosjektenes effekter i beste fall bare kunne fange opp en begrenset del av de relevante effektene. En utvidet analyse, som forsøker å inkludere slike effekter, er *samfunnseffekt-analyse*. Slike analyser omfatter både sosiale, miljømessige og økonomiske virkninger – der alle disse tre begrepene er bredt definert. Store offentlige investeringsprosjekter bidrar normalt til å skape nytte i forhold til de behovene og interessene som prosjektene er begrunnet med, men kan samtidig ha bieffekter som reduserer nytten i forhold til enkelte andre hensyn og interesser. En samfunnseffekt-analyse bør derfor belyse hvordan konsekvensene av et prosjekt slår ut for ulike befolkningsgrupper, og ikke bare effekten for samfunnet sett under ett. Miljøkonsekvensanalyse (Environmental Impact Assessment) og analyse av sosiale konsekvenser (Social Impact Assessment) er viktige delmetoder innenfor samfunnseffektanalyser. Økonomiske beregninger vil innenfor denne tilnærmingen fortrinnsvis bli utført som kostnads-virkningsanalyser, siden en stor del av de relevante effektene ikke lar seg måle og verdisette på en måte som oppfyller nytte-kostnadsanalysens krav til datainput.

Prestasjonsvurderinger ble tidligere utført med utgangspunkt i et nokså snevert produktivitetsbegrep. Etter hvert er det vokst fram en erkjennelse av at dette hadde mange uheldige effekter. Et eksempel på dette er innkjøpsavdelinger som ukritisk anskaffer deler fra useriøse leverandører fordi utvikling i innkjøpskostnader er et fokusområde i målesystemet. Mange forfattere har tatt et kraftig oppgjør med denne kulturen. Hovedbudskapet deres er at prestasjon må måles ved bruk av langt flere dimensjoner. Konseptet balansert målstyring er et uttrykk for dette. Ved å måle flere viktige dimensjoner av prestasjon, styres atferden i virksomhetene til også å fokusere på disse. Det såkalte DPSIR-rammeverket reflekterer også denne tankegangen. DPSIR står for *Driving forces, Pressures, States, Impacts* og *Responses*. Tiltak og prosjekter bør evalueres ut fra disse fem forholdene, og vil da vise sammenhenger og hvilke drivkrefter som bidrar til å skape fremtidige effekter. Denne kunnskapen kan også utnyttes i ex post-evalueringer av prosjekter. Er det kjent at forhold som samfunnskontakt, tilfredshet i nærmiljøet med prosjektet, industriell utvikling som et resultat av prosjektet, miljøpåvirkning (positiv eller negativ), osv. vil vurderes, vil det

uovergelig sikre at prosjektene også i etablering, planlegging og gjennomføring er mer bevisste disse forholdene.

Substantivet *benchmark* defineres som ”et på forhånd definert punkt, brukt som en referanse for målinger”. I forbindelse med prestasjonsvurderinger har benchmarking en viktig funksjon i å gi målinger et referansepunkt. En ex post-vurdering av ett enkelt prosjekt vil kunne avdekke tallverdier for forholdet mellom kostnad og nytte, trafikkvolum, brukertilfredshet, osv. Isolert sett gir selvsagt disse en viss innsikt i prosjektets prestasjoner. Ytterligere informasjon får en ved å sammenholde disse med de opprinnelige estimater, dermed kan man vurdere om estimatene var realistiske og om prosjektet evnet å nå disse. Virkelig interessant blir evalueringen om man kan sammenlikne resultatene med en serie andre prosjekter. Den såkalte referanseprognosemetoden er en form for strukturert kvantitativ benchmarking av forutsetningene for et prosjekt. Fremgangsmåten innebærer at man sammenligner det aktuelle prosjektet med en gruppe av tilsvarende prosjekter og utnytter foreliggende informasjon om disse. Felles for både benchmarking og referanseprognosemetoden er kravet om en database av et visst omfang og en grad av standardisering av evalueringskriterier/måleparametre, f. eks. mht. teknisk funksjonalitet og effektivitet, økonomisk effektivitet, sosial aksept, politisk aksept, miljømessig aksept og næringsmessig utvikling.

Eksisterende praksis – internasjonale og norske eksempler

Søking på internett, i biblioteker og i forskningsdatabaser viser at det alt overveiende av litteraturen om effektvurderinger av store statlige investeringsprosjekter dreier seg om ex ante-vurderinger. I kapittel 3 vises eksempler på eksisterende praksis som vi mener kan være relevante også med tanke på ex post-evalueringer. Eksemplene omfatter systemer for prestasjonsmåling i transportetater i amerikanske delstater, oppfølging av transportpolitiske mål i Sverige, systemer for vurdering av løsningskonsepter for kollektivtransport benyttet av organer på regionalt eller delstatsnivå i USA, karakterboken i den danske byggebransjen, og et metodeverk for evaluering og benchmarking av prosjekter utviklet av det private firmaet IPA.

Vi har ikke lyktes i å finne noen systematisk gjennomgang av eller samlet oversikt over tidligere utførte ex post-evalueringer av store statlige prosjekter i Norge. Derimot finnes det en rekke dokumenter av ulike slag som tar for seg ulike sider ved gjennomførte statlige investeringsprosjekter. Kapittel 4 presenterer hovedbudskapene fra et utvalg dokumenter vi finner relevante. Oversikten er ikke komplett. Et fellestrekk er at dokumentene er utarbeidet med ulike formål og derfor har ulike fokus og tar opp ulike tema. Evalueringer av følgende prosjekter omtales:

Prosjekt Nytt Rikshospital, IT-prosjekter i Trygdeetaten, Troll Olje-prosjektet, materiellinvesteringer i Forsvaret, rehabilitering av Det Kongelige Slott, regjeringens representasjonsbolig og Stiftsgården, en undersøkelse av planlegging og oppfølging av store veganlegg i Statens vegvesen, samferdselsmyndighetenes forslag om etterprøving av nytte-kostnadsanalyser av samferdselsprosjekter, utbyggingen av tre oljefelt (Åsgard, Visund og Jotun) på kontinentalsokkelen, tetting av Romeriksporten, anskaffelse av F-16 kampfly, IT-satsingen Winix i norsk skole, IT-prosjektet Tress 90, effektvurdering av alternative utbyggingsprosjekter for Nasjonalmuseet for kunst, arkitektur og design, og Samlet Plan for Vassdrag.

En generell observasjon fra gjennomgangen av tidligere større evalueringer, er at de sier relativt lite om de forholdene vi er opptatt av i denne utredningen: Evalueringene belyser i liten grad om riktig konsept ble valgt, og heller ikke eventuell bredde i alternative konsepter. De fleste evalueringene belyser i relativt liten grad nullalternativet, og i den grad de gjør det, gjøres det ikke på noen systematisk eller standardisert måte. Videre gjøres de fleste evalueringene med referanse til ex ante-beregninger for kostnader og eventuelt inntekter for prosjektet, men i liten grad med referanse til ex ante-beregninger av samfunnsmessig kostnad og nytte. Evalueringene handler dessuten i begrenset grad om samfunnsmessige effekter i et videre perspektiv, fordelingsvirkninger, osv.

Forslag til ny metode

Kapittel 5 beskriver vårt forslag til en metode for ex post-evaluering av store, statlige prosjekter. Vi diskuterer først hensikten med slik evaluering, etterfulgt av hva som vil være sammenlikningsgrunnlaget for evalueringen, når evalueringen bør finne sted og detaljer i selve metoden for evalueringen. Videre diskuteres hvilket datagrunnlag som kreves for å gjennomføre en slik evaluering, hvilke roller som må fylles i en ordning for systematiske evalueringer, og hvilke konsekvenser en slik ordning vil medføre for ulike aktører.

Hensikt

Hensikten med ex post-evalueringer er for det første å vurdere om de opprinnelige målsettingene for prosjektene faktisk blir oppnådd. I tillegg vil ex post-evaluering kunne gi viktige bidrag til å forebedre ex ante-analysene, til å øke bevisstheten om at ex ante-beregningene som investeringen bygger på må være realistiske eller ”edruelige”, og til å standardisere analyser og beslutningsunderlag ex ante. Gjennom ex post-evalueringer vil en også kunne bygge opp en erfaringsdatabase om effekter av store investeringer, som vil kunne være nyttig når nye prosjekter skal utformes og planlegges.

Det man ønsker å sette søkelyset på er altså evnen til å velge riktige prosjekter og konsepter, kvaliteten på analyser og beslutningsprosesser i tidligfasen, osv samt de oppnådde effektene av prosjektet. Sentrale spørsmål vil være: Var ex ante-analysene anlagt bredt nok? Ga prosjektet de planlagte effekter? Dette innebærer at evalueringen ikke bare bør være en produktevaluering, men også omfatte elementer av en prosessevaluering. Noen sentrale spørsmål som bør reises i forbindelse med planleggings- og beslutningsprosessen av store prosjekter kan være: Hvilke føringer ble satt for ex ante-analysene og konseptvalg? Har det skjedd utvikling/endringer i nytte-/kostnadsestimater utover i prosessen eller i design/utforming av konsept? Var ulike beregninger (nytte, kostnad, osv.) og diskonteringsfaktorer for realisert og alternative konsepter rimelige? Ble alternative konsepter utredet seriøst?

Referansepunkt og sammenlikningsgrunnlag

Når effektene oppnådd med prosjektet skal vurderes, er et sentralt spørsmål hva det sammenliknes mot. For å evaluere de reelle effektene av prosjektet, må situasjonen etter at prosjektet er gjennomført, sammenliknes med situasjonen man ville ha hatt på det samme tidspunktet uten at prosjektet var blitt gjennomført. En slik kontrafaktisk analyse er komplisert, og man kan derfor ikke kreve for høy presisjon. Også for spørsmålet om riktig konsept ble valgt, vil det relevante være å gjennomføre en kontrafaktisk analyse – i dette tilfellet opp mot hva situasjonen ville blitt om man hadde valgt ett av de andre alternative konsepter. Videre kan man sammenholde oppnådde effekter med den estimerte nytten slik den ble fremsatt før prosjektet ble vedtatt og igangsatt. Dette vil være en viktig indikator for kvaliteten på planleggingsarbeidet. Endelig kan man sammenlikne oppnådde effekter i prosjektet som evalueres med hva andre, liknende prosjekter har oppnådd. Dette gir en ekstern referanse gjennom en form for benchmarking.

En ren sammenlikning av tilstanden etter gjennomføring av prosjektet med før-situasjonen er selvfølgelig enklere å utføre enn de kontrafaktiske analysene som skal til for å avdekke prosjektets effekter. Slike enkle sammenstillinger kan gi erfaringsdata om typiske endringer i forbindelse med den aktuelle typen prosjekter, og kan dermed gi input til prognosearbeid i framtidige prosjekter. Slike enkle før-etter-sammenlikninger gir imidlertid ikke i seg selv holdepunkt for å si noe om *effektene* av et prosjekt.

Tidspunkt for ex-post-evaluering

Det går ofte en viss tid fra ferdigstillelsen av et prosjekt til nytteeffektene viser seg. Ofte fører oppstartproblemer til en *nedgang* i opplevd nytte like etter at ny infrastruktur er tatt i bruk. Hvis prosjektet evalueres for tidlig (innen ett år etter

avslutning), vil man kunne risikere å evaluere prosjektet mens det ennå ikke har realisert nytten. Hvis prosjektet på den annen side evalueres for sent (mer enn 4-5 år etter avslutning), vil det være vanskelig å påvise hvilke endringer som skyldes prosjektet og hvilke som skyldes andre trekk ved samfunnsutviklingen.

Prosjektorganisasjonen kan også på dette tidspunktet være så spredd at det kan være vanskelig å innhente data og informasjon. Tidspunktet for ex post-evalueringen vil nok derfor måtte variere fra prosjekt til prosjekt, men dersom det lar seg gjøre anbefales to år etter at prosjektet er fullført som et standardisert tidspunkt. Man har da lagt bak seg eventuelle garantiperioder og oppstartsproblemer, samtidig som effektene i mindre grad vil være forsterket eller motvirket av annen samfunnsutvikling eller nye prosjekter.

Mer spesifikt om datagrunnlag, vurderingsparametre og relevante momenter

Evalueringsmetodikken bør tilstrebe å fange opp så brede konsekvenskategorier at den kan dekke sterkt ulike alternative konsepter. Det er også en selvfølge at både positive og negative konsekvenser vurderes. Og som ellers i evalueringer vil denne typen ex post-vurderinger måtte inkludere fordelingskjevheter prosjektet måtte føre til. Til en viss grad vil ex post-evalueringen bestå i å repetere ex ante-analysen med reelle tall, for derigjennom å vurdere arbeidet som ble gjort i prosjektets tidlige fase. I prosjekter der konsepter endres underveis, bør evalueringen vurdere det realiserede konseptet, men ved sammenlikning av konseptalternativer må også det opprinnelige konseptet vurderes.

Det bør bygges opp en database/et bibliotek der erfaringene fra evalueringene samles. På flest mulig av evalueringstidspunktene bør data standardiseres og legges til rette for framtidig benchmarking, forskning og analyse. Data bør forvaltes slik at de kan være tilgjengelige for aktuelle framtidige gjennomførere av KS1, KS2 og ex post-evalueringer, for eksempel i regi av Senter for Statlig Økonomistyring (SSØ) eller forskningsprosjektet Concept.

For at ex post-evalueringene skal bli gode, forutsettes det et godt datagrunnlag, herunder nullpunktsmålinger, data om nullalternativet og alternative konsepter. Dette stiller krav til beslutningsunderlaget og tilhørende analyser. Det vil være nødvendig at det i KS1 og KS2 sikres et godt datagrunnlag for framtidig ex post-evaluering.

Vi har utviklet et forslag til vurderingsparametre for ulike typer prosjekter, gruppert innenfor tre standard kategorier: Investeringskostnader, driftskostnader og konsekvenser av prosjektet. Den sistnevnte kategorien omfatter bl.a. bruk/produksjonsdata for kjernevirksomheten i anlegget, direkte nytte for brukere,

sikkerhet og pålitelighet, miljøforhold, sosiale virkninger og regionale virkninger. Videre har vi inndelt prosjektene i fire sektorbaserte hovedkategorier: Samferdselsprosjekter, byggeprosjekter, forsvarsprosjekter og øvrige prosjekter (hovedsakelig IT-prosjekter). I kapittel 5 er det tatt med en tabell som viser en rekke relevante vurderingsparametre innenfor hver av de sektorbaserte hovedkategoriene.

Ansvar for og roller i ex post-evaluering

Det aktuelle fagdepartementet, i samarbeid med Finansdepartementet, bør ha hovedansvaret for at ex post-evalueringen gjennomføres. Det bør enten tas høyde i prosjektets budsjett for kostnadene ved evalueringen eller det må opprettes en budsjettpost for ex post-evalueringer for de tilfeller der evalueringene vil komme etter at prosjektet er fullført og ”lukket”.

Eksempler på gjennomføring av ex post-evaluering med ny metode

Kapittel 6 viser noen konkrete eksempler på hvordan ex post-evalueringer kan gjennomføres i tråd med den metodikken vi foreslår. Eksempelene omfatter prosjektet for bygging av nytt rikshospital, fire jernbaneprosjekter og et vegtunnelprosjekt. Rikshospital-eksemplet viser hvordan en bred, prosjektspesifikk evaluering kan gjennomføres. Jernbaneprosjektene og vegtunnelen er eksempler på hvordan en løpende oppfølging av mer konkrete parametere kan gjøres. I tillegg viser evalueringen av vegprosjektet hvordan en bredere nytte-kostnadsanalyse kan gjennomføres.

Konklusjoner

Rapportens hovedkonklusjoner er kort oppsummert i kapittel 7. Prosjektgruppen forslår at det etableres en ordning med ex post-evaluering av alle store, offentlige prosjekter, basert på følgende hovedprinsipper:

1. Ex post-evalueringen skal ha som formål å kartlegge effekter av det gjennomførte prosjektet sammenlignet med en situasjon der prosjektet ikke ble gjennomført (null-alternativet), og sammenlignet med effekter av alternative men ikke valgte konsepter som ble utviklet og vurdert i KS1-evalueringen.
Videre bør evalueringene inneholde:
2. En vurdering av gjennomføringen av prosjektet med hensyn til effektivitet, kvalitet og resultater. Gjennomføringen bør vurderes mot et utvalg relevante referanseprosjekter.

3. En vurdering av i hvilken grad nytten og kostnadene knyttet til det realiserte prosjektet faller sammen med de ex ante-beregninger som lå til grunn for beslutningen om gjennomføringen av prosjektet, samt forklaring av eventuelle avvik
4. En analyse av prosjektets samfunnsmessige effekter i et bredt perspektiv
5. En vurdering av om riktig konsept ble valgt, herunder om et annet konsept burde ha blitt valgt i lys av den nye kunnskap man har ervervet, både om det gjennomførte prosjektet og om den utviklingen som har skjedd i de samfunnsmessige og teknologiske rammene rundt prosjektet.
6. En samlet oppsummering av erfaringer fra prosjektet, herunder en formidlings-/læringsstrategi.

Summary

This report reviews existing impact assessment methods and practices for governmental projects, with an emphasis on methods for after-the fact evaluation (ex post impact assessment). Based on the review of existing methods and practices, methods for ex post evaluation of large-scale Norwegian governmental investment projects are proposed. The focus of the report is directed toward impacts of the projects at a societal level, i.e., what has been termed the external effectiveness of the projects. This is distinct from an assessment of the efficiency of the resource use within the project itself (internal effectiveness). The recommended method is illustrated by several examples of its possible use in the assessment of different types of projects.

Existing theories on project evaluation and impact assessment

Most of the literature on methods and practice for impact assessment applies to appraisals of the assumed impacts of alternative solutions before any decision is made regarding which alternative to implement, i.e. ex ante impact assessment instead of ex post. There is, however, considerable methodical overlap between ex ante and ex post impact assessment, regarding the relevant consequence categories, indicators for measuring effects against these categories, as well as the assessment of how beneficial or undesirable the respective effects are. In Chapter 2, four main groups of methods are reviewed: Cost-benefit analysis and related methods, broad analysis of society-level impacts, performance measuring, and benchmarking. The two first of these categories have been developed for and are primarily applied to ex ante assessment, but may also be used as tools for ex post evaluation. Performance measuring and benchmarking are methods having to a higher extent been developed for the purpose of ex post evaluation.

Cost-benefit analysis implies that benefits and costs are being monetarized to as great extent as can be defended scientifically. The initiative is considered to be profitable for society if the sum of all calculated benefits resulting from an initiative is higher than the calculated costs. Real economic impacts include effects that can readily be monetarized, physically quantifiable impacts that are difficult or impossible to monetarize, as well as non-quantifiable effects. Efforts are continually being made in order to enable monetarization of impact categories for which monetary values have hitherto not been assigned, notably by means of willingness-to-pay investigations.

However, the theoretical foundation of this type of analyses has been criticized from several camps. According to earlier reports from the Concept research program, the ability of cost-benefit analyses to summarize the impacts of an initiative in a meaningful way is higher, the lower is the importance of any distributional effects, the larger is the number of consequences that may be monetarized in a scientifically defensible way, and the lower is the extent to which the decision involves ethically difficult issues.

Since especially the social, but also many of the environmental impacts typically occurring as a result of large governmental works projects are difficult to assess with a high degree of precision, and even more difficult to value in economic terms, cost-benefit analyses can at best only comprise a limited part of the relevant impacts. A broader scope, allowing the inclusion of impacts that cannot be monetarized in a defensible way, is possible when conducting what we have called *broad analyses of society-level impacts*. Such analyses include social, environmental as well as economic impacts, all of which broadly defined. Large-scale governmental investment projects normally contribute to create benefits in relation to the needs and interests from which the projects have been justified, but may at the same time result in side-effects affecting other concerns and interests in a negative way. A broad analysis of society-level impacts should therefore elucidate how impacts of a project affect different population groups, and not only impacts at an aggregate social level. Environmental Impact Assessment and Social Impact Assessment are important partial methods within analyses of society-level impacts. Within this approach, the economic assessment will typically be carried out as a cost-impact analysis, as a large number of the relevant effects do not lend themselves to quantification and valorizing in a way meeting the requirements of a cost-benefit analysis.

Performance measurement was previously often based on a quite narrow concept of productivity. Gradually, awareness has grown that such analyses have numerous unfavorable implications. One example of such perverted effects is the purchase of commodities from unreliable suppliers in response to a strict focus of the performance measurement system on purchasing costs. Many authors have put forth serious criticism against this culture. Their main message is that performance must be measured against a far broader range of dimensions. The concept of balanced management by objectives is an example of this concern. By measuring a broader range of performance dimensions, the behavior of the agencies is directed toward a focus on these dimensions. The so-called DPSIR framework for environmental indicators reflects a similar way of thinking. DPSIR denotes *Driving forces, Pressures, States, Impacts and Responses*. Initiatives and projects should be evaluated against all these five dimensions. The analysis will then throw light on

relationships and driving forces contributing to produce future impacts. The above-mentioned way of reasoning can also be drawn on in ex post evaluation of projects. If project participants have been made aware in advance that communication with society, neighborhood satisfaction with the project, industrial development resulting from the project, environmental impacts, etc. will be subject to evaluation, they will probably direct more of their attention to such effects when establishing, planning and implementing the project.

The noun *benchmark* has been defined as ‘a predefined point, used as a reference for measurement’. In the context of performance measurement, benchmarking has as its main function to provide measurements with a point of reference. An ex post evaluation of a single project may elicit figures for costs-benefit ratios, amounts of traffic, user satisfaction, etc. Seen in isolation, this provides a certain insight into the performance of the project. Further information may be obtained by comparing this information with the original (ex ante) estimates, thereby enabling an evaluation of the realism of these estimates and the ability of the project to reach its goals. A comparison of the results with the evaluations of a number of similar projects may be even more interesting. The so-called reference class prognosis method is an example of a structured, quantitative benchmarking of the conditions of a project. The method implies that the project in question is compared to a group of similar projects, utilizing available information about the latter. Both benchmarking and the reference class prognosis method require a database of a certain amount and a certain degree of standardized evaluation criteria/measurement parameters, e.g. in terms of technical functionality and efficiency, economic effectiveness, social, political and environmental acceptance, and industrial development.

Existing practice – international and Norwegian examples

Searches on the Internet, in libraries and in research databases show that the vast bulk of literature on impact assessment of large governmental investment projects deals with ex ante analyses. In Chapter 3, some examples of existing practice are shown, which we believe will be relevant to ex post evaluation as well as to their original ex ante purpose. The examples include state-level systems for performance measurement in the USA, Swedish routines for monitoring the compliance of implemented transport policy measures with transport policy goals, systems used by state or regional authorities in USA for the evaluation of alternative public transport concepts, the so-called mark book of the Danish construction trade, and a framework for project evaluation and benchmarking developed by the consulting company IPA.

We have not succeeded in finding any systematic review or overview of previously conducted ex post evaluations of large governmental projects in Norway. However, a number of documents exist, where particular aspects of implemented governmental investment projects are addressed. Chapter 4 presents the main messages from a number of the most relevant among these documents. The overview does not claim to be complete. A common feature is that the documents have been produced for different purposes, thus varying in their focus and addressing different topics. Evaluations of the following projects are briefly presented: The project for a New National Hospital; ICT projects in the Social Welfare Agency; The Troll Oil project; defense equipment investments; renovation of the Royal Castle; the governmental representational residence and an old wooden representational building in Trondheim; the Highway Directorate's investigation of the planning and monitoring of large road investments; a proposal for ex-post evaluation of cost-benefit analyses of transportation investment projects; the development of three offshore oil fields (Åsgard, Visund and Jotun); the repair of leakages in a railway tunnel; the purchase of F-16 combat airplanes; the Winix ICT project in the public educational system; the ICT project Tress 90; impact assessments of alternative solutions for the National museum for art, architecture and design; and the Comprehensive Plan for Watercourses.

A general observation from the review of the above-mentioned projects is that they do only to a limited extent address the topics focused on in the present report. The evaluations are only to a little extent able to throw light on whether or not the right concept was chosen, and neither on the possible range of alternative concepts. Most of the evaluations hardly elucidate the zero alternative, and to the extent that this is done, the analysis is not conducted in any systematic or standardized way. Most of the evaluations refer to ex ante forecasts of direct costs and any revenues from the projects, but they seldom make reference to ex ante calculations of social costs and benefits. The evaluations address wider societal impacts, distributional effects etc. only to a limited extent.

Proposal for a new method

Chapter 5 describes our proposal for a method for ex post evaluation of large governmental projects. First, we discuss the purpose of such evaluations, followed by considerations about the appropriate contexts against which the project should be compared, the appropriate time of the evaluations, and more specific details of the evaluation method. Moreover, the data required in order to carry out the recommended evaluation is discussed, along with the roles to be filled in a system of

institutionalized ex post project evaluations, and the consequences of such an institution for different agents.

Purpose

The purpose of ex post evaluations is firstly to assess whether or not the original objectives of the projects were actually achieved. In addition, ex post evaluation may provide valuable contributions to the knowledge base for future ex ante analyses, increase the motivation among those who carry out the ex ante analyses for making realistic and sober assessments, and contribute to a standardizing of analyses and other information making up the base for decisions about project implementation. Through ex post evaluations it will also be possible to build a database about impacts of large governmental investments. Such a database may be useful when designing and planning future projects.

Thus, what the proposed ex post analyses aim to draw the attention toward is the ability to choose the right projects and concepts, the quality of analyses and decision-making processes in the early project phase, and the effects obtained through the projects. Key questions are: Was the scope of the ex ante analyses sufficiently broad? Did the project result in the planned effects? This implies that the evaluation should not be confined to a product evaluation, but should also include elements of a process evaluation.

Key questions to be raised concerning the planning and decision-making processes in connection with large projects include: Which terms were set for the ex ante analyses and the choice of concept? Have changes or developments occurred in the cost-benefit estimates during the course of the process, or in the design of the concepts? Were the various calculations (costs, benefits, etc.) and discounting rates for the implemented and alternative concepts reasonable? Were alternative concepts made subject to serious elucidation?

Points of reference and bases for comparison

When assessing the effects obtained through a project, a key question is the appropriate situation against which the situation after implementation should be compared. In order to evaluate the real effects of the project, the situation after implementation must be compared to the situation that would have occurred at the same time in the absence of the project. Such a counterfactual analysis is complicated, and it is therefore not appropriate to require a too high degree of precision in the assessments. Also the question of whether the right concept was chosen requires a counterfactual analysis – in this case with a comparison with the

hypothetical situations resulting if each of the alternative concepts had been chosen. Moreover, the obtained effects can be compared to the benefits and negative impacts predicted before the decision and commencing of the project. This will be an important indicator of the quality of the planning work. Finally, the impacts of the evaluated project can be compared to what other, similar projects have brought about. This provides an external reference through a form of benchmarking.

Needless to say, a mere comparison of the situation after the implementation of the project to the pre-project situation is simpler to conduct than the counterfactual analyses necessary to uncover the impacts of the project. Such simple juxtapositions can provide experiential data about changes typically associated with the relevant type of project, hence making up a source for inputs to forecasts in connection with future projects. However, such simple before-and-after comparisons are insufficient as a base for conclusions about the *effects* of a project.

Scheduling of the ex post evaluation

Often, some time has to pass after implementation before the positive effects of a project manifest themselves. Actually, start-up problems often result in a decline in perceived utility immediately after the opening of new infrastructure. If the project is being evaluated too early (e.g. within one year after completion), there is a risk that the evaluation takes place at a time when the project has not yet realized its benefits. On the other hand, if the evaluation takes place too late (e.g. more than 4-5 years after completion), it may be difficult to identify which changes are due to the project and which are results of exogenous developmental trends. The project organization may also be scattered at such a late time, thus rendering the collection of data and information difficult. Thus, the timing of the ex post evaluation will have to vary from project to project, but if possible, a recommendation is the end of the second year of operation as a standard scheduling. By then, any warranting periods and startup problems have usually been passed, while the effects have usually not yet been significantly amplified or counteracted by new projects or exogenous developmental trends.

More specifically about data sources, evaluation parameters and relevant aspects

The evaluation methods should aim at covering a sufficiently wide range of consequence categories to be applicable to essentially different concept solutions. Needless to say, both positive and negative impacts must be assessed, as must any distributional biases resulting from the projects. To some extent, the ex post evaluation will consist of a repetition of the ex ante analysis with real-life figures,

thus enabling an assessment of the accuracy of the work carried out during the early stage of the project. In projects where concepts are being changed during the process, the evaluation should assess the realized concept, but when comparing concept alternatives, the original concept should also be included in the comparison.

In order to accumulate experience from several project evaluations, a database or library of such information should be established. Data should be standardized at as high a number of the evaluation points as possible and be made available for future benchmarking, research and analysis. The data should be managed in such a way that they can be accessible for future quality assurance teams evaluating the ex ante analyses leading up to choices of concepts (KS1) and project designs (KS2), as well as those who carry out future ex post evaluations. The responsibility for this might be located to the Center for Governmental Economy Management (SSØ) or the Concept research project.

High-quality ex post evaluation requires a good information base, including zero-point measurements, data about the zero alternative and information about alternative concepts. This places demands on the amount and quality of data and related analyses provided as a base for decision-making. It is therefore crucial that the above-mentioned KS1 and KS2 assessments ensure a good information base for future ex post evaluation.

We have developed a proposal for evaluation parameters for different types of projects, grouped into three standard categories: Investment costs, operational costs, and consequences of the projects. The latter category includes, among other things, data for the use/production making up the core function of the installation, direct utility for users, safety and reliability, environmental conditions, social impacts and regional impacts. Moreover projects have subdivided the projects into four thematic main categories: Transportation projects, building projects, defense projects, and other projects (mainly ICT projects). Chapter 5 includes a table showing relevant evaluation parameters within each of these thematic categories.

Responsibility for and roles in ex post evaluation

The relevant ministry, in cooperation with the Ministry of Finance, should carry the main responsibility for initiating ex post evaluations of relevant projects. Funding covering the costs of evaluation should either be provided as a part of the budget of the project subject to evaluation, or alternatively be granted from a particular item in the National Budget. The latter is necessary to ensure funding possibilities in cases where the project organization has been dismantled before the initiation of the ex-post evaluation.

Examples of ex post evaluations according to the proposed method

Chapter 6 shows a few illustrative examples showing how ex post evaluations following our proposed method may be carried out. The examples include the construction of a new National Hospital, four railroad projects, and a road tunnel. The National Hospital case shows types of impacts and considerations relevant to include in a broad, project-specific evaluation. The railroad projects and the road tunnel exemplify how a continual monitoring of more specific parameters may be carried out. In addition, the evaluation of the road project shows how a broader cost benefit analysis can be carried out.

Conclusions

The main conclusions of the report are briefly summarized in Chapter 7. The project group proposes that a scheme of mandatory ex post evaluations of large governmental projects be established, based on the following main principles:

1. The purpose of the ex post evaluation is to uncover impacts of the implemented project, based on comparisons with the situation had the project not been implemented (the zero alternative), and with the effects of alternative, but not chosen concepts developed and assessed in connection with the KS1 evaluation.

Moreover, the evaluations should include:

2. An assessment of the project implementation in terms of effectiveness, quality, and results. The implementation should be evaluated against a selection of relevant reference projects
3. An assessment of the extent to which the benefits and costs associated with the realized project are consistent with the ex ante forecasts on which the decision to implement the project was based, and an explanation of any deviations
4. An analysis of society-level impacts, assessed from a broad perspective
5. An appraisal of whether or not the right concept was chosen. This appraisal should include an assessment of whether or not a different project should have been chosen, judged in the light of new knowledge obtained about the implemented project as well as the development of social and technological conditions relevant to the project
6. A comprehensive summarizing of experience from the project, including a strategy for communication and learning.

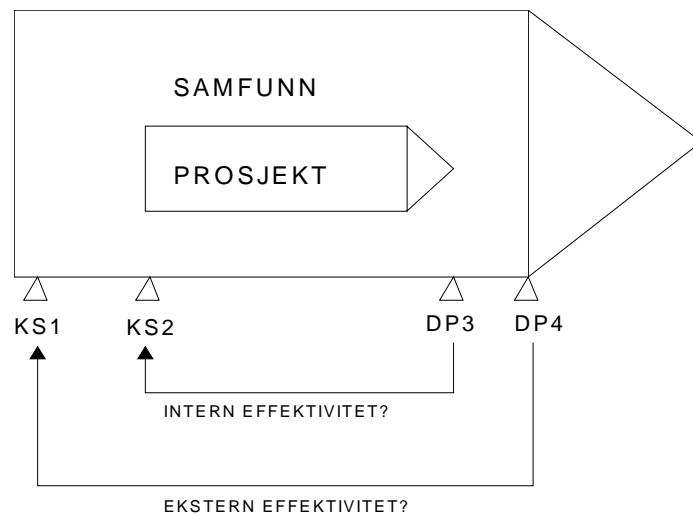
1 Innledning

1.1 Bakgrunn

Dette prosjektet har som mål å utvikle metoder for å foreta ex post-vurderinger av offentlige investeringsprosjekter. En viktig forutsetning for dette er innsikt i hvilke investeringer det offentlige foretar. Staten investerer i en rekke andre ting. Noen eksempler på områder der staten foretar store investeringer er:

- Samferdselsinfrastruktur (veier, broer, jernbanespor, flyplasser, osv.).
- Bygningsinvesteringer (sykehus, kulturhus, undervisningsbygg, fengsler, kontorer, osv.).
- Energiinfrastrukturinvesteringer (rørledninger, kraftlinjer, oljefeltinstallasjoner, osv.).
- Investeringer i forsvarsmateriell og – installasjoner (skip, helikoptre, fly, våpensystemer, osv.).
- IT-investeringer (programvare, maskinvare, større systemer, osv.).
- Investeringer i industriprosjekter (for eksempel Mongstad-raffineriet).

Statlige investeringer består altså av en svært heterogen gruppe prosjekter, som har svært ulike egenskaper og skal tjene ulike formål. Denne rapporten beskriver i korte trekk en metodikk for ex post-evalueringer av større offentlige prosjekter. I forhold til KS-regimet kommer denne evalueringen inn etter at prosjektet er slutført, der man skal evaluere virkninger opp mot det som ex ante-analysene ga som samfunnsmessige virkninger av prosjektet. Dette kan framstilles slik:



Figur 1 Resultatmåling med vekt på intern og ekstern effektivitet

Vi ser at KS1 (konseptvalg) og KS2 (prosjektutforming) gjøres før prosjektet eventuelt blir implementert. DP3 gjøres for å evaluere kvaliteten på selve gjennomføringen av prosjektet i henhold til kostnadsestimatene og prosjektets planlagte utforming. Vi evaluerer da med tanke på prosjektets interne effektivitet, nemlig om ressursene i selve prosjektet er brukt effektivt. DP3 henvender seg derfor primært til den ex ante-vurderingen som ble gjennomført i KS2.

DP4 er det evalueringspunktet som denne rapporten skal omfatte. Dette punkt skal i en viss tid etter prosjektets åpning måle hvilke effekter man fikk, målt opp mot det som vil være den kontrafaktiske situasjonen, nemlig situasjonen som ville ha oppstått for de måleparameterne som vi velger å måle opp mot, uten at prosjektet ble gjennomført. Da er målsettingene knyttet til prosjektets samfunnsmessige effekter. Virkningene knyttet til effektiviteten av selve gjennomføringen er behandlet i DP3. Det er dermed tiltakets eksterne effektivitet vi er ute etter å måle, nemlig hvordan prosjektet møter de samfunns mål og effektmål som ble formulert i KS1.

Definisjonsmessig betyr ex post enkelt sagt ”i etterkant”, engelsk ”after the fact” (Wikipedia, 2007), i motsetning til ex ante; ”i forkant”, engelsk ”beforehand”. I prosjektsammenheng mener vi med en ex post-evaluering dermed en vurdering av prosjektet i etterkant av at det er gjennomført, men ex ante-vurderinger er de antakelser og analyseres som gjøres før prosjektet startes. Hensiktene er dermed nokså ulike; en ex ante-analyse har som mål å klargjøre forutsetningene for prosjektet, definere dets målsetninger og forventede effekter, bidra til planleggingen av det, osv. I etterkant gjøres en ex post-evaluering primært for å vurdere hvilke

effekter prosjektet oppnådde, om disse er i tråd med de definerte målsetningene, om kvaliteten i gjennomføringen var som forventet, osv.

Det foreligger imidlertid ikke mye litteratur omkring eksisterende praksis og metoder for effektvurdering av store offentlige investeringsprosjekter etter at prosjektene er gjennomført. Det meste av det som er skrevet om metoder og praksis for effektvurderinger, gjelder vurderinger av ulike løsningsalternativers antatte effekter før det tas beslutning om hvilket alternativ som skal gjennomføres, dvs. ex ante-effektvurderinger i stedet for ex post. I den grad det er foretatt effektvurderinger i ettertid, har dette hittil oftest skjedd på ad hoc-basis.

Det betyr ikke nødvendigvis at metodene som kan anvendes henholdsvis ex ante og ex post er så ulike som disse forskjellene kan indikere. Som vi skal vise kan ofte de samme vurderingsformer benyttes, ex ante for å forutse effekter og nytte som kan oppnås, ex post for å verifisere om de inntraff. Til en viss grad vil anvendelsen av metodene fremstå noe ulikt, og i tillegg finnes det andre teknikker som bare er egnet for ex post-anvendelser. I dette kapittelet vil vi gi en oversikt over den eksisterende kunnskap som finnes på dette feltet, slik den er dokumentert i ulik litteratur.

1.2 Målsetninger

Delprosjektet omhandler ex post-effektvurdering av prosjekter som faller inn under Finansdepartementets kvalitetssikringsordning. Hovedhensikten er å legge grunnlag for å evaluere effekten av offentlige investeringer, som kan bidra til bedre tidligfasevurderinger og beslutninger i tiden som kommer.

Rapporten skal beskrive og vurdere forskning, kunnskap og praksis internasjonalt og i Norge vedrørende effektvurdering av store investeringsprosjekter og gi en anbefaling om hvordan dette bør gjennomføres for store statlige prosjekter i Norge. Anbefalingen skal ha format av en overgripende modell eller rammeverk for effektvurdering som kan anvendes på tvers av sektorer, med eksempler på hvordan dette kan operasjonaliseres på prosjekter innenfor enkelte sektorer, og hvilket datatilfang som vil kreves. Det skal søkes å oppnå en rimelig balanse mellom relevante målinger og realistisk innsats for å utføre, samle inn, lagre og strukturere den foreslåtte effektoppfølgingen.

1.3 Metode

Resultatene som presenteres i denne rapporten er fremkommet gjennom en kombinasjon av litteraturstudier, internettsøk og sammenstilling av bidragsytneres egne erfaringer.

Først er det utført litteraturstudier innefor flere perspektiver. Resultatene er presentert i kapittel 2. Denne delen av rapporten er i første rekke en syntetisering av eksisterende kunnskap fra tidligere forskning. Kapittel 3 er basert på litteraturstudier og internettsøk, med fokus på praktiske eksempler av effektvurderinger.

Kapittel 4 inneholder eksempler på utførte ex post-evalueringer av store statlige prosjekter i Norge. Vi har ikke lyktes i å finne noen systematisk gjennomgang av eller samlet oversikt over utførte ex post-evalueringer av store investeringer prosjekter. Kapittel 4 er basert på det man har funnet mest relevant ved gjennomgang av blant annet Riksrevisjonens oversikter og stortingsdokumenter fra de siste årene. Rapportens fremstilling utgir seg ikke for å være fulldekkende, eller basert på strukturert uttrekk fra en større populasjon av prosjekter. Likevel bør antallet prosjekter sikre en viss representativitet for denne type evalueringer. Man kan stille spørsmål ved om sammenstillingen av evalueringer er representativ for store statlige prosjekter generelt, da det er en tendens til at evalueringer får større oppmerksomhet, og derfor er lettere tilgjengelige, for prosjekter som får oppmerksomhet enn for andre prosjekter. En vanlig grunn til at prosjekter får oppmerksomhet er at de får utfordringer i forhold til interessenter, kostnader eller fremdrift, ofte en kombinasjon.

Kapittel 5 presenterer et forslag til nytt opplegg for ex post effektvurderinger, og er basert på de tidligere kapitlene og rapportforfatterens erfaringer. Eksempler på hvordan den nye metoden kan anvendes på konkrete prosjekter er presentert i kapittel 6. Erfaringene fra prosjektet Nytt Rikshospital i dette kapitlet er til stor del basert på tilgjengelig informasjon om prosjektet og det ferdige sykehuset. Oppfølgingen av de fire jernbaneprosjektene er basert på flere typer informasjonskilder, inkludert gjennomgang av dokumenter, deltakende observasjon og intervjuer.

Forskningsmetodiske utfordringer ved ex post-effektvurderinger taes opp i de respektive kapitlene av rapporten, bl.a. spørsmålet om kontrafaktiske vurderinger.

2 Eksisterende teori om ex ante og ex post-evaluering og effektvurdering

Avhengig av hva man tillegger som relevant innhold i ex post-vurderinger og hvordan man vurderer ulike eksisterende metoder, kan det nok argumenteres for hvorvidt en metode er relevant eller ikke for anvendelse ex post i prosjektsammenheng. Vi har imidlertid måttet foreta et utvalg, og vi startet innledningsvis med å definere fire grovere fagområder vi antok ville være aktuelle:

- 1) Samfunnsøkonomiske analyser, en samlebetegnelse for ulike metoder som tradisjonelt vurderer effekter av et prosjekt som det finnes etablerte måter å prissette, men etter hvert også effekter som tidligere har vært vanskelige å kvantifisere i økonomiske termer.
- 2) Analyser av samfunnsnytte, igjen et samlebegrep, for metoder for å fange opp også andre effekter for samfunnet av et prosjekt enn rent økonomiske forhold.
- 3) Analyser av samferdselsmessige effekter, dette området ble tatt med for å fange opp eventuelle metoder som er utformet for å vurdere mer operasjonelle samferdselsmessige forhold som ikke dekkes av de to foregående. Det reflekterer også det faktum at en stor andel av offentlige prosjekter i den størrelseskategori Concept arbeider med er innen samferdselssektoren.
- 4) Metoder og teknikker fra kvalitetsledelse og prestasjonsvurdering, der det vi visste at det fantes verktøy som i utgangspunktet ikke har vært utviklet for anvendelse mot prosjekter, men som etter hvert har vært forsøkt konvertert til slik bruk.

Hvert av disse områdene har i prosjektet vært representert av ett fagmiljø, og disse foretok deretter en utredning av hvilke metoder som fantes innenfor sitt felt. Denne øvelsen avdekket noe overlapp mellom feltene, spesielt mellom samfunnsøkonomiske analyser og samferdselsmessige vurderinger, men også mellom samfunnsøkonomiske analyser og vurderinger av samfunnsnytte. Etter å ha eliminert slik overlapp fremstod de følgende som relevante metoder å beskrive nærmere:

- Nytte-kostnadsanalyse, med andre varianter
- Samfunnseffekt-analyse
- Prestasjonsvurdering
- Benchmarking

Videre vil vi gjennomgå detaljer rundt disse ulike metodene, som dels er i bruk og dels anbefales brukt, da primært i dag i ex ante-effektvurderinger. Det må også poengteres at disse metodene i stor grad er utviklet for ex ante-analyser. Grunnen til at vi inkluderer dem her er at de gjerne kan anvendes ex post, men da basert på reelle data etter at prosjektet er fullført.

2.1 Nytte-kostnadsanalyse

En nytte-kostnadsanalyse er en systematisk kartlegging av fordeler og ulemper ved et tiltak (i mange fagtradisjoner benyttes ofte begrepet tiltak om den investeringen som vurderes, og som i vår diskusjon er ensbetydende med prosjekt). Nyttevirkninger og kostnader verdsettes i kroner så langt det er faglig holdbart. Dersom en legger sammen den beregnede verdien av alle effektene ved et tiltak og summen av disse er større enn kostnadene, sier en at tiltaket er samfunnsøkonomisk lønnsomt. Et standard utgangspunkt for samfunnsøkonomiske evalueringer er nytte-kostnadsanalyse, enten det dreier seg om ex post eller ex ante-evalueringer. Faglig sett er dermed metodeapparatet i ex post-evalueringer prinsipielt det samme som ved ex ante-vurderinger.

Det foreligger lite litteratur om spesifikke metoder og praksis for samfunnsøkonomisk evaluering av store investeringsprosjekter. Riktignok har Finansdepartementet utarbeidet en veileder for evalueringer (Finansdepartementet, 2005a). Dette er i stor grad skrevet ut fra et styringsperspektiv og tar i liten grad opp det metodiske knyttet til kartlegging og analyse av samfunnsøkonomiske effekter. Derimot finnes det en omfattende litteratur om samfunnsøkonomiske analyser ut fra et ex ante-perspektiv. Blant annet gir Finansdepartementet en generell oversikt i sin veileder for økonomiske analyser (Finansdepartementet, 2005b). Ellers har også Statens Vegvesen utviklet en metodikk beskrevet i *Veileder for etterprøving av prissatte konsekvenser av store vegprosjekt* (Statens vegvesen, Utbyggingsavdelingen, 2006).

Til grunn for analyser av denne typen ligger en tankegang om at den samlede velferden i samfunnet kan påvirkes av hvordan ressurser disponeres. Dersom en flytter ressurser fra en anvendelse til en annen, og samfunnets samlede nytte øker

mer som følge av den nye anvendelsen enn den reduseres av at ressursene tas ut av den gamle anvendelsen, øker samfunnets samlede velferd. Et offentlig prosjekt vil være samfunnsøkonomisk lønnsomt dersom det har en slik effekt, det vil si at det bidrar til at samfunnets netto nytte øker.

Med dette som utgangspunkt er oppgaven skjematisk sett å beregne samfunnets netto nytte av det aktuelle prosjektet. Det teoretiske grunnlaget for denne type analyse er blitt kritisert fra flere vinkler. Dette gjelder blant annet det nytteteoretiske utgangspunktet og bruken av nytte-kostnadsanalyser som beslutningsverktøy for samfunnsmessige disposisjoner. Finansdepartementets anbefalinger, riktignok rettet mot ex ante-vurderinger, sier at nytte-kostnadsanalyse kan velges når de viktigste plankonsekvensene er økonomiske og målt i kroner, og når man ønsker å sammenligne flest mulig virkninger ved hjelp av en måleenhet som alle kjenner (kroner) Tidligere arbeider i Concept-programmet peker på at samfunnsøkonomisk lønnsomhet gir en bedre oppsummering av et tiltaks konsekvenser jo mindre betydningsfulle fordelingsvirkningene er, jo flere konsekvenser vi finner det faglig forsvarlig å verdsette i kroner, og jo mindre innslag det er av etisk vanskelige spørsmål.

Håndtering av inntekter og nytte

Et viktig spørsmål er; når bør nytte beregnes i kroner? Finansdepartementet peker i sin veileder på at et kriterium for når en bør verdsette effekter i kroner, er om verdsettingen gir beslutningstakerne et bedre og mer utfyllende bilde av tiltakets effekter enn om man ikke foretar slik verdsetting. Dette kan være tilfelle for noen typer investeringer, for eksempel når det er snakk om å sammenligne alternativer av noenlunde samme art og der slik verdsetting er meningsfull, men kan i andre tilfeller være mindre relevant.

Realøkonomiske virkninger/effekter representerer endring i bruk og tilgang av realressurser i en økonomi. Med realressurser forstår vi goder som har en økonomisk verdi, enten de er prissatte eller ei. Som eksempel kan bedret framkommelighet i et transportsystem bidra til økonomiske utvikling fordi en kan utnytte eksisterende innsatsfaktorer bedre (gjennom sparte tids- og kjørekostnader) eller tiltrekke nye innsatsfaktorer gjennom allokering av ressurser dit hvor avkastningen er størst. Dette kan endre konkurranseforholdene for næringslivet. En effekt kan være økt effektivitet via utnyttelse av stordriftsfordeler (bedre utnyttelse av produksjonskapasiteten), eller et skjerpet konkurranseklima der et tidligere geografisk skjermet næringsliv blir eksponert for konkurranse fra andre regioner. Et annet eksempel kan være endringer i tilgjengelighet som ikke nødvendigvis

reflekteres i økt bruk av transport. Dette kan vi kalle opsjonsverdi. Dette kan eksemplifiseres ved at det i visse strøk av landet kan ha en reell verdi å vite at man raskt kan komme under medisinsk behandling, uavhengig av egen reiseaktivitet ellers.

Dersom noen av effektene skulle oppstå utenfor prosjektets influensområde, det område som prosjektet primært har hatt som målsetning å påvirke, så vil denne type effekter også være realøkonomiske, og komme i tillegg til de som er tatt inn i analysen. Definisjon og avgrensning av tiltakets kraftfelt er derfor viktig for å kunne fange opp relevante virkninger, både når det gjelder samfunnsøkonomiske virkninger og fordelingsvirkninger. Realøkonomiske virkning kan deles inn i tre hovedgrupper:

- 1) Kvantifiserbare og prissatte virkninger.
- 2) Virkninger som er kvantifiserbare i fysiske størrelser, men som ikke lar seg verdsette i kroner.
- 3) Ikke-kvantifiserbare virkninger.

Det blir arbeidet kontinuerlig for å prissette virkninger, det vil si å flytte virkninger fra gruppe 2 til gruppe 1. Det settes våren 2007 i gang et omfattende arbeid med å revidere og utvide grunnlaget for virkninger i gruppe 1 og 2. Pris er en skala, og det gir større grad av sammenlignbarhet dersom man klarer å måle virkninger langs samme skala. Eksempler på virkninger av prosjekter i gruppe 1 er tidskostnader, byggekostnader, billettinntekter, enkelte typer risikoreduksjon samt tiltakskostnader. Virkninger i gruppe 2 og 3 kan imidlertid også være viktige. Eksempler på slike virkninger er utslippstyper som ikke er prissatt og eventuell verdi av å ha mulighet til å reise eller av å kunne utvikle reiseintensive næringsaktiviteter (opsjonsverdi).

En annen viktig faktor som må vurderes ved en ex post-evaluering er ringvirkninger. Ofte brukes argumenter for gjennomføring av et prosjekt at det skaper arbeidsplasser eller gir ringvirkninger. Slike argumenter kan referere til at prosjektet skaper økt etterspørsel av ulike leveranser til prosjektet, utbygging av kommunikasjoner, servicefunksjoner av forskjellig slag, økt omsetning som følge av økt forbruk som følge av inntekter fra prosjektet, mv. Finansdepartementet sier i sin veileder at det generelt må stilles strenge krav til det empiriske grunnlaget for å kunne regne inn ringvirkninger i en samfunnsøkonomisk analyse.

Johansen (1977) peker på at ved å trekke inn ringvirkninger kan en få nær sagt et hvilket som helst prosjekt til å se samfunnsøkonomisk lønnsomt ut. Han viser til at det er nødvendig å være presis med hensyn til hva alternativene er. I områder med

sysselsettingsvansker kan det være relevant å trekke inn slike ting. En må da imidlertid stille spørsmålet om alternativet til gjennomføring av prosjektet er at det ikke kommer noen stimulans til ulike aktiviteter som gir økt sysselsetting. Hvis forskjellige aktiviteter kan stimuleres på andre måter, vil det riktige synspunktet være at de ressursene som brukes opp gjennom de såkalte ringvirkningene er knappe ressurser som trekkes bort fra andre anvendelser. Bærerne av de såkalte ringvirkningene, arbeidskraft, vare- og tjenesteleveranser til prosjektet, osv. er da kostnader som ikke skal behandles på noen annen måte enn som regulære kostnader ved prosjektet. I følge Johansen (1977) er det vanskelig å se at kjøpekraft skapt ved gjennomføring av et prosjekt og som stimulerer omsetningen av forbruksvarer i et område, skal føres inn i en lønnsomhetsberegning som en positiv faktor. Denne formen for stimulans kan en normalt få på andre måter uten å benytte ressurser gjennom et prosjekt.

Fordelingsvirkninger er et annet moment i vurderingen av nytte. Samfunnsøkonomiske analyser har tradisjonelt ikke behandlet virkninger for ulike konsumentgrupper, landsdeler eller næringer av offentlige tiltak, eller interessekonflikter knyttet til tiltakene. Analysen fokuserer på om tiltaket totalt sett er positivt for samfunnet, selv om bare noen grupper har direkte nytte av tiltaket. Finansdepartementet (2005b) legger imidlertid betydelig vekt på at fordelingsvirkningene av et tiltak beskrives i analysen. Spørsmålet vil dermed være om ulike grupper i samfunnet i forskjellig grad drar nytte av investeringer, og hvordan det i så fall skal telle med i vurderingen av prosjektet.

Vurdering av fordelingsvirkninger i en samfunnsøkonomisk analyse er i praksis relativt upløyd mark. Skal man korrigere nytteverdiene for for eksempel ulike sosiale lag eller inntektsgrupper? I forbindelse med beslutningen om nytt operabygg i Oslo kom det for eksempel opp i diskusjonen en påstand om at det i hovedsak er høyere inntektsgrupper som drar nytte av opera. Hvis det hadde vært riktig, skulle da nytten korrigeres, og i så fall i hvilken retning? Et annet spørsmål som kan være relevant for investeringer i distrikter er om nytten for fastboende skal vurderes annerledes enn nytten for for eksempel turister og hyttefolk, selv om betalingsvilligheten kan være lik. Dette vil kanskje særlig gjelde investeringer i transportinfrastruktur. Men det er også slik i dag at staten gjennom Telenors konsesjonsmessige forsyningsplikt støtter fasttelefoniinfrastruktur for fastboende, men ikke for hyttefolk.

Håndtering av kostnader

I en samfunnsøkonomisk analyse skal man bruke priser som reflekterer de realøkonomiske kostnadene ved å benytte ressurser i et prosjekt. Utgangspunktet er

at alle ressurser har en alternativ anvendelse. De realøkonomiske kostnadene ved å benytte for eksempel ressursen arbeidskraft i et prosjekt er lik verdien av denne arbeidskraften i andre prosjekter. Kalkulasjonsprisene i en samfunnsøkonomisk analyse skal altså reflektere alternativverdien av de ressursene som brukes. Svært ofte bruker man markedsprisene som kalkulasjonspriser, og ofte er de to prisene identiske. Ulike former for markedssvikt medfører imidlertid at det kan oppstå avvik mellom markeds- og kalkulasjonsprisen, for eksempel i markeder med monopoler, som følge av skatter og avgifter eller når det er arbeidsledighet.

Verdsetting av arbeidskraftinnsats er et viktig moment i kostnadssiden. Offentlige prosjekter vil konkurrere både med privat og annen offentlig virksomhet om arbeidskraft og vareinnsats. Private bedrifter og offentlig virksomhet vil vurdere etterspørselen etter arbeidskraft ut fra lønn før skatt inklusive arbeidsgiveravgift. Konsumentene vil imidlertid vurdere sitt arbeidstilbud i forhold til lønn etter skatt. Dersom offentlig bruk av arbeidskraft dels fortrenger arbeidsinnsats fra annen privat eller offentlig produksjon og dels fortrenger fritid blir kalkulasjonsprisen for arbeidskraft et veid gjennomsnitt av lønn før og etter skatt. I Finansdepartementets anvisninger anbefales som en hovedregel å bruke lønn inklusive skatt og arbeidsgiveravgift.

En annen betydelig innsatsfaktor i prosjekter er innsatsvarer. Innsatsvarer fortrenger annen privat og i noen grad offentlig bruk, og/eller medfører økt produksjon. Varene bør derfor i utgangspunktet verdsettes til markedspris, men eksklusive generelle avgifter som merverdiavgiften. Dette er i tråd med anbefalingene i Finansdepartementet (2005b).

Hvordan særavgifter skal behandles er mer komplisert. Særavgifter som bidrar til at prisen på samme gode kan bli forskjellig i ulike anvendelser bør ikke tas med i kalkulasjonsprisen. Avgifter som er innført for å delvis korrigere for såkalte eksterne virkninger, i første rekke globale og lokale miljøvirkninger, bør tas med dersom man ikke eksplisitt verdsetter miljøskadene i analysen. Det er viktig å merke seg at man altså ikke både kan inkludere miljøavgiften og anslaget på miljøkostnaden.

Hvis mulig bør man anslå miljøkostnadene, selv om dette metodisk er svært krevende, og trekke ut miljøavgiften dersom den finnes. Regelen om å verdsette på grunnlag av alternativverdier gjelder også eventuelle miljøkostnader. Korrekte kalkulasjonspriser for miljøskader er lik verdien av ren luft, rent vann og liknende.

En rekke innsatsfaktorer vil skaffes til veie gjennom import. For å kunne vurdere ulike alternativer i forhold til annen ressursbruk, må kostnadene reflektere

ressursenes verdiskaping i beste alternative anvendelse. Hva som er beste alternative anvendelse i det enkelte tilfelle må vurderes nærmere. Ofte vil import knyttet til en offentlig investering i liten eller ingen grad konkurrere med tilsvarende private investeringer. Dette betyr at kalkulasjonsprisen bør settes til pris ved grensen, dvs. eksklusive avgifter. Dette gjøres for eksempel ved verdsetting av kjøp av politibiler, hvor bilene verdsettes til pris eksklusive særavgifter. Denne kalkulasjonsprisen følger av at prisen ved grensen (eksklusive avgifter) er alternativkostnaden for landet, mens merverdiavgiften og evt. andre avgifter kan betraktes som rene overføringer internt i statsbudsjettet.

Når det er snakk om investeringer med stor importandel vil også hvilke valutakurser som brukes i praksis kunne ha betydning for kalkylene, både med hensyn til avveining mellom import og hjemmeproduserte varer og tjenester, og med hensyn til import fra ulike regioner der valutakursene vil utvikle seg ulikt, for eksempel USA, Eurosonen, Kina eller Japan. Så langt har imidlertid staten regnet seg som risikonøytral i forhold til valutausikkerhet og ikke gitt anvisninger om hvordan slik usikkerhet bør håndteres.

Skattefinansiering av offentlige prosjekter innebærer kostnader for samfunnet som må inkluderes i den samfunnsøkonomiske analysen. Skatten utgjør en kile mellom prisen til tilbyder og prisen til den som etterspør for eksempel arbeidskraft. Skatten bidrar derfor til vridninger i ressursbruken som gir et effektivitetstap for samfunnet. Skattefinansiering innebærer økt skatt, og derfor mer effektivitetstap for hver krone som investeres.

Analysen gir ulike anslag på hva den reelle kostnaden ved skattefinansiering er. Finansdepartementet anbefaler å bruke en skattekostnad på 20 øre per krone for netto økt offentlig finansiering som følge av et offentlig tiltak. Man skal korrigere for mulige effekter av tiltaket som bidrar for eksempel til å øke arbeidstilbudet, noe som vil øke de offentlige skatteinntektene og dermed redusere statens netto finansieringsbehov.

Spesielt om samfunnsøkonomisk verdsetting av miljøeffekter

Miljøområdet er kanskje det området der en har arbeidet lengst med å identifisere og beregne samfunnsøkonomiske effekter. På dette området er det utviklet en rekke metoder for å verdsette ikke prissatte miljøeffekter, slik at disse kan gå inn i samfunnsøkonomiske kalkyler. Blant annet gjelder dette hvordan man skal sette verdi på kostnader knyttet til ulike former for forurensninger. Samtidig er miljøspørsmålet blitt stadig viktigere og miljøeffekter av prosjekter får større

betydning. Dermed vil kravet til nøyaktighet i slike effektestimeringer øke, og per i dag er nok denne ikke god nok.

I mangel av markedspriser på utslipp er det nødvendig å finne andre, indirekte måter å anslå den samfunnsøkonomiske verdien av de eksterne effektene av ulike former for forurensninger. Fire metoder som ofte benyttes for å anslå miljøkostnader er:

- Skadekostnader. De fysiske skadene fra utslipp beskrives, og prisen på skader estimeres. For eksempel kan lønnskostnader benyttes til å verdsette antall dagers fravær fra jobb som følge av helseskader. Kostnader ved økt dødsrisiko fra skadelige utslipp kan anslås gjennom å se på hva vi faktisk betaler for å redusere dødsrisiko på andre områder. Skader på bygninger og materialer kan anslås ved hjelp av reparasjonskostnader eller lignende.
- Tiltakskostnader. De marginale kostnadene ved tiltak for å redusere de samme utslipp fra andre kilder eller anslag for kostnadene ved ytterligere reduksjoner kan tas som en indikasjon på hva samfunnet minst er villig til å betale for å redusere utslipp (implisitt verdsetting).
- Betalingsvillighetsundersøkelser (betinget verdsetting). En spør her om hva folk er villige til å betale for å unngå skadene fra en bestemt type utslipp, eller hvor mye de må ha for å akseptere skadene. En kan også be et ekspertpanel om å verdsette skadene.
- Hedoniske metoder. Boligprismetoden er den vanligste av de hedoniske metodene og bygger på en antagelse om at støynivå, luftkvalitet og andre ulemper påvirker markedsprisen på ellers like boliger. Ved å observere forskjeller i boligpriser mellom ulike områder kan en dermed anslå betalingsvilligheten for investeringer som reduserer miljø- og støvbelastning.

De forskjellige framgangsmåtene har en tendens til å gi ulike resultater. Dette indikerer at anslagene for miljøkostnadene er usikre. Usikkerheten kan skyldes mangel på kunnskap om kortsiktige og langsiktige negative effekter av utslippene, og vanskeligheter med å avsløre hvordan folk flest eller myndighetene verdsetter disse effektene eller hvor mye de er villige til å betale for å redusere eller unngå effektene. Kostnadene vil dessuten vanligvis endre seg over tid som følge av blant annet endret betalingsvillighet og fordi skadevirkningene kan endres når utslippene endres.

Kostnadsberegningssutvalget (1997) anbefalte i sin tid at en i verdsetting av miljøskader i størst mulig grad utnytter den informasjonen som ligger i observert adferd. Utvalget pekte på at betinget verdsetting er best egnet til verdsetting av lokale miljøulemper som for eksempel støy og støvplager, og at det må stilles særlig høye metodekrav til undersøkelsene. Et problem ved å ta utgangspunkt i observert adferd er at analysen da bygger på at eksisterende atferd er en rasjonell og

tilstrekkelig måte å forholde seg til miljøproblemene på. Dette kan være en forutsetning som kan diskuteres, for eksempel når det gjelder forbruk av fossilt brensel, klimagassutslipp og reduksjon av biodiversitet og naturens biologiske produksjonskapasitet.

Odgaard, Kelly & Laird (2005) har undersøkt hvilke nasjonale vurderingsmetoder som brukes i forbindelse med transportinfrastrukturprosjekter i Trans European Network (TEN). Undersøkelsen omfatter de 25 medlemslandene i EU samt Sveits, og retter spesiell oppmerksomhet mot nytte-kostnadsanalysers rolle i prosjektvurderingene. Planleggingen av TEN-prosjektene må anses som planlegging på et forholdsvis strategisk nivå, der den mer detaljerte prosjektplanleggingen ofte skjer for delprosjekter innenfor en større transportkorridor som utgjør en del av det transeuropeiske nettverket.

Undersøkelsen viser at det er stor variasjon mellom landene i måten prosjektvurderingene gjennomføres på. Men alle landene benytter nytte-kostnadsanalyser i en eller annen form i forbindelse med vurderinger av transportprosjekter. Denne metoden var blitt mer utbredt ved undersøkelsestidspunktet i 2004 enn da en tilsvarende undersøkelse ble gjennomført syv år tidligere. I en rekke av de nye, østeuropeiske landene brukes imidlertid nytte-kostnadsanalyser oftest eller bare for prosjekter der det søkes om EU-finansiering.

Selv om nytte-kostnadsanalyser brukes for transportprosjekter i alle de undersøkte landene, står denne metoden vanligvis ikke alene, men kombineres med for eksempel multikriterieanalyse (vekt- og poengmetode uten verdisseting av konsekvensene i form av pengeverdi), kvantitativ vurdering (måling av effektene på forholdstallsskala, men uten vekting), og/eller kvalitative analyser (poengmetode der konsekvensene kategoriseres på ordinalskala, uten vekting).

Det er også stor variasjon mellom landene i måten nytte-kostnadsanalyser brukes på. Dette gjelder både hvilke konsekvenskategorier som verdissetes og tas med i den økonomiske beregningen, hvordan viktige konsekvenskategorier verdsettes, og hvilke diskonteringsrater som benyttes. Som det framgår av Figur 2 går Danmark og Nederland lengst i å verdissete miljøvirkninger (herunder både støy, luftforurensning og klimagassutslipp), indirekte sosioøkonomiske effekter og andre vanskelig verdissetbare konsekvenstyper. Portugal, Estland og Litauen peker seg ut i den andre enden av skalaen med verdisseting av færrest ikke-markedsprissatte virkninger.

For transportinfrastrukturprosjekter er spart reisetid for brukerne av de nye anleggene ofte en dominerende post på nyttesiden. Verdissetingen av denne tiden

varierer imidlertid mye. For reiser til arbeidsplass varierer verdisetningen per minutt spart ved arbeidsreiser med en faktor 20 mellom den høyeste (maksimumssatsen i Østerrike) og laveste (minimumssatsen i Ungarn). Hvis vi bare sammenlikner henholdsvis maksimums- og minimumssatser, er det en variasjon med faktor 10 både for maksimumssatsene (mellom Østerrike og Hellas) og minimumssatsene (mellom Italia og Ungarn). For andre reiser enn reiser mellom bolig og arbeidsplass er det også store forskjeller, med variasjon med faktor 15 mellom høyeste og laveste sats. Det er store forskjeller også i verdisetningen av tidsbesparelser for godstransporten, trengselskostnader ut over selve tidsforbruket, miljøvirkninger og trafikksikkerhet. Verdisetningen av et dødsfall som følge af trafikkulykke varierer for eksempel med en faktor 10 mellom Finland og Slovakia, og verdisetningen av et tonn CO₂-utslipp varierer med en faktor 5 mellom Sverige og Finland. Kalkulasjonsrenten som brukes i analysene viser også stor variasjon: I fire av de undersøkte landene benyttes en kalkulasjonsrente på under 4 %, i 17 land ligger diskonteringsfaktoren mellom 4 % og 7 %, og i tre land er den over 7 %. Materialet viser at det ikke er noe klart regionalt mønster i variasjonen i kalkulasjonsrente.

		Byggekostnader	Forstyrrelser som skyldes bygging	Kostnader til vedlikehold, drift og administrasjon	Besparelser i reisetid for passasjerer	Brukerbetaling og inntekter	Driftskostnader for kjøretøyer	Fordeler for godstrafikk	Sikkerhet	Støy	Lufforurensning - lokal/regional	Klimaendring	Indirekte sosio-økonomiske effekter
Nord/vest	Østerrike												
	Belgia												
	Danmark												
	Finland												
	Frankrike												
	Tyskland												
	Irland												
	Nederland												
	Sverige												
	Sveits												
	Storbritannia												
Øst	Tsjekkiske republikk												
	Estland												
	Ungarn												
	Latvia												
	Litauen												
	Polen												
	Slovakiske republikk												
	Slovenia												
Sør	Kypros												
	Hellas												
	Italia												
	Malta												
	Portugal												
	Spania												

Fargekoder:

	Inkludert med pengeverdi
	Målt kvantitativt, kvalitativt eller ikke inkludert

Figur 2 Behandling av ulike konsekvenskategorier i nytte-kostnadsanalyser av transportprosjekter (der slike analyser benyttes), fordelt på ulike land, kilde Odgaard et al., 2005

Disse eksemplene viser at bruken av nytte-kostnadsanalyser ved vurdering av store offentlige transportinvesteringsprosjekter langt fra er noen eksakt vitenskap, og at det ligger et betydelig innslag av skjønn og politiske prioriteringer i de forutsetningene som analysene bygger på. Odgaard et al. (2005) peker på at den store variasjonen i forutsetninger særlig utgjør et problem ved vurdering av alternative løsninger for prosjekter som går gjennom flere land, slik tilfellet er med en rekke av korridorane i det transeuropeiske nettverket. En rekke forskere har dessuten pekt på at nytte-kostnadsanalyser innenfor transportsektoren historisk har hatt en innebygd tendens til å favorisere vegbygging for å imøtekomme etterspørselen etter mer vegkapasitet – også i situasjoner der det foreligger miljø- og samferdselspolitiske mål om å dempe denne etterspørselen og redusere trafikkmengdene (Kenworthy, 1990; Grove-White, 1997; Sactra, 1999; Arge,

Homleid og Stølan, 2000; Owens & Covell, 2002, Tennøy, 2003; Owens, Rayner og Bina, 2004; Nielsen & Fosgerau, 2005). Disse innvendinger mot nytte-kostnadsanalyser i ex ante-analyser vil for øvrig i stor grad også gjelde for anvendelser i ex post-evalueringer.

Under den større overskriften samfunnsøkonomiske analyser finnes det for øvrig flere metoder enn nytte-kostnadsanalyse, som nettopp ble beskrevet. De to vanligste, utover nytte-kostnadsanalyse, er:

- Kostnadseffektivitetsanalyse, er en systematisk verdsetting av kostnadene ved ulike alternative tiltak som kan realisere samme mål. Kostnadene verdsettes i kroner, og man søker å finne den mest kostnadseffektive måten å nå et gitt mål.
- Kostnads-virkningsanalyse, er en kartlegging av kostnader for ulike tiltak som er rettet mot samme problem, men der effektene av tiltakene ikke er helt like. En kan i slike tilfeller ikke uten videre konkludere med at det tiltaket som har lavest kostnader er å foretrekke, men må vurdere kostnadene ved ulike alternativer opp mot ulike virkninger av alternativene.

2.2 Samfunnseffekt-analyse

Som det delvis har vært nevnt under behandlingen av nytte-kostnadsanalyse, vil et prosjekt kunne medføre en rekke effekter, både positive og negative, som ikke direkte lar seg angi kvantitativt. En utvidet analyse, som forsøker å inkludere slike effekter, er samfunnseffekt-analyse.

Når vi her snakker om samfunnsnytte, anlegger vi et bredt perspektiv. Vi snakker om nytte for ulike befolkningsgrupper, og ikke bare nytte slik den for eksempel oppfattes av det politiske flertallet og dets representanter – selv om det siste selvsagt vil være viktig i et demokrati- og styringsperspektiv. Vi snakker dessuten om både sosiale, miljømessige og økonomiske aspekter av nytte – der alle disse tre begrepene er bredt definert. Og ikke minst snakker vi om både positiv og negativ nytte, vel vitende om at de færreste prosjekter bare har positive effekter. Store offentlige investeringsprosjekter bidrar normalt til å skape nytte i forhold til de behovene og interessene som prosjektene er begrunnet med, men kan samtidig ha bieffekter som reduserer nytten i forhold til enkelte andre hensyn og interesser.

Samtidig må det sies at denne typen bredere anlagte evalueringer av prosjekter, ex post, er lite utbredt. I de seneste årene er det tatt enkelte nasjonale initiativer for mer

systematiske ex post-vurderinger, men da gjerne i et snevrere perspektiv. I Storbritannia introduserte det britiske vegdirektoratet i 2001 det såkalte Post Opening Project Evaluation-programmet (POPE) for å evaluere effekter som følge av gjennomføring av store vegprosjekter. Disse evalueringene har begrenset seg til å omfatte samlet trafikkvolum, reisetid og ulykker, og er utført med forholdsvis beskjeden ressursinnsats (gjennomsnittlig 12.000 pund per prosjekt, tilsvarende i underkant av NOK 150.000). I 2004/2005 evaluerte det britiske vegdirektoratet om lag 50 prosjekter i henhold til denne ordningen (Oxera, 2005).

Ifølge Lauridsen (2001) praktiserte Norge og Sverige, men ikke Danmark ex post-evalueringer av store vegprosjekter ved inngangen til dette århundret. Disse ex post-evalueringene dreide seg dels om selve planprosessene, dels om evalueringsmetodene som ble brukt i ex ante-evalueringene. I Norge tok Samferdselsdepartementet i 2003 initiativ til mer omfattende før-etterundersøkelser av store samferdselsprosjekter, i første omgang avgrenset til prosjekter med en investeringsramme over 100 millioner NOK (Olsson, 2004). Ifølge Samferdselsdepartementet er det ønskelig at slike evalueringer omfatter alle sentrale parametre som inngikk i beslutningsgrunnlaget, inkludert kostnader, trafikktvikling, ulykker, miljø og regionale virkninger.

I Sverige er Banverket og Vägverket siden 1997 pålagt å utføre etterkalkyler av ferdigstilte investeringer. Statens Institut för Kommunikationsanalys (SIKA) har dessuten siden 1999 utført analyser av hvordan de transportpolitiske målene blir oppfylt innenfor hele transportsektoren i Sverige. Vi gir en nærmere omtale og drøfting av noen av disse ordningene som praktiseres rundt om i verden senere i rapporten.

Ved siden av de ovenfor nevnte, myndighetsinitierte rutinene for ex post-evalueringer har transportforskere og andre samfunnsforskere gjennomført en del studier av store offentlige investeringsprosjekter. Blant disse har særlig arbeidene til Flyvbjerg og hans kolleger oppnådd stor oppmerksomhet (Flyvbjerg, Holm & Buhl, 2002 og 2005; Flyvbjerg, Bruzelius & Rothengatter, 2003). Disse studiene har først og fremst gått i dybden når det gjelder transportprosjekter, men resultatene er også diskutert i lys av erfaringer med andre typer store offentlige investeringsprosjekter, bl.a. store offentlige bygninger og bygningskomplekser. I sine studier av transportprosjekter fokuserer Flyvbjerg og hans kolleger særlig på forholdet mellom prognostiserte og faktiske trafikk tall (antall passasjerer for baneprosjekter og antall kjøretøyer for vegprosjekter), og forholdet mellom budsjetterte og faktiske byggekostnader. For noen prosjekter suppleres disse analysene med vurderinger av i hvilken grad de miljøkonsekvensvurderingene som ble utført på forhånd, fanget opp

de faktiske miljøvirkningene av prosjektene. Flyvbjerg og hans kollegers undersøkelser dekker imidlertid et betydelig smalere spekter av konsekvenskategorier enn det en effektvurdering i forhold til samfunnsnytte vil måtte omfatte.

Vurderinger av samfunnsnytte ex ante

I mangel av eksisterende teori og praksis for bredere anlagte vurderinger av samfunnsnytte i etterkant, vil vi beskrive noen tilnærminger til slike vurderinger i forkant av et prosjekts oppstart. Hensikten er å vise hvilke effekter som kan tas med i slike vurderinger og hvordan helhetlige evalueringer kan utformes.

The Green Book

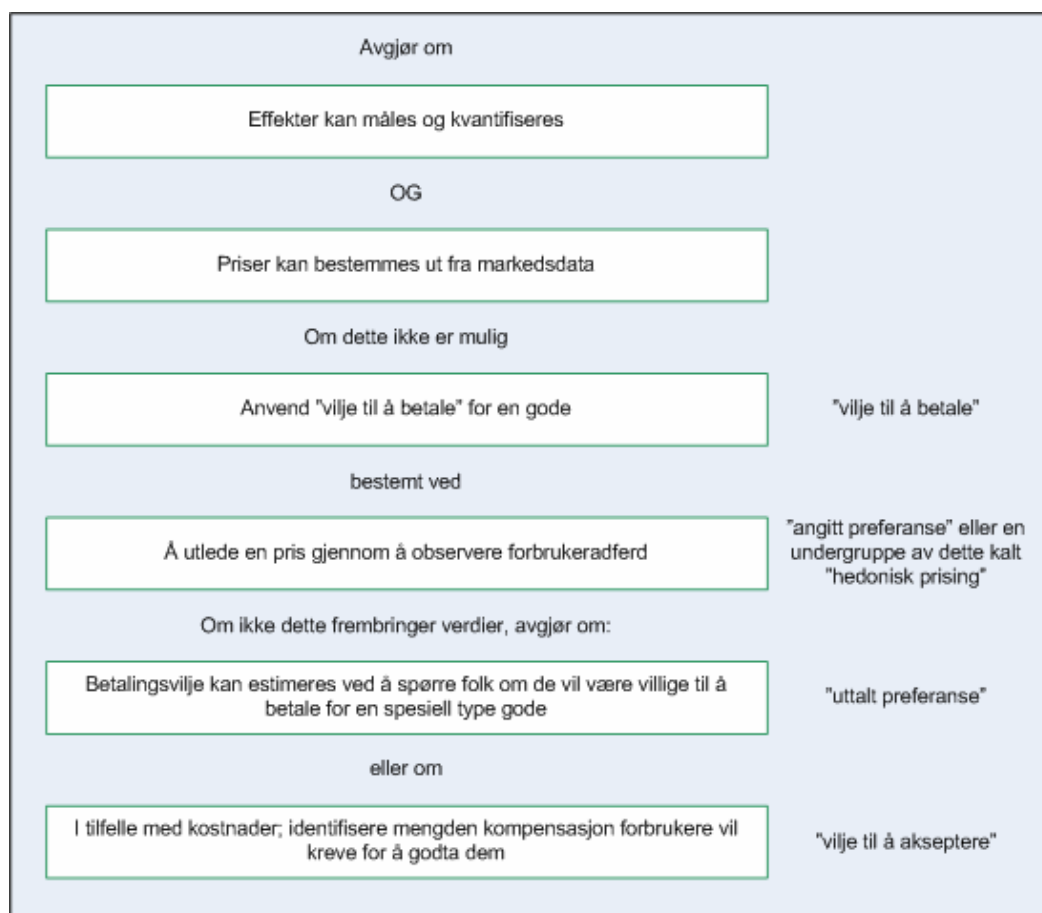
The Green Book er det britiske finansdepartementets veiledningshåndbok om vurdering av politikk, programmer og prosjekter (HM Treasury, 2004). Formålet med håndboken er å sikre at ingen politikk, programmer eller prosjekter vedtas uten at man først har funnet svarene på følgende spørsmål:

- Finnes det bedre måter å oppnå dette målet på?
- Kan disse ressursene brukes på bedre måter?

Håndboken beskriver hvordan økonomiske, finansielle, sosiale og miljømessige vurderinger av en politikk, et program eller et prosjekt bør kombineres. The Green Book omfatter både ex ante og ex post-effektvurderinger. Hovedvekten ligger imidlertid på effektvurderinger før planen, programmet eller prosjektet vedtas, og med fokus på effektvurderinger i den tidlige fasen av planleggingen. Begrepet ”option appraisal” brukes om vurderingene i denne fasen, til forskjell fra vurderinger av ”solutions”, som viser til vurderinger av bearbejdede versjoner av et valgt hovedkonsept.

The Green Book framholder vurdering av ulike løsningskonsepters nytte og kostnader som den grunnleggende metoden for å vurdere om et konsept er verdt å satse på. Formell nytte-kostnadsanalyse anbefales også som en viktig del av vurderingene. Generelt anbefaler The Green Book å prissette effektene i så stor grad som mulig. Håndboken inneholder en figur der den anbefalte framgangsmåten mht. prissetting av effekter beskrives, se Figur 3. Først må de som foretar effektvurderingen for hver av konsekvenskategoriene ta stilling til om effekten kan måles og kvantifiseres, og om det er mulig å fastsette priser ut fra markedsdata. Hvis

dette ikke er mulig, anbefales det å bruke betalingsvillighet (for fordeler) og aksepteringsvilje (for ulemper) til å fastsette prisene. Betalingsvilje anbefales dels undersøkt ved å se på hva folk faktisk har betalt for å få adgang til et tilsvarende gode ("revealed preferences"), og dels ved å spørre hva de hypotetisk ville være villige til å betale ("stated preferences").



Figur 3 Verdifastsettelse av effekter

Selv om The Green Book altså går inn for nytte-kostnadsanalyse som overordnet metode for effektvurderingene, er håndbokens anbefalinger mer nyanserte enn det som ofte kjennetegner litteraturen om slike analyser. Det legges blant annet stor vekt på å foreta særskilte analyser av hvordan fordeler og ulemper fordeler seg mellom ulike befolkningsgrupper. Dermed går man, i hvert fall delvis, klar av ett av de viktigste kritikkpunktene mot nytte-kostnadsanalyse som metode. I tillegg til å analysere fordelingsmessige konsekvenser, anbefaler The Green Book også justering for ulik forventet prisutvikling for forskjellige typer varer og tjenester, analyse av risiko, usikkerhet og verdien av fleksibilitet mht framtidige beslutninger, og ikke

minst korrigerende for vurderingsoptimisme¹. Det siste vil først og fremst være relevant for ex ante-vurderinger, men i den grad prosjektets samfunnsnytte påvirkes av antatte endringer i sentrale evalueringsparametere (for eksempel passasjertall) i perioden etter evalueringstidspunktet, vil korreksjon for vurderingsoptimisme også kunne være relevant i forbindelse med effektvurderinger ex post.

Selv om The Green Book går inn for å prissette så mange som mulig av konsekvensene, anerkjennes det at dette ikke vil være mulig for alle typer effekter. Håndboken understreker at de ikke-prissatte effektene ikke må glemmes, men identifiseres og vurderes opp mot de prissatte effektene. Boken peker på vekt- og poengmetoder som de mest utbredte måtene å håndtere ikke-prissatte effekter på, men peker også på enkle sjekklister med ytelseskrav i forhold til ulike kriterier som et alternativ.

The Green Books anbefaling av nytte-kostnadsanalyse som metode for effektvurderinger dreier seg om en ”mykere”, mer modifisert variant enn det som vanligvis foreskrives i metodelitteraturen. Likevel kan det være grunn til å problematisere om denne metoden egner seg til å vurdere hvor nyttig et gjennomført, stort statlig investeringsprosjekt er for samfunnet, og om prosjektet overhodet burde vært gjennomført. På transportområdet foreligger det for eksempel transportpolitiske målsettinger som springer ut av et ønske om å redusere negative miljøvirkninger av transporten. Noen av disse miljøvirkningene er av utpreget langsiktig karakter, for eksempel konsekvensene av utslipp av drivhusgasser. Det er svært vanskelig, for ikke å si umulig, å beregne den økonomiske verdien av forskjeller mellom alternative konseptløsninger mht. utslipp av drivhusgasser. Og selv om The Green Book opererer med gradvis redusert kalkulasjonsrente for langsiktige effekter, er også en lav diskonteringsrate av langsiktige, irreversible og alvorlige miljøskader problematisk (med mindre man forutsetter at ødelagt miljø blir kompensert gjennom vekst i forbruket av varer og tjenester). I situasjoner der nettopp slike vanskelig prissatte effekter kan være utslagsgivende for hvilken konseptløsning som politisk vil bli foretrukket, kan det virke ufornuftig å basere evalueringen av prosjektets verdi for samfunnet på nytte-kostnadsanalyse. Også det

¹ Basert bl.a. på de funn Flyvbjerg et al (2002) har gjort mht. avvik mellom prognostiserte og faktiske effekter av store transportinfrastrukturprosjekter, anbefaler The Green Book å justere anslagene for økonomiske utlegg, nytte og tidsforbruk for å korrigere for en utbredt tendens til overoptimistiske prognoser. Korreksjonene bør ifølge The Green Book skje ved å bruke empiriske data fra andre tilsvarende prosjekter. Dette likner mye på den tilnærmingen Flyvbjerg et al. (2006) anbefaler: Å sammenlikne det foreslåtte prosjektet med statistiske data for økonomiske utlegg og nytte (for eksempel passasjerbelegg eller trafikkvolum) innenfor et tilstrekkelig stort utvalg av sammenliknbare prosjekter. Sammenliknet med Flyvbjerg et al's anbefalte ”referanseklasseprognosemetode”, er anbefalingen i The Green Book likevel mindre ambisiøs, idet den åpner for mer kvalitativ og case-basert sammenlikning der man ikke nødvendigvis gjør bruk av statistiske gjennomsnitt eller sannsynlighetsfordelinger.

faktum at det er vanskelig å lage særlig presise kontrafaktiske vurderinger av en rekke av de relevante effektene, taler mot å benytte metoder som er avhengige av høy grad av presis kvantifisering².

Guidelines for Social Impact Assessment

Disse retningslinjene er utarbeidet av en gruppe samfunnsforskere og utgitt av US Dept. of Commerce, National Oceans and Atmospheric Administration og National Marine Fisheries Service, og hører med blant de få dokumentene man kan finne på Internett om hvordan sosiale virkninger av offentlige tiltak bør vurderes. Vurdering av sosiale virkninger av prosjekter og tiltak beskrives dels som et supplement til miljøkonsekvensanalyse, og dels som ”den samfunnsvitenskapelige komponenten av miljøvurdering”. Samfunnsmessige konsekvenser beskrives som:

”consequences to human populations of any public or private actions that alter the ways in which people live, work, play, relate to one another, organize to meet their needs and generally cope as members of society. The term also includes cultural impacts following changes to norms, values, and beliefs that guide and rationalize their cognition of themselves and their society.”

I dokumentet fokuseres det på sannsynlige konsekvenser av “policy actions” og visse typer offentlige tiltak. Heftet hevdes å være det første systematiske og tverrvitenskapelige dokument med retningslinjer og prinsipper som kan hjelpe offentlige organer og den private sektor til å bruke samfunnsmessige konsekvensvurderinger til å gjøre bedre beslutninger i saker som faller inn under den nasjonale miljølovgivningen.

Dokumentet argumenterer mot sentretning om økonomiske virkninger og den utbredte tanken om at ”penger vil kompensere for eventuelle negative virkninger”. Notatet peker på at analyser basert på rådende økonomisk teori som regel så bort fra

Fordeling av fordeler og ulemper

Hvilken viktighet folk tillegger sine nabolag og lokalsamfunn

Langvarige sosiale nettverk.

USAs miljølovgivning sier eksplisitt at ikke bare direkte effekter, men også indirekte og kumulative effekter (inkl. både estetiske, historiske, kulturelle, økonomiske,

² Når det gjelder transportprosjekter, bygger dessuten noen av de vanligste beregningsmodellene som det ville være nærliggende å benytte til slik kvantifisering, på meget forenklede, og i noen tilfeller misvisende, forutsetninger om hvilke årsak-virkningssammenhenger som skaper effektene. Se Næss, 2005 og 2006 for en kritikk.

sosiale og helsemessige forhold) må vurderes. Konsekvenser i forhold til urbefolkning framheves særskilt. Dokumentet trekker parallell fra miljøkonsekvensvurderingenes fokus på sårbare arter og økosystemer til sosiale konsekvensvurderingers fokus på sårbare grupper i befolkningen. Dokumentet peker videre på at ”sosial konstruksjon av virkeligheten”, holdninger, motstand og konflikter må tas alvorlig. Det å stemple opposisjon som ”følelsesladd” og ”feilinformert” vil akselerere konflikter og hindre samarbeid om løsning av problemer.

Dokumentet anbefaler en komparativ metode for å forutsi virkninger av de tiltakene som vurderes. Først å forsøke å forstå tidligere atferd blant individer og lokalsamfunn i det området som berøres av tiltaket. Dernest brukes en komparativ vurderingsmetode for å studere begivenhetsrekkefølgen i et samfunn hvor tilsvarende endringer er blitt gjennomført, og på bakgrunn av dette ekstrapolere hva som sannsynligvis vil skje i andre lokalsamfunn der tilsvarende tiltak er planlagt.

Dokumentet peker på fem typer informasjon om konsekvenser som særlig viktige i en vurdering av sosiale effekter, nemlig informasjon om endringer i befolkningssammensetninger, lokalsamfunns- og institusjonsstrukturer, politiske og sosiale ressurser (for eksempel nettverk), atferdsendringer blant individer og familier, og felles lokalsamfunnsressurser. Dokumentet anbefaler at konsekvenstyper og faser sammenstilles i en matrise som utgangspunkt for en analyse der oppgaven er å si noe om størrelsen og betydningen av effekten innenfor hver av matrisens ruter.

Dokumentet gjennomgår ti hovedtrinn som anbefales fulgt ved utarbeiding av en vurdering av sosiale effekter. Disse er vist i Tabell 1. I figuren har vi valgt å beskrive den anbefalte framgangsmåten mer detaljert for enkelte av trinnene enn for andre. Etter vår oppfatning er de mest detaljert beskrevne trinnene særlig relevante i forhold til debatten om hvilke metoder som egner seg best i effektvurderinger av store offentlige investeringsprosjekter. I de originale retningslinjene som figuren bygger på, er detaljeringsgraden mer omfattende også for de trinnene som i Tabell 1 er behandlet summarisk.

1. Offentlig involvering – utvikle en effektiv, offentlig plan for å involvere alle potensielle berørte parter. Start helt i begynnelsen av planleggingsfasen for prosjektet.
2. Identifisering av alternativer
3. Baseline-studium
4. Omfang: Identifiser hele bredden av de sannsynlige samfunnsmessige effekter
 - Metode:
 - Gjennomgang av eksisterende relevant litteratur
 - Offentlige undersøkelser
 - Teknikker for å sikre offentlig deltakelse
 - Relevante kriterier for å velge ut relevante effekter:
 - Sannsynligheten for at hendelsen skal inntreffe
 - Antall personer, inkludert urbefolkning, som vil berøres
 - Varigheten på effekter
 - Verdien av goder og kostnader for berørte grupper (intensiteten av effektene)
 - I hvilken grad påvirkningen er reverserbar eller kan håndteres
 - Sannsynligheten for at det påføres følgeeffekter
 - Referanser til nåværende og fremtidige strategiske, politiske beslutninger
 - Usikkerhet beheftet ved de mulige effekter
 - Tilstedeværelse eller fravær av kontrovers rundt det foreslåtte tiltaket
5. Vurdering av de estimerte effekter – undersøk de sannsynlige effekter med bakgrunn i
 - Forventede forhold uten tiltaket (nullalternativet)
 - Forventede forhold med tiltaket
 - Forventede effekter, som kan sees på som differansen mellom en fremtid med og uten det foreslåtte tiltaket,
 - Basert på fem hovedkilder til informasjon:
 - Data fra forslagsstillerne bak prosjektet
 - Data fra tidligere erfaringer med liknende tiltak presentert i litteratur eller i andre kilder
 - Befolkningsdata og annen sentral statistikk
 - Dokumenter og sekundære kilder
 - Feltundersøkelser, inkludert intervjuer, høringer, gruppemøter og undersøkelser blant befolkningen generelt
6. Forutse responser på effekter – Vurder viktigheten av de identifiserte samfunnsmessige effekter
7. Indirekte og kumulative effekter
 - F.eks. vekst i befolkning eller varehandel
8. Endringer i alternativer
9. Håndtering – utvikle en håndteringsplan. Rekkefølge av tiltak:
 - a) forsøk å unngå negative effekter
 - b) forsøk å minimere de som ikke kan unngås
 - c) kompensere for negative effekter
10. Overvåkning – utvikle et overvåkningsprogram

Tabell 1 Ti trinn i utarbeiding av en vurdering av sosiale effekter (Social Impact Assessment – SIA)

Mange av de virkningene som det vil være viktig å fokusere på i en Social Impact Assessment, er meget vanskelige å kvantifisere på forhånd. Det gjelder for eksempel sosiale virkninger av permanente endringer i sysselsettingen lokalt, av innflytting av arbeidskraft i byggeperioden, eller av økt pendling på grunn av lettere tilgjengelighet til arbeidsplasser utenfor lokalområdet. Heller ikke i ettertid vil kvantifisering av slike virkninger være noen enkel og uproblematisk oppgave. Enda vanskeligere vil det være å prissette slike sosiale konsekvenser. Nytte-kostnadsanalyse nevnes da heller ikke som anbefalt metode for Social Impact Assessment. I stedet anbefales en bred belysning av ulike typer sannsynlige og mulige konsekvenser, jf. ovenfor. Dokumentet understreker at det er bedre å gi omtrent riktige anslag om viktige tema enn å gi nøyaktig riktige anslag om uviktige temaer. Relevans prioriteres altså klart foran målbarhet.

FNs protokoll om strategiske miljøvurderinger

Som ledd i arbeidet med en FN-konvensjon om miljøvurderinger i en grenseoverskridende kontekst ble det i 2003 vedtatt en protokoll om strategiske miljøvurderinger (United Nations, 2003). Protokollen opererer med et bredt miljøbegrep som også omfatter helse. Protokollen definerer ”miljøeffekt, inkludert helseeffekt”, som enhver virkning på miljøet, herunder menneskers helse, flora, fauna, biodiversitet, jordsmonn, klima, luft, vann, landskap, naturarealer, materielle verdier, kulturarv og den gjensidige påvirkningen mellom disse faktorene. Målet med protokollen er å ”bidra til en høy grad av miljøvern (inkl. helsevern) gjennom å sikre miljøhensyn i utviklingen av planer og programmer, bidra til miljøhensyn i utarbeidingen av politikker og lovgivning, etablere klare, transparente og effektive prosedyrer for strategisk miljøvurdering, tilrettelegge for befolkningsdeltakelse i strategiske miljøvurderinger, og på denne måten integrere miljø- og helsehensyn i tiltak og virkemidler for å fremme bærekraftig utvikling” (Article 1).

Strategisk miljøvurdering (Strategic Environmental Assessment – SEA) defineres som ”the evaluation of the likely environmental, including health, effects, which comprises the determination of the scope of an environmental report and its preparation, the carrying out of public participation and consultations, and the taking into account of the environmental report and the results of the public participation and consultations in a plan or programme”. Som det framgår ovenfor, nevnes befolkningsdeltakelse eksplisitt som et viktig mål både i protokollen og som en del av definisjonen av strategisk miljøvurdering. Transparens i prosedyrene understrekes spesielt. Definisjonen omfatter dessuten oppfølgingen av vurderingsrapporten så vel som befolkningsdeltakelsen i de planene og programmene som senere vedtas. Definisjonen omfatter derfor ikke bare fasen med

forberedelse av beslutninger, men også selve beslutningsfasen og innholdet i beslutningene (i form av vedtatte planer eller programmer).

Oppsummering

Innenfor den foreliggende litteraturen om effektvurderinger er det en spenning mellom den delen som anbefaler nytte-kostnadsanalyse som hovedmetode, og den delen som fokuserer på andre vurderings- og sammenlikningsmetoder. Dette gjelder både for analyser på prosjekt- og på konseptnivå. Når det gjelder analyser på konseptnivå, går *The Green Book* i Storbritannia lengst i å anbefale økonomisk verdisetting og nåverdiberegning av effektene. Samtidig anerkjenner denne håndboken at ikke alle relevante effekter kan prissettes, og at det i tillegg til nyttekostnadsanalysene må gjøres en rekke andre analyser for at effektvurderingen skal være fullstendig.

Siden særlig de sosiale, men også mange av de miljømessige effektene som typisk kan opptre som følge av store offentlige investeringsprosjekter, er vanskelige å fastslå med høy grad av presisjon, og enda vanskeligere å prissette, vil nytte-kostnadsanalyse av prosjektenes effekter i beste fall bare kunne fange opp en begrenset del av de relevante effektene.

2.3 Prestasjonsvurdering

Disse metodene har fundament i fagfeltene Total Quality Management og måling. Gjennom de siste tiårene har disse fagfeltene generert kunnskap og erkjennelser som kan utnyttes i utforming av metodikk for ex post-evalueringer av prosjekter.

Behovet for et sett med balanserte måleparametre

Én primær årsak til at det mer ensidige produktivitetsbegrepet ble erstattet av prestasjon, var erkjennelsen av at å måle såpass ensidig som produktivetsdefinisjonene foreskrev hadde mange uheldige effekter (Bredrup, 1995; Sink og Tuttle, 1989; Kaplan og Norton, 1996). Det finnes mye empirisk forskning som dokumenterer at de faktorer og forhold i en virksomhet som det settes fokus på gjennom å foreta målinger, får oppmerksomhet (se for eksempel Sink og Tuttle, 1989; Hronec, 1993). Faktisk kan bruken av måleparametre ha sterke adferdsendrende effekter, i en slik grad at ellers rasjonelle personer gjør disposisjoner de åpenbart må forstå ikke er fornuftige. Eksempler på dette er innkjøpsavdelinger som ukritisk anskaffer deler fra useriøse leverandører fordi

utvikling i innkjøpskostnader er et fokusområde i virksomhetens målesystem eller bedrifter som skjærer sterkt ned på utviklingsaktiviteter fordi målesystemet for tiden har et sterkt kostnadsfokus (Andersen og Fagerhaug, 2001).

Mange forfattere har tatt et kraftig oppgjør med denne kulturen (Kaplan og Johnsen, 1992; Bredrup, 1995; Sink og Tuttle, 1989, osv.). Hovedbudskapet til de fleste alternative prestasjonsmodeller, som gradvis ble lansert for å erstatte produktivitetsbegrepet, var også nettopp at prestasjon måtte måles ved bruk av langt flere dimensjoner. Sterkest uttrykkes dette kanskje i nettopp konseptet balansert målstyring, og tanken er selvsagt at ved å måle flere viktige dimensjoner av prestasjon, styres adferden i virksomhetene til også å fokusere på disse. Måler man både kundetilfredshet, medarbeidertilfredshet, sikkerhetsforhold, miljøpåvirkning, osv. og virkelig setter fokus på disse forholdene gjennom virksomhetens målesystem og bruk av dette, vil ansatte ledelsen og ansatte søke å oppnå gode resultater for alle disse områdene.

Denne argumentasjonen er like gyldig for evaluering av prosjekter. I den grad de evalueringskriterier som vil anvendes når et prosjekt skal evalueres i ettertid er kjente for prosjektet og prosjektledelsen på forhånd, vil dette i større eller mindre grad styre deres adferd. Vet man at budsjettoverholdelse er det aller viktigste for prosjekteier, bevilgende myndigheter og kanskje media, vil fokus være nettopp på å holde budsjettet, ofte med negative konsekvenser for kanskje sikkerhet, kvalitet på det som bygges eller andre forhold. Tilsvarende diskusjon kan føres for ensidig fokus på andre dimensjoner, enten dette er kvalitet eller ”høyverdige” forhold som medarbeidertilfredshet eller estetikk. Tradisjonelt evalueres prosjekter ut fra relativt enkle kostnads-/nyttevurderinger, og dette kan åpenbart føre til et fokus på å holde kostnader lave, all den stund dette normalt er enklere enn å øke nytten. Å evaluere prosjekter med et balansert sett av kriterier, som er kjente på forhånd, er en viktig lærdom som kan avledes fra utviklingen innen prestasjonsmåling.

Behovet for å måle både ”leading indicators” (prestasjonsdrivere) og ”lagging indicators” (resultatindikatorer)

Relatert til diskusjonen om balanserte måleparametre, er spørsmålet om tidsdimensjonen i målinger. En del forskere har sett på konsepter for såkalt ”tidlig varslings”, det vil si muligheter for å oppfatte svake signaler om at fremtidige problemer/kriser er under utvikling (Ansoff, 1975). Ansoffs utgangspunkt var at slike hendelser representerer diskontinuiteter som ikke var en kulminering av en gradvis utvikling, men som likevel ofte kunne spores tilbake til svake, utydelige signaler som kunne detekteres om man var oppmerksomme på dem.

I andre sammenhenger mener man slike overraskende problemer/hendelser i realiteten er et resultat av en trend som har utviklet seg over en viss tid, og som dermed kan identifiseres ved ekstrapolering av målinger av relevante forhold (Davis, 1975; Stallworthy, 1979).

En tredje oppfatning av tidlig varsling finner vi i begrepene *leading* og *lagging* (som kanskje kan oversettes med ledende og forsinkede på norsk) prestasjonsindikatorer, også dette begreper lansert av Kaplan og Norton (1996). I motsetning til troen på at fremtidige problemer/hendelser er et resultat av enten diskontinuiteter i utviklingen eller gradvise trender, bygger dette på en oppfatning av årsaks-og-virkningskjeder. Tanken er at å måle sent i et hendelsesforløp kun gir innsikt i et resultat som allerede er et faktum og ikke kan gjøres noe med, selv om målingene påviser ugunstige resultater. Et typisk eksempel er økonomiske målinger, gjerne månedlig eller kvartalsvis. Disse er et resultat av alle transaksjoner i hele perioden, og gjør det umulig å vite hva som skapte resultatet eller påvirke det i ettertid. Slike målinger er typiske resultatindikatorer som bare kan vise symptomer på at noe er galt eller godt.

Måler man heller kundetilfredshet, er dette ett steg tidligere i hendelseskjeden og et forhold som påvirker det økonomiske resultatet. Ved siden av å gi en tidligere varsling om den kommende utvikling innen økonomien i virksomheten, gjør måling av kundetilfredshet at virksomhetens ledelse og ansatte fokuserer på nettopp å gjøre kundene fornøyde, noe som da vil være en prestasjonsdriver for bedriften. Går man enda lengre tilbake, gjennom de fire prestasjonsperspektivene i balansert målstyring, skjønner man at det vil gi enda bedre fokus på de drivende forretningsprosessene om man måler kvaliteten/prestasjonsnivået i utførelsen av de interne prosessene, som skaper de produkter eller tjenester som er grunnlaget for kundetilfredsheten og derigjennom økonomien. Endelig vil utviklingen av den kunnskap som må til for å utføre disse prosessene godt, være et fundamentalt forhold som kan predikere fremtidig suksess for virksomheten. Igjen gir målinger langs hele denne hendelseskjeden som resultat at alle stadiene får oppmerksomhet og man unngår et kortsiktig fokus.

Det såkalte DPSIR-rammeverket, først utviklet av OECD som SPR, reflekterer også denne tankegangen. DPSIR står for *Driving forces, Pressures, States, Impacts* og *Responses*. Tiltak og prosjekter bør evalueres ut fra disse fem forholdene, og vil da vise sammenhenger og hva drivkrefter som må være tilstede for å skape fremtidige effekter.

Denne kunnskapen kan også utnyttes i *ex post*-evalueringer av prosjekter, i tråd med diskusjonen tidligere om at kjente evalueringskriterier i forkant vil styre adferden til

de som senere skal evalueres. Ved siden av at evalueringskriteriene er balanserte i form av prestasjonsdimensjoner, bør de også være balanserte mellom prestasjonsdrivere og resultatindikatorer. Vet prosjektledelsen at det kun vil benyttes kriterier for tid og budsjett i evalueringen, blir det ofte et kortsiktig fokus på å overholde grenseverdier for disse størrelsene. Er det derimot kjent at forhold som samfunnskontakt, tilfredshet i nærmiljøet med prosjektet, industriell utvikling som et resultat av prosjektet, miljøpåvirkning (positiv eller negativ), osv. vil vurderes, vil det uvergelig sikre at prosjektet også i etablering, planlegging og gjennomføring er mer bevisste disse forholdene. De fleste av disse må sies å være prestasjonsdrivende og mer langsiktige og vil i så måte kunne bidra til bedre prosjekter. Gjennom å sikre at grunnleggende forutsetninger som aksept i nærmiljøet eller at målene med prosjektet er harmoniserte blant interessentene, kan mange vanskeligheter avverges i en tidlig fase og slik kan innsikten i evalueringskriterier på forhånd drive utviklingen av prosjektet. I tillegg gir det prosjektledelsen mulighet for å selv drive løpende prestasjonsmåling av de samme kriteriene, og dermed forsterke denne påvirkningen på adferden.

De psykologiske/motivasjonsmessige effekter av måling

Vi har flere ganger så langt i diskusjonen basert resonnementene på at det som måles i en virksomhet, får fokus og oppmerksomhet og at ledelsen og ansatte ønsker å maksimere de målte prestasjonsnivåene (evt. minimere negative aspekter ved målt prestasjon). Dette er nok et stykke inn på de faglige enemerker til organisasjonspsykologer og andre fagfelt med mer innsikt i menneskelig og organisasjonsmessig adferd enn dem som tradisjonelt har drevet forskningen innen måling.

Snur man på denne observasjonen, er den motsatte effekten at det som ikke omfattes av målinger, ikke prioriteres like høyt eller slett nedvurderes (Andersen og Fagerhaug, 2001). Dette var mye av grunnlaget for at prestasjon fortrenget produktivitetsbegrepet, gjennom at en sterk oppmerksomhet på mer kortsiktige økonomiske forhold relatert til produktivitet førte til at mer langsiktige prestasjonsdrivere ble undervurdert eller helt forkastet.

Det finnes sågar forskning som mener å ha påvist at simpelthen det faktum at det gjennomføres målinger av en aktivitet, påvirker adferden til dem som gjennomfører aktiviteten. Sink og Tuttle (1989) demonstrerte dette med maling av vegger i et rom, der malerne som løpende ble målt på fremdrift og gitt tilbakemelding på hastigheten, var raskere enn dem som ikke fikk slike målinger. Det skal også sies at dette forsøket ikke kontrollerte kvaliteten på malingsarbeidet... For øvrig har mange

sammenliknet måling og de effekter dette gir for motivasjon og adferd med idrett, der de aller fleste, om ikke alle, har ulike mekanismer for å måle og sammenlikne utøvernes prestasjoner, det være seg tid, poenger, lengde, osv.

Lærdommen i forhold til temaet ex post-evaluering av prosjekter er selvsagt at måling synes å påvirke adferd, i positiv eller negativ retning, avhengig av hvordan måleparametrene er definert. Dette støtter dermed den tidligere argumentasjonen om at å definere evalueringskriteriene på forhånd og gjøre disse kjent for prosjektet, enten spesifikt i hvert tilfelle basert på prosjektets ulike målsetninger, eller som generiske kriterier som gjelder for alle prosjekter, kan ha positiv effekt på prestasjonene til prosjektet. Ved å definere balanserte kriterier som også tar hensyn til prestasjonsdrivere for de ønskede resultater og effekter, kan simpelthen eksistensen av kriteriene, til og med uten at det nødvendigvis gjennomføres en ex post-evaluering, være med og gi et bedre prosjekt.

Det nyttigste å vite, er ofte vanskelig å måle

Dette er en annen erkjennelse som er gjort i løpet av de siste par tiårene innenfor fagfeltet prestasjonsmåling. Én av grunnene til at økonomiske forhold ble så sterkt vektlagt i tidligere tilnærminger for produktivitetsmåling var nok ganske sikkert at slike tall var lett tilgjengelige. Faren ved denne tendensen er selvsagt at man i systemer for prestasjonsmåling måler det som faktisk lar seg måle mest mulig objektivt, og som ofte er størrelser innen økonomi, tidsforbruk eller andre konkrete faktorer. Dette trenger ikke nødvendigvis være det som er viktig for virksomhetens måloppnåelse eller langsiktige suksess og dermed bidra til å forskyve fokus i feil retning. Den kjente kvalitetsguruen Deming uttrykte dette ved at i faktabasert ledelse er normalt de viktigste forhold å ha kunnskap de vanskeligste ("unknowables") (Deming, 1982).

Innen måling har man, i takt med at prestasjonsbegrepet har blitt utvidet med nye dimensjoner, også hatt behov for å finne teknikker for å måle forhold av mer subjektiv karakter. Dette kan være forhold som tidligere har vært nevnt, kundetilfredshet, miljøeffekter, osv. Slike størrelser er normalt ikke like lett å måle nøytralt og presist, og løsningen har vært å finne andre måter å innhente slike data på, for eksempel gjennom bruk av spørreskjema, surrogatindikatorer, komposittmålinger, osv.

Dette er igjen lærdommer man kan trekke nytte av i ex post-evalueringer av prosjekter. Omsatt i praksis betyr det at slike evalueringer ikke bare trenger baseres på objektive tallverdier for kostnad, antall brukere eller liknende. Det kan være like

relevant å innhente informasjon gjennom spørreundersøkelser blant brukere om fornøydhet med prosjektets resultatmål eller andre, mer kvalitative forhold.

Oppnå ”presise” målinger

Det har ofte vist seg at én og samme måleparameter kan få svært ulik nytteverdi avhengig av hvordan den spesifikt defineres, hvor ofte den måles, for hvilken tidsperiode, osv. Innen kvalitetsledelse og prestasjonsmåling har man etter hvert etablert en viss erfaring med hvordan slike målinger bør gjennomføres for å bli mest mulig presise og nyttige til bruk i organisasjonene. Noen relevante momenter i forhold til effektivitet av prosjekter kan være:

- Forholdet mellom målefrekvens og -periode og tidspunktet for måling.
- Vurdering av målekostnader.
- Automatisert versus manuell måling/datafangst.
- Sikring av målekvalitet og robusthet.

For det første punktet kan målefrekvens defineres som hvor ofte det foretas målinger, mens måleperiode er tidsperioden med data som inkluderes i målingen (Andersen og Fagerhaug, 2001). En målefrekvens på månedlig datainnsamling med en måleperiode på et kvartal vil dermed si at det en gang i måneden foretas beregnes et gjennomsnitt for de foregående tre månedene.

I en løpende virksomhet er målefrekvens et relevant spørsmål, målinger av tidsforbruk for en aktivitet kan for eksempel gjøres hver time, daglig, månedlig, osv. Frekvensen avhenger typisk av to faktorer, målekostnadene og grad av fluktusjon av det som måles. I effektivitet av prosjekter er dette kanskje ikke like viktig, normalt vil vurderingen gjøres én, i høyden to ganger. Et svært relevant spørsmål er derimot når effektivitet vurderingen foretas. Samtidig har også målefrekvens en viss betydning, da en slik ex post-effektivitet vurdering alltid vil baseres på ulike data samlet inn om prosjektet og dets resultater. For et vegprosjekt vil trafikkvolum være et sentralt kriterium i vurderingen av prosjektet, og trafikkvolum er normalt ikke fullstendig stabilt. Resultatet av vurderingen vil dermed avhenge både av når vurderingen gjennomføres (i forhold til den mer langsiktige utviklingen) og målefrekvensen for trafikkmålingene (dersom vurderingen bygger på flere målinger foretatt på ulike tidspunkter).

Typisk vil realisert nytte av et prosjekt svinge betydelig fra slutfasen av prosjektet og utover i tid. Mot slutten av et prosjekt og umiddelbart etter kan ofte problem

knyttet til driftsetting medføre lavere nytte enn før prosjektet. Etter hvert kommer man over startvanskene og nytten øker, til slutt forhåpentligvis opp mot eller over det forventede nivået. Foretas målingen i fasen med startvansker, kommer selvsagt prosjektet dårlig ut. Senere vil det fremstå bedre, men jo lenger ut i tid etter avsluttet prosjekt målingen gjøres, jo vanskeligere er det å med sikkerhet tilordne effektene kun til prosjektet og ikke andre faktorer.

Uansett når i prosjektets utvikling evaluering gjøres, vil målefrekvensen og -periode for datainnsamlingen ha betydning. Måles trafikkvolum på en veistrekning ukentlig og med en periode på en dag, blir resultatet svært forskjellig om målingene gjøres i helg eller på hverdag. Generelt er noen anbefalinger fra måling i løpende virksomhet det kan være verdt å ta med videre (Andersen og Fagerhaug, 2001):

- Måleperioden bør ikke være kortere enn tiden mellom hver måling, da dette vil gi bare stikkprøver.
- Svært korte måleperioder gjør det vanskelig å se trender da målingene vil fluktuere hele tiden.
- Svært lange måleperioder kan skjule trender som faktisk er til stede da endringer vil ha liten påvirkning på gjennomsnittet.

Som nevnt henger spørsmålet om målefrekvens nært sammen med kostnader innhenting av målingene medfører. Jo høyere frekvens, jo flere målinger. Dette henger selvsagt også sammen med hvordan målingene foretas og kostnadene per måling. Er man avhengig av å bruke intervjuere som foretar utspørringer av et større antall brukere av et prosjekts resultater, er det kanskje bare praktisk mulig å foreta én måling. Kan en automatisk trafikk teller løpende måle trafikkvolum på en bro uten at dette gir særlige kostnader, kan det samles inn data for en lengre periode. Generelt er det et viktig skille mellom målinger som må foretas manuelt og automatisert datainnsamling, selv om det kan være nyanser også mellom disse.

I EU-prosjektene ENAPS (ENAPS-konsortiet, 1999) og APM (APM-konsortiet, 2002) ble det gjort store studier av datainnsamling for prestasjonsvurdering i produksjonsbedrifter og en klar konklusjon var at automatisert måling burde tilstrebes der dette var mulig. Dette ga både lavere målekostnader og i stor grad også høyere kvalitet i datamaterialet. Kvalitetsøkningen skyldes at målemetoden standardiseres og dermed elimineres variasjoner som skyldes at ulike personer foretar målinger på ulike måter, enten målingene er gjort gjennom intervjuer eller telling. Dette gjør dermed datamaterialet mer robust for feilkilder og andre avvik,

selv om automatiserte målinger selvsagt ikke alltid er verken hundre prosent pålitelig eller praktisk mulig å få til.

Målinger som skiller mellom indre og ytre påvirkningsfaktorer

Vi har så langt diskutert både prestasjonsdrivere versus resultatindikatorer og at karakteristika ved prosesser og prosjekter oppviser statistiske fordelinger. Et siste forhold knyttet til hvordan målinger gjøres, er skillet mellom påvirkningsfaktorer for resultater som kan og ikke kan påvirkes i særlig grad. I prestasjonsmåling kommer til uttrykk gjennom at man forsøker å definere måleparametre som er et resultat av hvordan de interne prosessene i en virksomhet gjennomføres og som dermed reflekterer prestasjoner som er et direkte resultat av kvaliteten i det arbeidet som gjennomføres.

Det motsatte vil da selvsagt være forhold som i liten eller bare til en viss grad kan påvirkes av bedriften selv, for eksempel vil salgstall være et resultat av kompliserte sammenhenger mellom generelle økonomiske konjunkturer, konkurrenters produkter, kvaliteten på arbeidet som gjøres av distributører, grossister og detaljistledet, men naturligvis også av kvaliteten på bedriftens eget produkt og derigjennom både prosesser for produktutvikling og produksjon. Problemet med å kun måle slike faktorer, som gjerne er resultatindikatorer på et overordnet nivå, er at det er vanskelig å skille hva som egentlig reflekterer organisasjonens prestasjoner, dermed er det vanskelig å skape motivasjon gjennom slike målinger, vanskelig å se endringer i egne prestasjoner, osv.

Dette henger også delvis sammen med spørsmålet om attribusjon, som generelt i evalueringssammenhenger dreier seg om i hvilken grad de resultater som registreres er en følge av det tiltak som evalueres eller ikke (se for eksempel OECD, 2000 og Samset, 2003). Jo mer overordnet man måler og jo lengre tid som går mellom tiltaket er gjennomført og målingen foretas, jo mer uklar blir sammenhengen mellom evalueringsobjektet og de målte resultatene.

Et generelt mål med evalueringer er at man skal evaluere tiltaket så presist som mulig, i forhold til de spørsmål man stiller i evalueringen. For prosjekter kan man minst se for seg følgende evalueringsspørsmål:

- 1) Ble prosjektet gjennomført på en god måte?
- 2) Var de estimerte kostnader og effekter realistiske?
- 3) Oppnådde prosjektet de estimerte effekter?
- 4) Bidro prosjektet til samfunnsnytte?

Alle disse vil kreve litt ulike tilnæringer til måling av effekter. Spørsmål 1 og 2 evaluerer typisk kvaliteten i prosjektgjennomføringen, og bør primært baseres på målinger som hensyntar indre påvirkningsfaktorer, men utelater forhold som prosjektet ikke har kunnet påvirke. Spørsmål 3 er en kombinasjon av en evaluering av kvaliteten på estimatene for prosjektet og gjennomføringen, mens spørsmål 4 dekker prosjekteiers evne til å velge riktig prosjekt. Spesielt for spørsmål 4 vil det være mer naturlig å trekke inn ytre påvirkningsfaktorer, da valg av prosjekter også bør ta høyde for ekstern utvikling som påvirker verdien av prosjektet.

Eierstyring, ansvarliggjøring og måling

I følge Stame (2006) er det engelske begrepet "Governance" relatert til styringsprosesser 'the process of governing', til forskjell fra selve styringsinstitusjonene. Samset, Berg og Klakegg (2006) viser at "governance regimes" omfatter prosesser og systemer som på vegne av den finansierende part skal sikre vellykkede investeringer. I tråd med Grünfeldt og Jakobsen (2006) velger vi å bruke ordet "eierstyring" for det som på engelsk kalles Governance, Corporate governance eller Good governance, men da eierstyring som prosess. Accountability kan oversettes med "ansvar" eller "ansvarliggjøring". Etablering av systemer for ansvarliggjøring er et vanlig element av eierstyring.

Hvordan henger eierstyring og ansvarliggjøring sammen med måling?

Prestasjonsmåling, og offentliggjøring av måleresultatene er et sentralt element innen ansvarliggjøring, spesielt i land som Storbritannia og USA. Transparens er et sentralt tema for å oppnå avsvareliggjøring. I følge de Lancer Julnes (2006) ble Government Performance and Results Act (GPRA) etablert i USA i 1993 på initiativ av daværende visepresident Gore (<http://www.whitehouse.gov/omb/mgmt-gpra/gplaw2m.html#h5>). Hensikten var å bruke prestasjonsmåling til oppfølging av om målsettingene til de føderale etater ble oppfyllet. GPRA pålegger etatene å etablere og rapportere kvantitative prestasjonsmål. Selv om dette system fokuserer på etater som helhet og programmer, som kan være grupper av prosjekter, så innebærer det at utfallet av etatenes prosjekter også følges opp.

2.4 Benchmarking som virkemiddel i prestasjonsvurdering

Benchmarking er i seg selv et omfattende konsept, og vi skal ikke bruke mye tid på generelle beskrivelser her. Kort fortalt defineres substantivet benchmark som "et på forhånd definert punkt, brukt som en referanse for målinger". I dette ligger at sammenlikning er et nøkkelord for benchmarking. Dette kan ta mange former og

verktøyet brukes på mange ulike måter i forskjellige sammenhenger, alt fra å sammenlikne resultater oppnådd av ulike filialer i landsdekkende banker til å rangere karakterer fra grunnskoler til å lære fra virksomheter som har funnet nye måter å utføre oppgaver på.

I forbindelse med prestasjonsvurderinger har benchmarking en viktig funksjon i å gi målinger et referansepunkt (se for eksempel Camp, 1989). En virksomhet som måler kundetilfredsstillelse kan ha resultater som varierer mellom for eksempel 92 og 94 %, og være fornøyd med det. Ved å måle dette løpende, kan bedriften se om dette måltallet utvikler seg over tid, i positiv eller negativ retning, og ta aksjon i henhold til dette. Målingen blir derimot enda mer verdifull om man har innsikt i tilsvarende tall for andre virksomheter, gjerne konkurrenter. Ligger den argeste konkurrenten i snitt på tall mellom 97 og 98 %, gir dette en helt annen betydning for egne resultater. Slik benchmarking setter ellers isolerte målinger inn i en sammenheng og på en "skala".

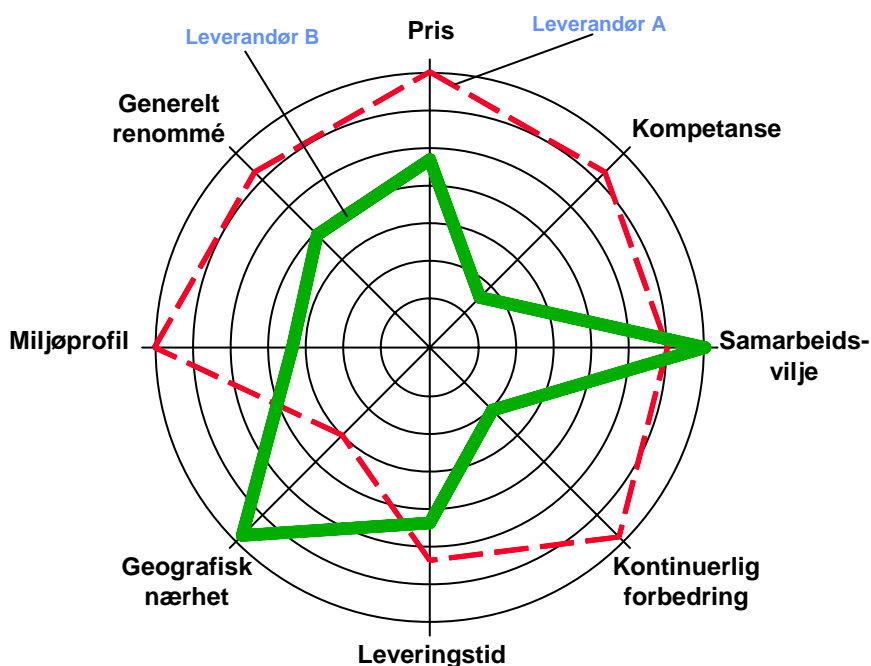
Dette er også høyst relevant for prosjekter. En ex post-vurdering av ett enkelt prosjekt vil kunne avdekke tallverdier for forholdet mellom kostnad og nytte, trafikkvolum, brukertilfredshet, osv. Isolert sett gir selvsagt disse en viss innsikt i prosjektets effekter. Ytterligere informasjon fås ved å samholde disse med de opprinnelige estimater, dermed kan man vurdere om estimatene var realistiske og om prosjektet evnet å nå disse. Virkelig interessant blir evalueringen om man kan sammenlikne resultatene med en serie andre prosjekter. Da vil man kanskje finne at prosjektet både utvilsomt bidro til samfunnsnytte og mer enn overgikk de opprinnelige estimater, men like fullt ga halvparten så mye "nytte per krone" som et liknende prosjekt gjennomført et annet sted i landet. Årsakene til dette kan være mange, alt fra beskjedne ambisjoner, dårlig gjennomført prosjekt, feilaktig valgt prosjekt eller lokale forhold som begrenser den potensielle nytten av prosjektet. Uansett gir en slik benchmarking økt innsikt og grunnlag for å vurdere prosjektet enn om det evalueres helt isolert.

I benchmarking er det utviklet flere analysemetoder for å tolke resultater fra benchmarking, fra enkle gapanalyser i tabell- (se for eksempel Spendolini, 1992) eller diagramform (for eksempel Andersen og Pettersen, 1995) til avanserte kvantitative teknikker som for eksempel datainnhyllingsanalyse. Felles for disse er at de forsøker å tolke benchmarkingdata for å konkludere med hva som er beste resultat, hvor store er avvikene, osv., selv om tilnærmingene er nokså ulike. To typiske former for enklere teknikker er vist under. En ren tabellarisk fremstilling brukes for å sammenlikne en del nøkkeltall for de objekter som sammenliknes, og kan gi en viss innsikt i likheter og forskjeller (se Figur 4). Ulempen er, spesielt ved et større antall objekter eller måleparametre, at slike tabeller fort blir uoversiktlige.

Prosjekt \ Måleparameter	A	B	C	D
Budsjettoverskridelse	11%	13%	-2%	48%
Tidsforsinkelse	3 uker	-4 mnd	1 mnd	3 mnd
Antall søksmål	2	4	Ingen	1
Trafikkvolum i tunnel	78.000	11.000	2.000	148.000
% av opprinnelig estimat	78%	83%	95%	62%

Figur 4 Sammenlikningsmatrise for benchmarkingpartnere

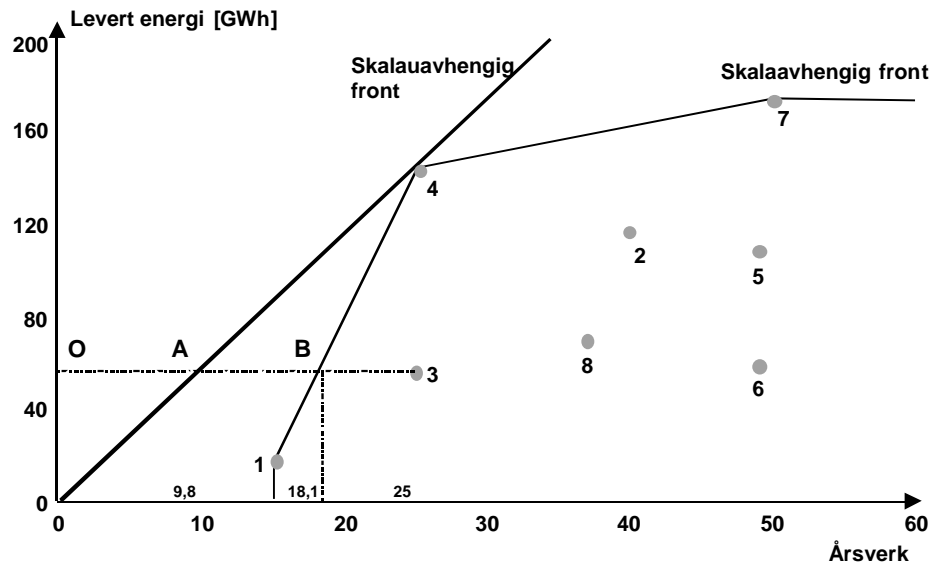
Styrken til grafiske verktøy er forbedret oversiktighet, og spesielt såkalte radardiagram er mye brukt i benchmarkinganalyser. Som Figur 5 viser (diagrammet viser en benchmarking av ulike potensielle leverandører for en kontrakt), så tillater en slik fremstilling raskere tolkning av hvilket objekt som har høyest prestasjon og hvor gapene opptrer.



Figur 5 Radardiagram for benchmarkinganalyse (der beste prestasjon ligger lengst fra sentrum av diagrammet)

Selv om dette, og andre typer diagrammer, gir en oversikt over de objekter som sammenliknes, er dette fortsatt en relativt enkel analyse. For eksempel sammenliknes prosjekter/leverandører direkte mot hverandre, uten å ta hensyn til ulikheter i

størrelse. Det finnes flere teknikker som tar slike forhold med i vurderingen, og såkalt datainnhyllingsanalyse (engelsk data envelopment analysis, DEA) er mye brukt i benchmarking (Thanassoulis (2000) viser for eksempel bruken av metoden for vannforsyning). Metoden forklares for øvrig enklest med et eksempel, som vist i Figur 6:



Figur 6 Datainnhyllingsanalyse (DEA)

Her er åtte kraftverk sammenliknet i forhold til effektivitet i energiproduksjon, representert ved forholdstallet årsverk som kreves for å produsere hver GWh. Disse åtte er plottet i et diagram, representert ved sirklene 1-8. Linje for skalauavhengig front viser at totalt sett er kraftverk 4 mest effektivt gjennom å levere høyest mengde energi per årsverk. Ved å også tegne innhyllingskurven med skalaavhengig front, tas det hensyn til kraftverkernes størrelse. Innenfor segmentene for små, mellomstore og store kraftverk er nå henholdsvis kraftverk 1, 4 og 7 de mest effektive. På denne måten justeres konklusjonene for eventuelle stordriftsfordeler og andre forhold som gjør direkte sammenlikning på tvers av størrelseskategorier mindre relevant. Diagrammet kan også anvendes til å beregne gapet opp til de mest effektive for de andre kraftverkene, som her for nummer 3. Med cirka 60 GWh kraft produsert med 25 årsverk er det ikke av de beste i denne gruppen. Den horisontale linjen viser gapet til fronten for beste prestasjoner, hvor vi ser at for å være mest effektiv blant de små kraftverkene måtte dette ha produsert samme mengde energi med 18,1 årsverk. For å være best totalt skulle tallet vært 9,8 årsverk.

Uansett hvordan data analyseres i forbindelse med gjennomføring av en ex post-vurdering, forutsetter bruk av benchmarking tilgang på data. I noen grad finnes dette, både i form av offentlig publiserte evalueringer, i dokumenter innad i ulike etater og fra andre kilder. Vedrørende kostnader finnes flere erfaringsdatabaser, inkludert Jernbaneverket, Statsbygg, Vegvesenet. Vurdering av tidligere kostnader, referansesjekk, er også en del av KS2-regimet, der det er utført for eksempelvis nytt Operabygg. Det er mindre tradisjon for erfaringsbaser på ”nyttensiden”. For transportprosjekter har flere land (inkludert Sverige, Norge, Storbritannia) har pålagt rapportering av utvalgte parametere, og hadde kunnet etablert erfaringsbaser.

Det eksisterer så vidt oss bekjent ikke noen systematisk database med data generert gjennom ex post-evalueringer for offentlige prosjekter (i privat sektor finnes det nok mer av dette, blant annet gjennom IPA, som omtales i kapittel 3.6). I den grad det i Norge vil standardiseres krav og praksis for gjennomføring av evalueringer av store prosjekter etter ferdigstillelse, bør det vurderes om ikke det også skal etableres en slik database. Dette vil da tillate også mer systematisk benchmarking ved gjennomføringen av evalueringene og kan dermed løfte nivået på evalueringene utover øvelser for enkeltprosjekter. Spørsmålet om en slik database kommer vi nærmere tilbake til senere i rapporten.

Én bruk av benchmarking på denne måten er for øvrig svært likt det Flyvbjerg, Holm og Buhl (2005) omtaler som referanseprognosemetoden i store transportrelaterte prosjekter. Referanseprognosemetoden innebærer at man sammenligner det aktuelle prosjektet med en gruppe av tilsvarende prosjekter og utnytter foreliggende informasjon om disse. Metoden består av følgende tre trinn i vurderingen av et enkeltstående prosjekt:

- 1) Identifisering av en relevant gruppe av tidligere prosjekter. Antallet bør være stort nok til å muliggjøre statistiske analyser, men prosjektene må være sammenlignbare med det prosjektet som skal vurderes.
- 2) Etablering av en sannsynlighetsfordeling for den valgte prosjektgruppen. Dette krever adgang til troverdige data for et tilstrekkelig antall prosjekter.
- 3) Sammenligning av det vurderte prosjektet med fordelingen innenfor hele gruppen av utførte prosjekter, med sikte på å finne det mest sannsynlige resultatet for det vurderte prosjektet.

Referanseprognosemetoden er en form for strukturert kvantitativ benchmarking av forutsetningene for et prosjekt. Fremgangsmåten forutsetter tilgang til data for et forholdsvis stort antall prosjekter. Merk at metoden forutsetter fokus på sammenligning mellom før- og ettersituasjonen, og i mindre grad sammenligning

mellom planlagt situasjon og faktisk etter utført prosjekt. Igjen er dette en metode som tas med videre i arbeidet, som et mulig metodisk element i vår anbefaling.

Felles for både benchmarking og referanseprognosemetoden er kravet om en database av et visst omfang, noe som dermed gjør dette spørsmålet enda mer relevant. Dette krever i sin tur en grad av standardisering av evalueringskriterier/måleparametre, ellers mister man mye av muligheten for sammenligning og benchmarking. Her har for eksempel forskningsprogrammet IMEC gjort arbeid, hvor følgende kriterier ble benyttet ved benchmarking av seksti store prosjekter verden rundt (Miller og Lessard, 2001):

- Teknisk funksjonalitet og effektivitet
- Økonomisk effektivitet
- Sosial aksept
- Politisk aksept
- Miljømessig aksept
- Industriell utvikling

3 Eksisterende praksis for ex post-vurderinger

Vi har så langt forsøkt å konsentrere oss om eksisterende teori som kan være relevant for ex post-vurderinger, selv om det har vært referert til hvordan denne anvendes i praksis i ulike sammenhenger. I dette kapitlet vil vi presentere en mer inngående oversikt over det vi kjenner til av eksisterende ordninger for bruk av effektvurdering internasjonalt og som kan være relevante for ex post-vurderinger av prosjekter. Disse representerer selvsagt ikke nødvendigvis en komplett oversikt over hva som gjøres i verden, men er det vi har oversikt over eller funnet frem til så langt. Hensikten er å vise bredden i hvordan dette gjøres i andre land, som et grunnlag for å bedre forstå hvordan våre anbefalinger senere i rapporten er fremkommet. I neste kapittel gis også en oversikt over tidligere gjennomførte etterkantsvurderinger av prosjekter i Norge.

Søking på internett, i biblioteker og i forskningsdatabaser viser at det alt overveiende av litteraturen om effektvurderinger av store statlige investeringsprosjekter dreier seg om ex ante-vurderinger. I den grad ex post-vurderinger av prosjektenes effekter er blitt utført, dreier det seg særlig om transportinfrastruktur-prosjekter. Imidlertid er ex post-vurderinger heller ikke innenfor denne prosjektkategorien særlig utbredt. Det har for eksempel fram til for ganske få år siden ikke eksistert nasjonale rutiner for ex post-evaluering av transportinfrastrukturinvesteringer – verken på prosjekt- eller programnivå – i noen av EU-landene. I de seneste årene er det tatt enkelte nasjonale initiativer for slike vurderinger.

3.1 Transportetater i amerikanske delstater

Baird og Stammer (2000) beskriver hvordan de delstatlige transportetatene i USA har etablert et system for prestasjonsmåling. Målingene inkluderer oppfølging av effektmål både på transportsystemene (reisetid, ulykker, punktlighet, osv.) og mer overordnede målsettinger som livskvalitet, regional utvikling og miljøeffekter.

Washington State utførte i 2004 en sammenstilling av hva de oppfattet som best practice hos andre delstaters transportetater (Washington State Department of Transportation, 2006). Sammenstillingen viser at flere delstater dels oppsummerer den generelle utviklingen i årsrapporter, dels beskriver status for utførte, pågående

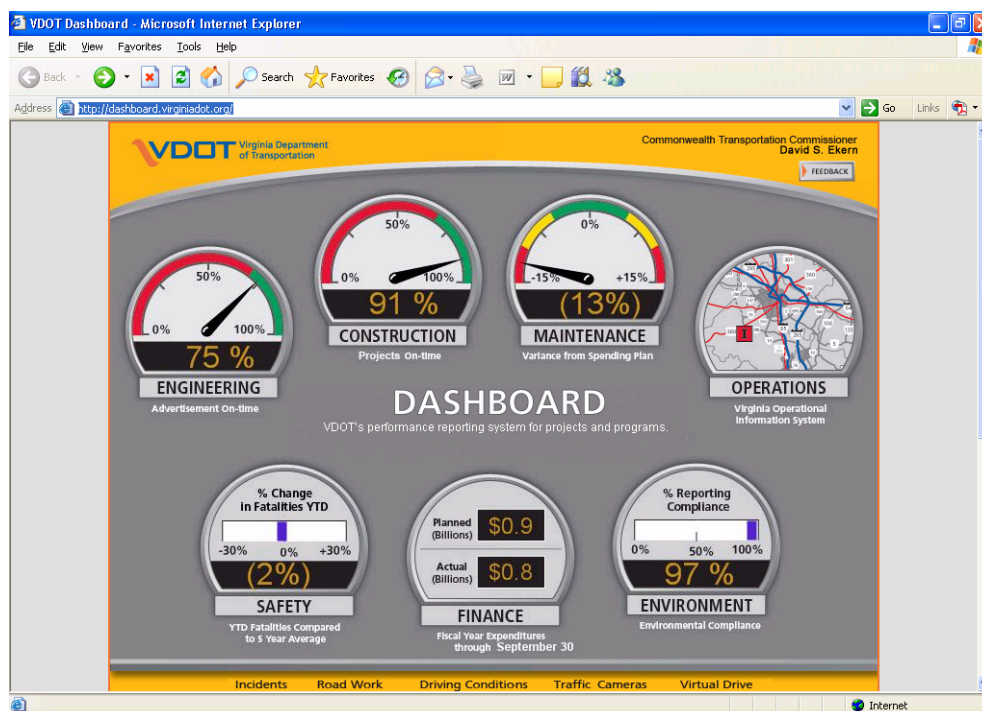
og planlagte prosjekter. Prosjektoversiktene er fokusert på tid, kostnad og ulemper i byggetiden. Effekten følges i første rekke opp for transportsystemene som helhet, noe som til stor del innebærer at summen av utførte prosjekter følges opp.

Minnesota State Department of Transportation bruker et hierarki for nivå og tidsperspektiv på prestasjonsmålingen (<http://www.dot.state.mn.us/dashboards/>), se Figur 7.



Figur 7 Struktur for måling i Minnesota, kilde: <http://www.dot.state.mn.us/dashboards/plan-n-prog1.html>

Virginia State Department of Transportation har bokstavelig talt et instrumentpanel for oppfølging av etaten (<http://dashboard.virginiadot.org/>), se Figur 8. På et overordnet nivå måles prosjektaktiviteten ved antall (andel i %) prosjekter som er i rute tidsmessig. Det finnes også kartbasert oppfølging av trafikksituasjonen med stengte kjørefelter på motorveier, ulykker med mer.



Figur 8 Dashboard fra Virginia <http://dashboard.virginiadot.org/>

Gosling (2000) henviser til en konferanse der prestasjonsmåling i transport var temaet. Det var stor enighet om noen områder, inkludert:

- Man bør måle oppnådde resultater (outcomes) heller enn utført transportarbeid (outputs)
- Prestasjonsmåling har en utvikling fra å relatere seg til service standarder til å fokusere på kundetilfredshet
- Prestasjonsmåling skal gi relevant beslutningsgrunnlag til strategi- og beslutningsprosesser
- Prosessen å utvikle prestasjonsmålesystemer krever en etablering av en konsensus blant involverte parter som er ressurskrevende, men viser seg verdifull i forhold til implementering og akseptens

I følge Gosling (2000) var det også noen områder der det var tydelig at det ikke var enighet, inkludert:

- Om prestasjonsmålingen skal basere seg på et sett med indikatorer som er felles for alle aktuelle transportslag eller bli mer skreddersydd til de ulike etatene

- Avveining mellom å bruke tilgjengelige data versus å fokusere på det man egentlig skulle ville måle, men ikke har mulighet til i nåsituasjonen

3.2 SIKAs oppfølging av transportpolitiske mål i Sverige.

SIKA (Statens Institut för Kommunikationsanalys) har siden 1999 utført analyser av hvordan de transportpolitiske målene blir oppfylt innen hele transportsektoren i Sverige (SIKA, 2003 og 2004a). Oppfølgingen gjøres ikke prosjektspesifikt, men basert på definerte politiske målsettinger innen hele transportsektoren. SIKA (2003 og 2004) mener at det generelt sett er mulig å følge utviklingen av de transportpolitiske målene, men at underlaget for en slik oppfølging er mangelfullt. Oppfølging kan kun gjøre for avgrensede områder. Spesielt mangler samfunnsøkonomisk beslutningsgrunnlag for flere viktige oppfølgingskategorier. SIKA (2004b) har også laget forslag til indikatorer for oppfølging av de transportpolitiske målene.

Målsettinger innen følgende områder følges opp av SIKA:

- Tilgjengelighet
- Likestilling
- Transportkvalitet
- Trafikksikkerhet
- Miljø
- Regional utvikling

3.3 Sound Transit

Sound Transit er et offentlig organ opprettet for å planlegge, bygge og drive regionale kollektivtransportssystemer og -tjenester i Puget Sound-området, som er regionen omkring Seattle på USAs nordvestkyst. Denne administrasjonen har utarbeidet et sett med vurderingskriterier til bruk i beslutninger om valg mellom ulike kollektivtransportløsninger (transit technologies). Det er altså snakk om en vurdering av løsninger på konseptnivå, ettersom både buss, sporvogn, metro, forstadsbane m.m. vurderes opp mot hverandre. (Derimot vurderes ikke kollektivsatsning opp mot vegbygging og privat biltrafikk som alternativ.) Den anbefalte metoden er å vurdere ulike alternativer i forhold til følgende kriterier: Framtidig passasjerbelegg, kapitalkostnader, driftskostnader, reisetid og punktlighet, mobilitet

og sammenknytting av aktuelle start- og målpunkter for reiser, sammenheng med øvrige deler av kollektivtransportsystemet, arealbruk og utbygging, kundenes opplevelse av tilbudet, risiko, miljøfordeler. Det er ikke satt opp noen vektning mellom disse kriteriene, og heller ikke angitt noe spesifisering av hvordan alternativenes virkninger skal måles. I stedet slås det fast at noen av vurderingskriteriene vil være ”ganske kvantitative”, mens andre vil være ”mer subjektive, relatert til politiske nøkkelbeslutninger og hvordan vi ønsker å håndtere vekst [i trafikken]”.

Administrasjonen i Sound Transit fører en løpende dialog med det politisk valgte styret og et særskilt oppnevnt ”ekspertvurderingspanel” om den konkrete bruken av disse kriteriene (se for eksempel Sound Transit Expert Review Panel, 2005). Blant annet har ekspertpanelet kommet med kommentarer og forslag til forbedringer i måten framtidige passasjertall anslås på. Panelet har også generelt etterlyst en større grad av kvantifisering av effekter enn administrasjonen i utgangspunktet hadde lagt opp til. Denne typen dialog med mer uavhengig (?) ekspertise kan bidra til større åpenhet om og robusthet i vurderingene, ettersom de som utarbeider forutsigelsene om alternativenes effekter blir ”kikkert i kortene” av andre fagfolk. Både den konkrete kritikken og korrigeringen de eksterne fagfolkene eventuelt bidrar med, og vissheten om at man vil bli utsatt for ekstern kommentering og kontroll, vil kunne bidra til å motvirke eventuelle fristelser til å lage ”biased” forutsigelser som stiller foreslåtte prosjekter i et bedre lys enn det reelt er grunnlag for. På den annen side kan det, hvis ekspertpanelet og styret har sterke normative oppfatninger for eksempel om ønskeligheten av en sterkest mulig utbygging av kollektivtransporten i regionen, ligge en risiko for at panelets og styrets påvirkning kan bidra til overoptimistiske og fordreide vurderingen. Verdien av dialogen mellom planleggerne og ekspertpanelet (og styret) vil derfor være sterkt avhengig av at panelet representerer en ekspertise som reelt er uavhengig (og at styret avstår fra å legge press om å få inn bestemte, politisk fordreide analyseforutsetninger).

3.4 Tennessee Rail Systems

Dette er en del av delstatsadministrasjonen i Tennessee i USA og har utarbeidet en manual med retningslinjer for ”strategisk prosjektevaluering” (Tennessee Rail Systems, 2001). Begrepsbruken ”strategisk prosjektevaluering” viser – i likhet med håndbokens øvrige innhold – at det er effektvurderinger i den tidlige fasen av et jernbaneprosjekt retningslinjene dreier seg om³.

³ I likhet med prosjektevalueringene i regi av Sound Transit er det bare løsningskonsepter innenfor kollektiv transport som vurderes, og fokuset er dessuten konseptuelt smalere, idet det bare er jernbanealternativer som

Nivå 1	Økonomisk effekt
	<ul style="list-style-type: none"> • Nye midler inn i den lokale økonomien <ul style="list-style-type: none"> – Nye kapitalinvesteringer – Nye driftsmidler
Nivå 2	Sysselsetting
	<ul style="list-style-type: none"> • Direkte sysselsetting som følge av prosjektet • Indirekte sysselsetting som et resultat av prosjektet
Nivå 3	Skatteeffekt
	<ul style="list-style-type: none"> • Lokal eiendomsskatt
Nivå 4	Driftsmessige, sosio-økonomiske og miljømessige effekter
	<ul style="list-style-type: none"> • Driftseffektivitet <ul style="list-style-type: none"> – Effektivitetsforbedring – Besparelser i reisetid – Kostnadsbesparelser for transportører – Utvidelsesmulighet • Sosioøkonomisk/miljømessig <ul style="list-style-type: none"> – Evne til å betjene næringslivet – Konsistens med regionale utviklingsplaner – Tilpasning til lokalsamfunnet – Sikkerhet for befolkningen – Ytre miljø – Støtte fra befolkningen og etater
Nivå 5	Avkastning på investert offentlig kapital
	<ul style="list-style-type: none"> • Offentlig og privat investert kapital • Investeringer som følge av prosjektet • Reduksjon i vogntogtrafikk • Besparelser i drivstoff • Effekt på transportkostnader • Kostnader for drift av jernbane • Effekter på reisetid • Reduserte kostnader til vedlikehold av veier • Reduserte kostnader knyttet til ulykker • Reduserte kostnader knyttet til kø • Statlige avgifter og inntekter • Andre kostnadsbesparelser

Tabell 2 Tennessee-modellen for evaluering av prosjekter

inngår i vurderingene (mens for eksempel bussalternativer ikke er inkludert). Det kan derfor diskuteres om det egentlig er vesensforskjellige løsningskonsepter som vurderes i Tennessee Rail Systems analyser. På den annen side ser analysene ikke ut til å dreie seg bare om å identifisere ulike traseer mellom gitte start- og slutt punkter, men om å vurdere jernbaneprosjekter i ulike transportkorridorer opp mot hverandre.

De vurderte prosjektene evalueres etter Tennessee-modellen opp mot fem hovedkriterier eller nivåer: Økonomisk virkning, jobbskaping, virkning på lokal skatteinntang, driftsmessige og sosioøkonomisk/miljømessige effekter, og avkastning på de statlige kapitalinvesteringene. Hvert av disse hovedkriteriene er inndelt i delkriterier, flest for nivå 4 og 5, som vist i Tabell 2:

En rekke av de økonomisk relaterte virkningene (både under nivå 1, 3 og 5) beregnes ved å multiplisere antatte fysiske virkninger med enhetspriser. Disse hentes fra manualer utgitt av andre myndigheter. Håndboken angir detaljert hvilke kilder som skal brukes i forbindelse med kvantifisering og verdsettingen av hver konsekvenskategori. Bruken av enhetspriser for forventede fysiske virkninger gjelder for eksempel priser for trafikkulykker og for spart tid. Som en rekke forfattere har påpekt (bl.a. Lohmann, 1997; Ackerman & Heinzerling, 2004; Næss, 2006), knytter det seg metodologiske og vitenskapsteoretiske problemer til denne type prissetting. I Tennesseees prosjektevalueringssystem vil disse prisene være eksogene faktorer som tas for "face value". I den grad disse enhetsprisene er de som generelt aksepteres i administrasjonen på tvers av sektorer og partipolitiske skillelinjer, kan det være en pragmatisk rimelig løsning – hvis man i det hele tatt mener det er nødvendig å gi seg inn på øvelsen med å sette prislapp på virkninger som ikke umiddelbart kan måles i pengeverdi og hvor det ikke finnes markedspriser å ta utgangspunkt i..

Ikke-kvantifiserte og ikke-verdsatte virkninger (først og fremst effektene under nivå 4) vurderes i forhold til en 5-trinns karakterskala med følgende trinn:

- 1) Uakseptabel ytelse eller negativ effekt
- 2) Dårlig, men akseptabel ytelse eller effekt
- 3) Nøytral eller ingen effekt
- 4) Moderat positiv effekt
- 5) Betydelig positiv effekt

For hvert av kriteriene er det utarbeidet korte, verbale beskrivelser av hva som typisk karakteriserer vedkommende karaktertrinn. Også de kvantifiserte og økonomisk verdsatte effektene under nivå 1, 3 og 5 innplasseres på denne skalaen.

Det skjer så en vekting av kriteriene innenfor hvert av de ulike nivåene av analyser, inklusiv under-kriteriene. Den maksimale samlede vektede verdien av alle effektene skal være 100. Det vil si at med 5 som maksimumsskåre for hvert kriterium, skal vektsummen for alle kriteriene være 20. vektningen er basert på politiske prioriteringer fastsatt av delstaten.

Dermed etableres et slags grunnlag for samlet vurdering og sammenlikning av prosjekalternativer, gitt at de politisk fastsatte vektallene virkelig avspeiler de politiske prioriteringene. Det prosjektet som har den høyeste poengsummen, vil i så fall utvetydig være det som bør foretrekkes.

Et blikk på datamaterialet for to konkrete analyser som er foretatt (Tennessee Department of Transportation, 2002) kan imidlertid gi grobunn for en viss tvil om hvor nøyaktig vektene avspeiler prioriteter, og dermed også om grunnlaget for å bruke vektene i multiplikasjoner med alternativenes scores på de ulike kriteriene. For eksempel har hadde hvert under-kriteriene på nivå 4 samme vekt (0,4) i den nevnte analysen. Kriteriene er imidlertid endret litt i de foreliggende retningslinjene, sammenliknet med dem som ble brukt i analyse vi har funnet fra 2002. Det er mulig at vektingene etter hvert som systemet brukes og man vinner erfaring, blir mer sofistikerte. Et grunnleggende problem med denne typen vekt- og poengmetoder er at de for å gi et rasjonelt grunnlag for rangering, forutsetter at ingen kriterier som tillegges politisk betydning, er utelatt, og at politikerne (eller eventuelle øvrige personer som står for vektingen) er ærlige når de angir hvilken vekt de vil tillegge ulike hensyn⁴.

I tillegg til vurderingen på poenggivning og vekter, omfatter vurderingsopplegget i Tennessee også en nytte-kostnadsanalyse av kriteriene under nivå 5. Det framgår ikke klart av materialet hvordan for eksempel diskontering av framtidige effekter utføres. Men siden nytte-kostnadsanalysen ikke omfatter langsiktige miljøvirkninger (og heller ikke de øvrige sosiale og miljømessige effektene under nivå 4), er problemene knyttet til diskontering langt mindre alvorlige enn de ville være hvis også langsiktige og irreversible miljøvirkninger var med i analysen. Nytteløshetsanalysen er med den valgte avgrensingen klart en partiell analyse, som i effektvurderingen inngår som en delvurdering innenfor et hovedrammeverk basert på konsekvensanalyse med poenggivning og vekter.

3.5 Karakterboken i den danske byggebransjen

Byggebransjen i de fleste land har tradisjonelt vært plaget med at det er enkelt å etablere virksomheter og at en rekke useriøse aktører til enhver tid er aktive i markedet. Disse kan det være vanskelig å identifisere og styre unna, og mange byggherrer har fått til dels store problemer i sine prosjekter gjennom bruk av slike

⁴ Det er ikke utenkelig at politikere noen ganger unnlater å lage eksplisitte målformuleringer eller tilkjennegi hvor stor eller liten vekt de reelt legger på visse politiske hensyn, fordi en synliggjøring av slike prioriteringer i noen tilfelle kan tenkes å frastøte velgere.

aktører. Spesielt for offentlige byggherrer, der det kan være vanskelig å argumentere seg bort fra å velge laveste anbud, har dette vært et betydelig problem, siden mange slike aktører ofte bevisst legger sin tilbud lavt.

I Danmark har man forsøkt å motvirke dette, og andre problemer relatert til leverandørenes prestasjonsnivåer, ved å innføre en såkalt karakterbok (Byggeriets Evaluerings Center, 2005). Prinsippet er kort fortalt at alle entreprenører som ønsker å ha oppdrag for offentlige byggherrer, får en karakterbok der oppdragsgiverne etter hvert oppdrag fyller inn sine vurderinger av entreprenøren. Disse vurderingene dekker et antall ulike dimensjoner; tid, kvalitet, arbeidsmiljø og kundetilfredshet. Ved neste tilbud, må oppdatert karakterbok legges ved tilbudet, og gir dermed byggherren innsikt i entreprenørens historikk og prestasjon, som et grunnlag for å unngå å velge aktører som ikke synes å holde mål. Ordningen forvaltes for øvrig av Byggeriets Evaluerings Center. Dette må kunne sees på som en form for ex post-evaluering i prosjekter, dog bare av leverandører enkeltvis, og bygger i stor grad på prinsipper fra prestasjonsmåling.

3.6 Independent Project Assessment: Evaluering og benchmarking av prosjekter

Firmaet IPA (Independent Project Analysis) tilbyr tjenester der prosjekter kan evalueres og benchmarkes. Dette er basert på en omfattende database firmaet har bygd opp over flere år, og metodene anvender både målebaserte teknikker for evaluering, statistiske metoder og benchmarking. Metodeverket skal, i følge IPA (2006), kunne anvendes til både konkurrentbenchmarking, vurdering av systemer for prosjektledelse over tid og evaluering av individuelle prosjekter. For evaluering av prosjekter gjelder dette primært ex ante, på et relativt tidlig stadium, og da med fokus på å gi prosjektet anbefalinger for fremtiden.

Tilnærmingen bygger på en antakelse om at utfallet av et prosjekt kan forutses ved å forstå de historiske sammenhengene mellom prosjekt drivere, som typisk er ulike karakteristika, teknologi og praksis for prosjektledelse, og de endelige resultatene fra prosjektet. I databasen finnes det kvantitative, empiriske data fra et stort antall prosjekter (per dato 10.000 prosjekter for bygging av prosessanlegg, 900 prosjekter for utbygging av olje- eller gassfelter og 200 IT-prosjekter) og som kan brukes for å identifisere disse sammenhengene. Ved å samle data om det konkrete prosjektet som skal analyseres, kan det sammenliknes med liknende prosjekter i databasen og konklusjoner trekkes. Eksempler på funn kan være at utviklingen av prosjektplanene

per dato er kommet kortere enn for snittet av vellykkede prosjekter eller at det synes som prosjektet ikke har fanget opp ny teknologiutvikling som kunne gi gevinster.

4 Eksempler på ex post- evalueringer i Norge

Vi har ikke lyktes i å finne noen systematisk gjennomgang av eller samlet oversikt over tidligere utførte ex post-evalueringer av store statlige prosjekter i Norge.

Derimot finnes det en rekke dokumenter av ulike slag som tar for seg ulike sider ved gjennomførte statlige investeringsprosjekter. Vi har gått gjennom blant annet Riksrevisjonens oversikter og stortingsdokumenter fra de siste årene. I dette kapitlet presenterer vi hovedbudskapene fra et utvalg dokumenter vi finner relevante.

Oversikten er ikke komplett. Et fellestrekk er at dokumentene er utarbeidet med ulike formål og derfor har ulike fokus og tar opp ulike tema. Det er videre en tendens til at de omtalte prosjektene er ”problemprosjekter”, og nettopp problemene kan være årsaken til at de er utarbeidet. Det er langt mellom evalueringer av ”vellykkede” investeringsprosjekter, men det gir neppe grunnlag for å trekke en konklusjon om at det er få prosjekter som er vellykkede.

I henhold til Økonomireglementet skal prosjekter evalueres når de er gjennomført. Dersom vi hadde gått ”et trinn ned”, det vil si til departementenes og etatenes arkiver, ville vi trolig ha funnet langt flere evalueringer enn de som har blitt saker på Stortingets bord.

4.1 Prosjekt Nytt Rikshospital

Merknad: Omtalt bl.a. i St.prp. nr. 47 (1993-1994))

Prosjektet for bygging av nytt rikshospital har vært gjenstand for evaluering, og tas derfor med i denne oversikten. Samtidig anvendes dette prosjektet som et eksempel på bruk av vår anbefalte metode i kapittel 6.1, der det også gjengis detaljer fra den allerede gjennomførte evalueringen av prosjektet. Presentasjonen av evalueringen i dette kapitlet er derfor gjort svært kort.

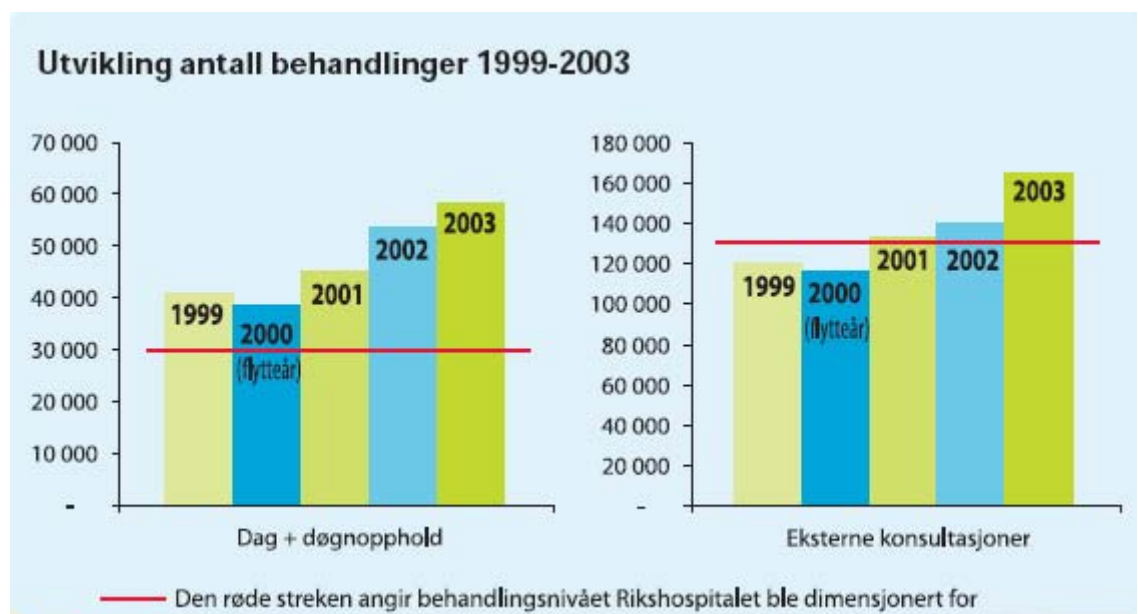
I Riksrevisjonens Dok. nr 3:1 (2003-2004) vises det til at det som oppfølging av denne saken er etablert en ordning med ekstern kvalitetssikring av store statlige byggeprosjekter. Ordningen administreres av Finansdepartementet. Det er stilt krav

om ekstern kvalitetssikring av alle investeringer over 500 mill. kroner før prosjektet fremmes for Stortinget. Kvalitetssikringen vil også gå inn på andre viktige forutsetninger for å sikre god gjennomføring og oppfølging av det enkelte prosjekt.

Byggingen av nytt Rikshospital på Gaustad har vært gjenstand for en evaluering forholdsvis kort tid etter at byggingen ble ferdigstilt (Whist, Bergsland, Jordanger, Ore og Øvrebø, 2001). Evalueringen fokuserte på gjennomføringen av prosjektet.

Rikshospitalet utførte en egen sammenstilling av effektene av det nye sykehuset tre år etter innflytting (Rikshospitalet, 2004). Målinger blir her presentert innenfor følgende områder (se Figur 9):

- Antall vitenskapelig artikler, doktorgrader og patenter
- Antall behandlinger
- Antall dager på ventelister
- Oppfyllelse av pålegg fra helsetilsynet



Figur 9 Effektutvikling Rikshospitalet (Rikshospitalet 2004, side 10)

4.2 Nytteeffekter av IT-prosjekter i Trygdeetaten

Riksrevisjonen la i 2005 fram en undersøkelse av nytteeffekter av IT-prosjekter i Trygdeetaten⁵. Undersøkelsen konkluderer med et Rikstrygdeverket i liten grad kan dokumentere oppnådde nytteeffekter av IT-investeringene. Systemer og rutiner for å sikre gevinster fra IT-prosjekter fungerer ikke tilfredsstillende. Riksrevisjonen peker på mangler ved risikohåndtering, kommunikasjon og oppfølging av IT-prosjekter i trygdeetaten.

Målet med undersøkelsen var å belyse trygdeetatens arbeid med å realisere forventede nytteeffekter fra IT-prosjekter. Riksrevisjonen har gjennomgått fire IT-prosjekter som Rikstrygdeverket hadde store forventninger til. Blant annet ble det undersøkt hvordan trygdeetaten planlegger, følger opp og dokumenterer forventede nytteeffekter fra IT-prosjekter.

Undersøkelsen viser at det er vesentlige svakheter ved mål- og resultatstyringen på IT-området i trygdeetaten. IT-prosjektene i trygdeetaten hadde ambisiøse mål for hvor mye som skulle spares, men disse var lite realistiske. Nytteanslagene ble etter hvert vesentlig nedjustert, men dette skjedde lenge etter at prosjektene ble igangsatt. Klare og realistiske mål er en forutsetning for et forsvarlig beslutningsgrunnlag og god planlegging. Rikstrygdeverket rapporterer generelt om god måloppnåelse av IT-satsningen i årsrapportene til departementet uten at det foreligger dokumentasjon som viser på hvilket grunnlag rapporteringen er gjort.

Riksrevisjonen reiste spørsmål ved om Rikstrygdeverket har gjort tilstrekkelig for å sikre at nytteeffekter fra IT-prosjekter blir realisert som forutsatt. Rikstrygdeverket utarbeider ikke planer for gevinstrealisering, og ansvaret for å realisere nytteeffekter blir i betydelig grad overlatt til ytre trygdeetat.

Undersøkelsen viser at det er mangler ved risikohåndtering, kommunikasjon og oppfølging av IT-prosjekter i trygdeetaten. For eksempel ble det fra hjelpemiddelsentralene i forbindelse med innføringen av en ny IT-løsning etterlyst mer veiledning og bedre tilrettelegging fra Rikstrygdeverket for å kunne ta i bruk systemet.

⁵ Riksrevisjonen Dok. nr 3:2 (2005-2006)

4.3 Troll Olje-prosjektet

ECON analyserte i 2006 Troll Olje-prosjektet⁶ med hensyn til økonomiske virkninger, verdiskaping og teknologiutvikling på oppdrag for Norsk Hydro ASA.

Ved utgangen av 2005 var de samlede drifts- og investeringskostnadene 66 milliarder løpende kroner. I samme periode hadde salget av oljeproduksjonen utgjort vel 204 milliarder kroner.

Da oljeprosjektet ble godkjent av Stortinget i 1992, var det forventet at feltet skulle produsere 350 millioner fat olje. Troll Olje har frem til utgangen av 2005 produsert 1.070 millioner fat olje. Operatøren Norsk Hydro mener at den totale oljeproduksjonen på Trollfeltet kan komme opp mot 2 000 millioner fat olje. Den mangedoblingen av utvinnbare reserver som har skjedd på Troll Olje siden prosjektet ble vedtatt utbygd, kan langt på vei tilskrives teknologiutviklingen i prosjektet. De løpende inntektene fra Troll Olje har ved oversteg ved utgangen av 2005 de løpende kostnadene med rundt 137 milliarder kroner.

De samfunnsmessige virkningene av Troll Olje ble vurdert som betydelige. Prosjektet har gitt store økonomiske overskudd, styrket norske og internasjonale leverandørbedrifter og bidratt til viktige teknologiske forbedringer både innen bore- og kompletteringsteknologi og undervannsproduksjon. Mange av de teknologiske løsningene som ble utviklet i forbindelse med Troll Olje-prosjektet, videreutvikles og tas i bruk på nye felt både i Norge og andre land. Troll Olje er et eksempel på at mye av den verdiskapingen som skjer på norsk sokkel kan tilskrives til teknologi- og kompetanseutvikling. Prosjektet er blitt en referanse for norsk oljeindustri i internasjonale markeder, og har vært en faktor i arbeidet med å vise mulighetene for økt oljeutvinning. Troll Olje har også bidratt til at myndighetene – og oljeselskapene – setter stadig høyere utvinningsmål.

Uten press fra norske myndigheter hadde neppe Troll Olje blitt realisert. Utviklingen av ny teknologi skjedde i et svært tett samarbeid mellom operatøren Norsk Hydro og leverandørbedriftene. Så å si all teknologi- og produktutviklingen i forbindelse med Troll Olje-prosjektet har skjedd i norske leverandørselskap, eller i norske datterselskaper eller filialer av internasjonale leverandørselskap.

De teknologiske utfordringene knyttet til å produsere oljen i Trollfeltet var særlig knyttet til tre forhold.

⁶ ECON-rapport 2006-022

- 1) Oljen befinner seg i et tynt oljeførende lag over hele feltet.
- 2) Det oljeførende laget har en meget stor utstrekning.
- 3) Havdypet er på om lag 340 meter, noe som i det tidsrommet Trollfeltet ble planlagt utbygget, var vurdert å være teknologisk krevende.

Da arbeidet med planene for utviklingen av det enorme gassfeltet startet i begynnelsen av 1980-årene, fantes det ikke teknologi som kunne gjøre det lønnsomt å produsere oljen. For å få økonomi i boreoperasjonene ble det klart at en måtte bore horisontale brønner med stor presisjon og over lange avstander. Boringen måtte skje fra flytere på dypt vann, noe som ikke var gjort tidligere andre steder. Feltet måtte utvikles med få flytende installasjoner og med mange havbunnsinstallasjoner for å få tilstrekkelig lønnsomhet. Utbyggingskonseptet krevde flerfasetransport av store mengder væske over lange avstander.

Lønnsomhetskravene gjorde det nødvendig med teknologiutvikling på flere områder. Analysen identifiserte fem nøkkelteknologier der det foregikk teknologiutvikling i prosjektet:

- Boring og komplettering av brønner
- Reservoarteknologi
- Undervannsløsninger
- Stigerør (risere), flowlines, rensing av produsert vann og flerfasetransport.
- Plattformen

Betydningen av teknologiutviklingen på Troll Olje varierer mellom de fem områdene. På enkelte områder var det helt nødvendig med teknologiske gjennombrudd, mens på andre områder var det behov for forbedringer og tilpasninger av eksisterende teknologi slik at den kunne utnyttes under nye forutsetninger.

4.4 Nitten materiellinvesteringsprosjekter i Forsvaret

Riksrevisjonen la i 2005 fram en undersøkelse av 19 materiellinvesteringsprosjekter i Forsvaret⁷. Formålet med undersøkelsen var å vurdere måloppnåelsen i Forsvarets materiellinvesteringsprosjekter i forhold til styringsparametrene tid, kostnad og kvalitet. Følgende problemstillinger stod sentralt i undersøkelsen:

⁷ Riksrevisjonen Dok. nr 3:16 (2004-2005)

- 1) Hvilke avvik fra planlagt tid, kostnad og kvalitet foreligger ved materiellinvesteringer i Forsvaret?
- 2) Hvilke årsaker kan forklare manglende måloppnåelse i Forsvarets materiellinvesteringsprosjekter?

Utvalget omfatter prosjekter av vesentlig karakter, og er fordelt på alle de tre forsvarsgrenene. Prosjektene i utvalget har en total kostnadsramme på ca. 18 milliarder kroner. Seks prosjekter tilhører Hæren med total kostnadsramme på ca. 9 milliarder kroner. To prosjekter tilhører Luftforsvaret med total kostnadsramme på ca. 3 milliarder kroner, og elleve prosjekter tilhører Sjøforsvaret med total kostnadsramme på ca. 6 milliarder kroner.

Undersøkelsen viser at flere betydelige materiellinvesteringsprosjekter ikke har nådd målene for levering av materiell. Sju av 19 prosjekter er forsinket med mer enn fire år og fire prosjekter har en økning i kostnadsrammen på mer enn 50 prosent. I fire av de ti prosjektene som er undersøkt for kvalitet, er materiellet levert med mangler. Undersøkelsen viser at planlagt leveringstidspunkt er den styringsparameteren som flest prosjekter i undersøkelsen ikke oppnår.

Undersøkelsen viser videre at det i svært liten grad er utarbeidet erfaringsrapporter som grunnlag for erfaringsoverføring.

4.5 Rehabilitering av det kongelige slott, Regjeringens representasjonsbolig og Stiftsgården

I 2001 fikk Stortinget Dokument nr. 3:7 (2000–2001) Riksrevisjonens undersøkelse vedrørende Statsbyggs rehabilitering av Det Kongelige Slott, Regjeringens representasjonsbolig i Parkveien 45 og Stiftsgården i Trondheim⁸.

Formålet med Riksrevisjonens undersøkelse var å vurdere økonomiske og framdriftsmessige forhold ved rehabiliteringen av Det Kongelige Slott, Regjeringens representasjonsbolig i Parkveien 45 og Stiftsgården i Trondheim, samt å belyse årsaker til avvik. Det ble lagt vekt på å undersøke Statsbyggs prosjektstyring, AADs styring av Statsbygg, FINs rolle og de oppdragsgivende departementenes behandling av rehabiliteringsprosjektene.

Undersøkelsen avdekket betydelige kostnadsoverskridelser ved rehabiliteringen av Det Kongelige Slott og Regjeringens representasjonsbolig i Parkveien 45.

⁸ Riksrevisjonen Dok. nr 3:1 (2004-2005)

På bakgrunn av mangler i kostnadsanslagene og regnskapene for arbeidene ved Stiftsgården, var det ikke mulig å fastslå om det var kostnadsoverskridelser i dette prosjektet.

Kostnadsoverskridelsene ved rehabiliteringen av Slottet kunne i stor grad tilskrives at FIN, AAD og Statsbygg undervurderte usikkerheten i prosjektet og organiserte prosjektet slik at det ble en stor grad av parallellitet i prosjektering og bygging.

Kostnadsoverskridelsene på arbeidene ved Parkveien 45 kunne blant annet tilskrives utvidelser i prosjektet og økte administrasjonskostnader og at Statsbygg undervurderte prosjektets kompleksitet, kvalitetskrav, antikvariske krav og prisstigning samt kostnader ved forsinkelser.

Saken ble behandlet i Stortinget 10. mai 2001. Kontroll- og konstitusjonskomiteen viste i sin innstilling blant annet til at vesentlige årsaker til kostnadsoverskridelsene i rehabiliteringsprosjektene har vært mangelfull planlegging, sviktende prosjektstyring og svakheter ved organiseringen av prosjektene. Komiteen merket seg at undersøkelsen avdekket vesentlige svakheter i Statsbyggs faglige arbeid vedrørende utarbeidelsen av forslag til kostnadsramme. Videre merket komiteen seg at verken Finansdepartementet (FIN) eller Arbeids- og administrasjonsdepartementet (AAD) fikk tilstrekkelig informasjon om usikkerheten i prosjektet på et tidlig nok tidspunkt. Komiteen fant det klanderverdig at AAD ikke iverksatte tiltak for å sikre mer realistiske kostnadsrammer.

I etterkant iverksatte AAD og Statsbygg tiltak som gir bedre beslutningsgrunnlag for fastsettelse av prosjektenes kostnadsramme og vil styrke kvaliteten på prosjektstyringen. Samtidig er det iverksatt tiltak for å tilpasse organiseringen og gjennomføringen av prosjektene til de kravene som gjelder for rehabilitering av kulturhistoriske eiendommer.

4.6 Planlegging og oppfølging av store veginvesteringer

Riksrevisjonen la i 2002 fram en undersøkelse av planlegging og oppfølging av store veganlegg i Statens vegvesen⁹. Undersøkelsen pekte på at mangelfulle kostnadsoverslag, sviktende plangrunnlag og omreguleringer er sentrale årsaker til at store vegprosjekter har hatt store kostnadsoverskridelser. Problemene er blitt forsterket av at Statens vegvesen har manglet et egnet verktøy for økonomisk styring av utbyggingsprosjekter.

⁹ Riksrevisjonen Dok. 3:3 (2002-2003)

Målet med revisjonen var å undersøke planprosessen, prosjektoppfølgingen og ressursallokeringen i forbindelse med Statens vegvesens gjennomføring av store vegprosjekter. Riksrevisjonen undersøkte følgende prosjekter: Trekantsambandet, Europavei 16 Aurland – Lærdal, Riksvei 555 Nygårdstangen – Gyldenpris, Europavei 134 Drammen – Mjøndalen og Riksvei 159 Knatten – Vigernes.

Riksrevisjonens undersøkelse viser at detaljerte kostnadsoverslag i enkelte tilfeller synes å mangle helt. I tillegg finnes det eksempler på at reguleringsplaner for deler av et prosjekt ikke var godkjent da prosjektet ble satt i gang. I sine kommentarer sier Samferdselsdepartementet at det er viktig at prosjekter ikke blir fremmet før reguleringsplanen er vedtatt, og at kostnadsoverslag er utarbeidet på bakgrunn av vedtatt plan.

Undersøkelsen avdekket omfattende omprosjekteringer og omreguleringer for flere av prosjektene, noe som delvis har kommet i stand etter initiativ fra berørte kommuner. Riksrevisjonen peker her på viktigheten av god kontakt mellom vegmyndighetene og berørte kommuner ved utbygging av store vegprosjekter. Dette er nødvendig både for å sikre effektive planprosesser og for at andre viktige samfunnshensyn som miljøhensyn, blir ivaretatt.

Statens vegvesen har manglet et egnet verktøy for styring og oppfølging av utbyggingsprosjekter. Dette er blitt vurdert som et stort problem i de fem prosjektene som inngår i undersøkelsen, noe som har vært kjent for Vegdirektoratet i lang tid. Samferdselsdepartementet er enig i at det ikke har vært gode nok styringsverktøy og rapporteringsrutiner i plan- og gjennomføringsperioden for de prosjektene som er blitt gjennomgått, og vil derfor arbeide videre med å utvikle verktøy, styringssystemer og kompetanse.

4.7 Etterprøving av nytte-kostnadsanalyser av samferdselsprosjekter

Det fremgår av Nasjonal transportplan 2006-2015 (St.meld. nr. 24 (2003-2004)) at det ikke er etablert omfattende rutiner for etterundersøkelser av investeringer/investeringsstrategier. Samferdselsdepartementet tok derfor i 2003 initiativ til at det skal gjøres en faglig etterprøving av store norske samferdselsprosjekter, herunder en analyse av i hvilken grad vurderinger og beregninger som gjøres i forkant av beslutninger gjenspeiler virkeligheten etter at prosjektene er tatt i bruk.

Nasjonal transportplan 2006-2015 (St.meld. nr. 24 (2003-2004)) fremhever behovet for å utvikle systemer for å få fram og analysere kunnskap om de realiserede virkningene av investeringspolitikken. Det var ønskelig å dekke både virkninger som kan måles direkte, eksempelvis trafikkmengder, framkommelighet, ulykker og miljøforhold, og virkninger som ikke kan måles direkte, eksempelvis ulike former for regionale effekter. Systemet skulle være sammenliknbart på tvers av sektorer. I den forbindelse ble både Jernbaneverket og Vegdirektoratet bedt om å utvikle og teste metodikk for etterprøving av konsekvenser av store samferdselsanlegg.

Vegdirektoratet har utviklet et opplegg for norske prosjekter basert på erfaringer fra New Zealand (Vegdirektoratet, 2003). Jernbaneverket har også deltatt i utviklingen av metode for etterprøving. Vegdirektoratet (2003) omtaler at ideelt sett burde man etterprøve også null-alternativet som var sammenligningsgrunnlaget ved beslutning, det vil si ikke å utføre det aktuelle prosjektet. Etterprøving av null-alternativet må i tilfelle skje mot en sammenliknbar vegstrekning der ingen tiltak er blitt utført og ingen ytre forhold blitt vesentlig endret. I det forslatte opplegget for etterprøving foreslås at null-alternativet ikke etterprøves særskilt da det er utfordrende både ressurs- og metodemessig. Vegdirektoratet (2003) foreslår at de utføres 2 til 5 år etter ferdigstillelse. Det innebærer at det fortsatt er usikkerhet rundt forventede virkninger i fremtiden. Trafikktall og hastighet oppgis å være forholdsvis enkelt å måle i ettertid. Ulykkeskostnader derimot er beheftet med stor usikkerhet både i estimat i før- og ettersituasjon da datagrunnlaget kan være for lite for å få statistisk holdbare resultater.

De områder som det foreslås å utføre etterprøving av er (Vegdirektoratet, 2003):

- Utbyggingskostnader
- Tidsforbruk
- Trafikktall
- Ulykkeskostnader
- Miljøkostnader (kun dersom de bidrar med mer enn 20% av samlet nytte, eller var et hovedargument for prosjektet)

COWI har på oppdrag av Vegdirektoratet testet metodikken for store veganlegg (COWI 2004).

I Storbritannia introduserte det britiske vegdirektoratet i 2001 det såkalte Post Opening Project Evaluation-programmet (POPE) for å evaluere effekter som følge av gjennomføring av store vegprosjekter. Disse evalueringene har begrenset seg til å omfatte samlet trafikkvolum, reisetid og ulykker, og er utført med forholdsvis beskjeden ressursinnsats (gjennomsnittlig 12.000 pund per prosjekt, tilsv. i underkant av NOK 150.000). I 2004/2005 evaluerte det britiske vegdirektoratet om lag 50 prosjekter i henhold til denne ordningen (Oxera, 2005).

4.8 Feltutbyggingene Åsgard, Visund og Jotun

I 30 av de i alt 32 feltutbyggingsprosjektene som ble gjennomført på 1990-tallet var det kostnadsoverskridelser. Riksrevisjonen har undersøkt de betydelige overskridelsene som har funnet sted i Åsgard-prosjektet, som er den største feltutbyggingen som har vært gjennomført på norsk sokkel. Undersøkelsen som Riksrevisjonen la fram i 2001 omfatter også Visund og Jotun¹⁰.

Undersøkelsen avdekket svakheter i planene og underlagsdokumentasjonen for prosjektene. Riksrevisjonen mente at myndighetene hadde hatt reelle muligheter til å påvirke utbyggingsprosjektene særlig ved behandling av planene. Etter Riksrevisjonens oppfatning hadde departementet et ansvar for å påse at alle vesentlige forhold rundt utbyggingsprosjektene ble fullstendig utredet, og at sentrale premisser for kostnadsberegningene var rimelige og realistiske.

I planene var det ikke gitt en helhetlig framstilling av hvordan rettighetshaverne hadde kommet fram til kostnadsanslagene. Det presenteres samlede kostnadsanslag, og det oppgis enhetsrater for noen elementer, men dette er ikke tilstrekkelig for å kunne etterprøve anslagene på et overordnet nivå. Riksrevisjonen mente at de framlagte planene burde inneholde kostnadsanslag som lar seg etterprøve. Riksrevisjonen pekte også på at det var uklart hvilket ansvar myndighetene faktisk tar ved behandlingen av kostnadsestimaterne i forbindelse med planprosessen og ved godkjenningen av plan for utbygging og drift.

4.9 Tetting av Romeriksporten

I St prp nr 1 Tillegg nr 8 (1998-99), jf Budsjett-innst S nr 13 (1998-99) orienterte Regjeringen Stortinget om status for framdrift og kostnader i arbeidet med tettingen av Romeriksporten. Av proposisjon framgår det at i løpet av sommeren 1998 hadde en ekspertgruppe på oppdrag fra NSB Gardermobanen AS utarbeidet et program for den videre ettertetting av Romeriksporten. På det tidspunkt vurderte NSB

¹⁰ Riksrevisjonen Dok. nr 3:8 (2000-2001)

Gardermobanen AS det slik at hvis ekspertgruppens program ble fulgt fullt ut og en oppnådde de ønskede resultater, ville Romeriksporten kunne åpne for trafikk ca 1. september 1999.

NSB Gardermobanen AS anslo økningen i de samlede ekstrakostnadene ved å gjennomføre dette programmet til 550 mill. kr. De samlede ekstra kostnadene i forbindelse med ettertettingen ville dermed øke fra de tidligere angitte 750 mill. kr til 1,3 mrd. kr. Ekspertgruppens program var lagt opp slik at det kunne gjennomføres trinnvis. Det ble i proposisjonen gitt uttrykk for at det var usikkerhet om hvor omfattende det videre tettearbeidet ville bli.

Effekten av ettertettingen avtok utover året i 1998 slik at redusert innlekkasje pr. investert krone ble mindre etter hvert. Endelig kostnad for ekstraarbeidene var ved utgangen av 1999 på om lag 1,3 mrd kr.

Det ble gjennomført omfattende effektmåling av tettearbeidet knyttet til vannlekkasjene og endringer i grunnforholdene i boligområder.

4.10 Anskaffelse av F-16 kampfly i perioden 1975-98

Riksrevisjonen la i 1999 fram sin erfaringsrapport av anskaffelsen av F-16 kampfly i perioden 1975–98¹¹. Rapporten var basert på en multinasjonal erfaringsrapport utarbeidet av representanter fra riksrevisjonene i Belgia, Danmark, Nederland og Norge i fellesskap. Formålet med rapporten var å samle erfaringer fra dette omfattende prosjektet som kan brukes i forbindelse med eventuelle lignende framtidige prosjekter.

Totalt sett ble det opprinnelige F-16-programmet vurdert som vellykket både sett ut fra en teknisk og en økonomisk side. Den totale prisen har hele tiden holdt seg innenfor den avtalte maksimalprisen, og brukerne av flyet synes å være meget godt fornøyd med dets operative egenskaper.

Det er imidlertid enkelte sider av programmet som etter Riksrevisjonens vurdering ikke synes å ha fungert etter forutsetningene, herunder gjenkjøpsavtalen. Når man i ettertid går tilbake, viser det seg at Norge kom klart dårligst ut av gjenkjøpsavtalen med et gjenkjøp på kun 44,8 %, mot det totale gjennomsnitt på 58,8 %.

¹¹ Riksrevisjonen Dok. 3:7 (1998-99)

Det er videre sannsynlig at man ved å velge tre produksjonslinjer ved produksjonen av F-16-flyene valgte en dyrere løsningen enn om man hadde holdt seg til én produksjonslinje. Delte produksjonslinjer ble imidlertid valgt blant annet for å ivareta gjenkjøp til de europeiske landene. Når man i ettertid ser at Norge kom dårligst ut av gjenkjøpsavtalen, er det grunn til å stille spørsmål ved om den valgte løsning ble vel kostbar for Norge.

Ordningen med endelig avslutning og oppgjør av kontraktene har heller ikke fungert tilfredsstillende. Riksrevisjonen mente at det ikke er akseptabelt at det skal gå 14 år fra siste leveranse av fly og til de siste kontrakter er endelig oppgjort. En slik langvarig kontraktperiode vanskeliggjør både økonomistyring og revisjon.

4.11 Winix

KUFs IT-satsing startet med St.meld. nr. 39 (1983-1984) Datateknologi i skolen. I juni 1984 opprettet departementet et sentralt sekretariat, Datasekretariatet (DS), som en selvstendig enhet direkte underlagt KUF på linje med fagavdelingene. DS skulle gjennomføre det forsøks- og utviklingsarbeidet som ble beskrevet i et fireårig handlingsprogram (1984-1987).

Handlingsprogrammet ble fulgt opp av en handlingsplan som det ble gjort rede for i St.meld. nr. 37 (1987-1988) Om datateknologi i skole og opplæring.

Høsten 1989 var det en oppfatning at en av de utviklede programvarene, Winix, hadde et kommersielt potensial som kunne gi grunnlag for videre IT-satsing i skolen. Den første avtalen om salg og markedsføring av Winix-produktene til en verdi av ca 3 mill. kroner ble inngått med et svensk selskap 1. november 1989. Denne avtalen ble sagt opp av KUF, og en ny avtale som totalt skulle innbringe ca 100 mill. kroner, ble i januar 1991 inngått med et amerikansk firma.

Fra 1. august 1990 ble ansvaret for IT-satsingen fordelt på departementets FoU-seksjon og de sakkyndige rådene. Denne organisasjonsformen varte til 1. august 1992. I januar 1992 etablerte KUF et statlig aksjeselskap, Winix A/S, til å stå for kommersialiseringen av Winix-produktene. I oktober 1993 fattet generalforsamlingen vedtak om å avvike selskapet.

Etter initiativ fra KUF ble IT-satsingen evaluert i flere omganger, av Organization of Economic Cooperation and Development (OECD) høsten 1987, og av stiftelsen International Movement Towards Educational Change (IMTEC) som leverte 16 delrapporter. Evalueringen fra OECD var meget positiv. IMTEC ga i sin rapport nr.

8 bl.a. uttrykk for at handlingsprogrammet hadde en organisasjonsmodell for skolefornyelse som innebar en «systemkonflikt» mellom Datasekretariatet og de sakkyndige rådene med hensyn til ansvars- og arbeidsområder.

KUF oppnådde gjennom sin IT-satsing utvikling av ca 100 pedagogiske programvarer, gjennomføring av forsøk med bruk av pedagogisk programvare innen ulike fag og skoleslag, kompetanseheving blant lærerne, samt spredning og integrering i arbeidet med informasjonsteknologi i skole og opplæring.

Riksrevisjonen uttalte¹² at KUFs IT-satsing ble gjennomført uten at det var stilt tilstrekkelige krav til økonomiske og administrative konsekvensanalyser, andre forberedelser og prosjektstyring som en offentlig investering av denne størrelsesorden burde ha foranlediget. For perioden 1984-1992 ble det disponert samlet ca 420 mill. kroner over KUFs budsjett til informasjonsteknologi. Av dette var medgått ca 150 mill. kroner til utvikling av programvarer, hvorav Winix-programmet kostet ca 88 mill. kroner, inklusiv 15 mill. kroner til aksjekapitalinnskudd i Winix A/S.

4.12 Tress-90

En avtale om konsulentoppdrag i forbindelse med utvikling, test- og pilotfasen i prosjektet Nyskriving TRESS-90 ble inngått mellom Rikstrykdeverket (RTV) og Tryggedata AS (TD) den 7. april 1992. Etter avtalen skulle oppdraget påbegynnes 1. april 1992 og avsluttes 1. mars 1993. Oppdraget skulle utføres etter en fast pris på kr 26.950.000.

I sin gjennomgang av saken pekte Riksrevisjonen¹³ på at flere forhold knyttet til avtalte forutsetninger, leveranser og gjennomføring av prosjektet førte til at det oppsto vesentlige forsinkelser og påløp vesentlige merkostnader i forhold til opprinnelig avtale. Dette var forsinkelser og merkostnader som påløp uten at disse var underlagt de avtalte prosedyrer om tillegg og endringer til avtalen og om forlengelse av frister og dekning av merkostnader.

Høsten 1992 ble det innledet forhandlinger mellom RTV og TD for å komme fram til en avtale om ansvars- og kostnadsfordeling av merutgiftene som hadde påløpt samt å avtale innskjerpede rutiner for å styre krav om tilleggs- og merarbeid. I protokoll av 5. november 1992 ble det inngått avtale om at ny ferdigdato skulle være 1. juni 1993. Videre ble det inngått avtale om fordeling av merutgifter som hadde

¹² Innst. S nr 113 (1994-1995)

¹³ Innst. S nr 190 (1995-1996)

påløpt t.o.m. 1. september 1992 og fordeling av de økonomiske konsekvenser av denne fristforlengelsen. Denne avtale medførte at fastprisavtalen ble utvidet med kr 5.200.000.

I november/desember 1992 gjennomførte SINTEF en ekstern vurdering av prosjektet. SINTEFs hovedkonklusjoner var at prosjektet var ute av kontroll og at det ikke ville være mulig å gjennomføre dette innen den frist som var satt for ny ferdigdato. I desember 1992, da TD varslet om ytterligere forsinkelser, ble det fra RTV stilt krav om at Norsk Informasjonsteknologi AS (NIT), som garantist for TD, skulle sikre gjennomføringen av prosjektet innenfor de avtalte tids- og kostnadsrammer. Da NIT ikke ville gi slike garantier, besluttet RTV å stanse prosjektet fra 12. mars 1993.

4.13 Effektivurdering av alternative utbyggingsprosjekter for Nasjonalmuseet for kunst, arkitektur og design

Gjennomgangen av effektivurderingsmetodene brukt i dette planarbeidet bygger på et notat av Strand (2006). I likhet med Tennesseees retningslinjer for vurdering av jernbaneprosjekter, bygger denne effektivurderingen på en vekt- og poengmetode. Men i motsetning til Tennesseees vektning, er vektningen i analysen av Nasjonalmuseums-alternativene utført av en gruppe bestående av representanter fra departement, Nasjonalmuseet og utredningsteamet fra Møreforskning/metier. Det er derfor ikke snakk om noen representativ gruppe, og det kan stilles alvorlige spørsmål ved hvilken legitimitet en slik selvutnevnt gruppe har til å foreta prioritering på vegne av storsamfunnet mellom ulike verdihensyn. Et annet vesentlig problem i disse analysene er at alternativene som utredes, skiller seg fra hverandre bare mht. bygningsdesign og lokalisering innenfor et snevert geografisk område (Tullinløkka i Oslo og tilstøtende gater). Hvor godt denne lokaliseringen passer inn i forhold til overordnede byplanmessige hensyn (f. eks i forhold til samordnet areal- og transportplanlegging), inngikk ikke i vurderingen, og vesensforskjellige alternative lokaliseringer ble ikke undersøkt.¹⁴

Disse problemene blir påpekt i en kvalitetssikringsrapport Terramar/Asplan Viak utført som ledd i KS1-rutinene. Den sistnevnte rapporten viser også hvor følsomme konklusjonene i den opprinnelige rapporten er i forhold til endrede vektninger mellom kriteriene. Med den ikke-representative vektingsgruppen vil dette være

¹⁴ I forhold til de Rikspolitiske retningslinjene for samordnet areal- og transportplanlegging må imidlertid en lokalisering på Tullinløkka betraktes som gunstig, sammenliknet for eksempel med lokalisering i ytre by eller Oslos nabokommuner.

problematisk, og det er god grunn til å spørre om ikke vekting av kriterier i en slik situasjon fungerer mer tilslørende enn klargjørende.

4.14 Samlet Plan for vassdrag

Et tidlig eksempel på strategisk effektvurdering i Norge er arbeidet med Samlet Plan for vassdrag i første del av 1980-årene. Dette planarbeidet ble gjennomført som respons på kritikk fra miljøorganisasjoner og miljøforvaltningen, som hadde framholdt at vedtak om enkeltstående vassdragsutbygginger medførte risiko for at konfliktfylte prosjekter med store miljøskader ble gjennomført i stedet for alternative utbyggingsprosjekter med færre konflikter og mindre omfattende konsekvenser for natur og miljø. I Samlet Plan-arbeidet ble vurderingene av konsekvenser for naturvern, friluftsliv, kulturminner og regional økonomi utført som en integrert del av planarbeidet, og det kan derfor diskuteres om dette var et eksempel på effektvurdering av planer og programmer eller en effektvurdering i arbeidet med planer og programmer. Strengt tatt representerte vel Samlet Plan-arbeidet det siste. Uansett er metodikken som ble brukt i Samlet Plan, verd å nevne. Utbyggingskostnader ble beregnet i millioner kroner og kraftproduksjonen i GWh/år. For hver av de øvrige hovedkategoriene av konsekvenser (bl.a. naturvern, friluftsliv, kulturminner, landbruk, reindrift, vilt, fiske og regional økonomi) ble inngrepets omfang angitt langs en 4-delt skala. Det berørte områdets verdi i forhold til vedkommende interesse ble også angitt langs en 4-delt skala. Disse vurderingene ble så sammenfattet til poeng på en skala for konsekvens i forhold til hver interesse (9-delt skala¹⁵ fra +4 til -4, med 0 for ingen endringer) og samlet (16-delt skala). Deretter skjedde en ny sammenstilling i en matrise med konsekvensgrad og kraftverksøkonomi som de to dimensjonene, begge målt langs en 4-delt skala. Prioriteringen mellom de vurderte utbyggingsprosjektene skjedde så i henhold til på forhånd oppsatte regler for hvilke verdikombinasjoner i matrisen som innebar første, andre og tredje prioritetsklasse for utbygging, og hvilke prosjekter som overhodet ikke burde realiseres.

Samlet Plan-metodikken er et interessant eksempel på et alternativ til nytte-kostnadsanalyse i effektvurderinger på overordnet, makro-skala plannivå. Metoden som ble brukt i Samlet Plan er imidlertid blitt kritisert for å være uoversiktlig og for å føre til en prioritering av mange små utbygginger som samlet gir større skade på miljø- og kulturverdier enn om man hadde valgt et begrenset antall store, men konfliktfylte utbygginger (Gottschalk & Wenstøp, 1985). Dette fordi vurderingskriteriene innebar at elektrisitetsgevinsten skulle vurderes opp mot

¹⁵ For de fleste variablene ble ikke hele skalaen brukt i praksis.

konfliktgrad i stedet for mot samlet skadeomfang, som Gottschalk og Wenstøp anbefalte.

4.15 Hvilke erfaringer kan vi høste fra tidligere evalueringer?

I lys av mandatet for denne utredningen er de viktigste observasjoner fra de tidligere større evalueringene vi har gjennomgått følgende:

- Evalueringene belyser i liten grad om riktig konsept ble valgt, og heller ikke eventuell bredde i alternative konsepter.
- De fleste evalueringene belyser i relativt liten grad nullalternativet, og i den grad de gjør det, gjøres det ikke på noen systematisk eller standardisert måte.
- De fleste evalueringene gjøres med referanse til ex ante-beregninger for kostnader og eventuelt inntekter for prosjektet, men i liten grad med referanse til ex ante-beregninger av samfunnsmessig kostnad og nytte.
- Evalueringene handler i begrenset grad om samfunnsmessige effekter i et videre perspektiv, fordelingsvirkninger, osv.

Når dette er sagt, er det ikke ment som noen kritikk av de enkelte gjennomgåtte evalueringene, men som en påpekning at evalueringene gjennomgående sier relativt lite om de forhold som vi er opptatt av i denne utredningen, og ikke minst en påpekning av behovet for en systematisk evalueringspraksis.

5 Foreslått prosedyre og metode for ex post-evaluering

Dette kapittelet beskriver vårt forslag til en metode for ex post-evaluering av store, statlige prosjekter, bestående av både en produktevaluering (av leveransene fra prosjektet) og en prosessevaluering (av ex ante analyser og gjennomføring av prosjektet). Vi diskuterer først hensikten med slik evaluering, etterfulgt av hva som vil være sammenlikningsgrunnlaget for evalueringen, når evalueringen bør finne sted og detaljer i selve metoden for evalueringen. Vi vil også diskutere hvilket datagrunnlag som kreves for å gjennomføre en slik evaluering og roller som må fylles i en ordning for systematiske evalueringer. Til slutt vil det vurderes hvilke konsekvenser en slik ordning for ex post-evaluering av offentlige prosjekter vil medføre for ulike aktører.

5.1 Hensikten med ex post-evaluering

Vi startet utredningen av en systematisk ex post-evaluering med utgangspunkt at dette skulle supplere de to etablerte punktene for kvalitetssikring, de såkalte KS1 og KS2. I ex post-evalueringen ville hensikten være å etterprøve effektene av prosjektet, i forhold til de antatte effekter vedtatt for prosjektet.

Etter hvert som diskusjonen pågikk i løpet av utredningen, har vi blitt klar over at etableringen av en slik ordning, nærmest et KS3, vil ha mange ulike hensikter: Noen av disse er direkte, andre mer indirekte som en følge av at en slik ordning gjennomføres systematisk og konsekvent. Vi har identifisert de følgende, og bredden i disse bør tas hensyn til når kostnadene ved en slik ordning vurderes:

- Ex post-evaluering skal sørge for at både kvalitet og relevans for de gjennomførte ex ante-analyser vurderes. Dette innebærer både metodikk, datagrunnlag og anvendelse av disse i prosjektets tidligfase. Målsetningen er at systematiske vurderinger av dette vil være et virkemiddel for å forbedre disse på sikt, gjennom å skape innsikt i hvordan ulike metoder for ex ante-analyse virker.
- Mane til edruelighet og redelighet i ex ante-analyser/-estimer. Dette oppnås primært gjennom at det er kjent på forhånd at det for store prosjekter vil bli gjennomført en ex post-evaluering, en form for ”ris bak

speilet” som gjør taktisk estimering av nytte og kostnader mindre attraktivt (se for eksempel Flyvbjerg et al for studier av såkalt politisk budsjettering). Dette burde også kunne bidra til å høyne den etiske/faglige standarden blant de involverte aktører, gjennom en konsekvent praktisert ordning for ex post-effektvurdering

- Gjennom konsistente vurderinger ex post-fullførelsen av prosjekter vil en slik ordning bidra til å disiplinere og standardisere de ex ante-analyser som gjennomføres og det beslutningsunderlag som legges til grunn for disse.
- En åpenbar hensikt er å vurdere de oppnådde effekter av prosjektet - ble de opprinnelige målsetningene (og eventuelt utvidede) oppnådd. Dette vil legge til rette for en informert samfunnsdebatt om store prosjekter og samfunnets bruk av penger på ulike tiltak.
- Systematiske ex post-evalueringer gir en unik anledning til å samle inn og bygge erfaringsdata om effekter som oppnås gjennom store prosjekter, hvilke avvik fra opprinnelige estimater som observeres, osv.. Dette vil være verdifull informasjon for ex ante-analyser for kommende prosjekter

I sum vil vi understreke at hensikten ikke er å evaluere ulike aktørers innsats i prosjektet, men totalt sett sørge for at det konsekvent gjennomføres helhetlige vurderinger av hele prosjektet og dets prosesser i den hensikt å forbedre disse.

Videre kan man diskutere hva man skal evaluere for å oppnå disse hensiktene med en slik ordning. Noen muligheter er:

- Nasjonens evne til å velge de riktige prosjekter, altså i hvilken grad politikere og forvaltning prioriterer landets begrensede ressurser på de prosjekter som gir høyest nytte for befolkningen.
- Kvaliteten på de gjennomførte ex ante-analysene og de definerte målsetninger for prosjektet.
- Gjennomføringsevnen til prosjektorganisasjonen.

Sistnevnte er i mindre grad relevant, vi mener en slik ordning i hovedsak skal være en vurdering av prosjektenes arbeid i en tidlig fase. Det man ønsker å sette søkelyset på er altså evnen til å velge riktige prosjekter og konsepter, kvaliteten på analyser og beslutningsprosesser i tidligfasen, osv samt de oppnådde effektene av prosjektet, ikke gjennomføringsevnen av prosjektet. Dermed må evalueringen måtte ta hensyn til spørsmål som:

- Var ex ante-analysene anlagt bredt nok, i forhold til å vurdere andre konsepter, andre prosjekter, i forhold til det som i ettertid har vist seg å være

den reelle utviklingen, osv. For eksempel kan prosjektet for kjøp av nye kampfly være for låst i dette som konsept dersom Norge i stedet viser seg å ha mer behov for utstyr for internasjonale operasjoner.

- Ga prosjektet de planlagte effekter? For å vurdere dette må de grunnleggende forutsetningene for prosjektet presiseres i ex ante beslutningsunderlag slik at det ferdige prosjektet kan vurderes opp mot disse.

Disse to hovedspørsmålene representerer hver sin tilnærming i forhold til å evaluere prosess eller produkt, altså de analyser og vurderinger som har vært gjennomført i prosjektet og dets tidlige fase versus de rene effektene av prosjektet. Normalt tar en ex post-evaluering i stor grad mål av seg om å være en produktevaluering av det ferdige prosjektets leveranser, da det er disse samfunnet er opptatt av. Det er dog flere grunner til at det kan være nyttig og relevant å også inkludere elementer av en prosessevaluering i en slik ordning. For det første kan et prosjekt som har gjennomført utmerket håndverk i ulike faser av prosjektet oppleve at resultatene ikke blir som forventet som følge av eksterne forhold prosjektet vanskelig kunne forventes å råde over eller forutse. Like viktig er det at prosessevaluering i større grad muliggjør mange av de nevnte hensikter med ex post-evaluering, rettet mot å forbedre de analyser og vurderinger som gjøres i prosjektene og i forbindelse med fremming av nye prosjekter. En ren effektvurdering vil i mindre grad fange opp forhold som kan gi slike forbedringer i fremtidige prosjekter. Noen sentrale spørsmål som bør reises i en prosessevaluering kan være:

- Hvilke føringer ble satt for ex ante-analysene og konseptvalg?
- Har det skjedd utvikling/endringer i nytte-/kostnadsestimater utover i prosessen eller i design/utforming av konsept? For eksempel ble det politisk vedtatt å bygge et pasienthotell som en del av Rikshospitalet etter at prosjektet var startet, noe som ga store konsekvenser for gjennomføringen av prosjektet.
- Grad av rimelighet i ulike beregninger (nytte, kostnad, osv.) og diskonteringsfaktorer for realisert og alternative konsepter?
- Seriositet i utredningen av alternative konsepter, eller har disse snarere vært pliktlop for å tilfredsstille kravet om slike utredninger?

Uansett hvordan en ordning for ex post-evalueringer utformes er et sentralt poeng at, i likhet med KS1/KS2, er sannsynligvis den aller viktigste effekten at det faktisk eksisterer et forutsigbart regime som gjør at prosjektene vet det vil komme en evaluering i etterkant. Det er ikke derfor bare en teoretisk mulighet for at prosjektet vil gåes etter i sømmene, noe som stort sett bare skjer i virkelige fiasko-prosjekter,

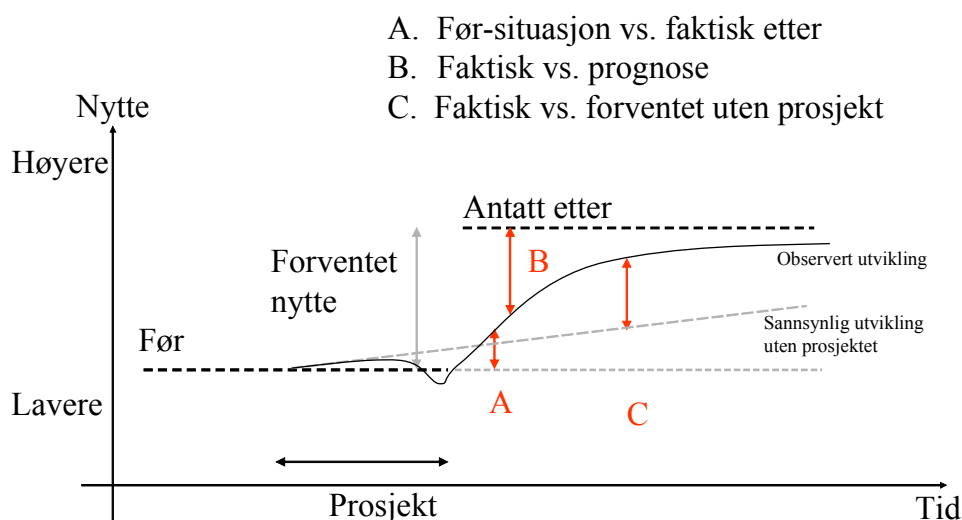
men det er 100% sikkert at det vil gjennomføres en etterkantsvurdering av prosjektet. Selv om hensikten med denne er læring for fremtiden og ikke henge ut noen eller finne syndebukker, må tilstedeværelsen av en slik ordning forventes å ha en oppdragende effekt. Dersom ex post-evalueringer til stadighet avdekker at ulike etater systematisk overvurderer gevinstene ved et prosjekt, undervurderer kostnadene, gjennomfører slette utredninger i tidligfasen eller andre kritikkverdige forhold, vil dette nødvendigvis måtte få konsekvenser. Ingen vil dermed ønske å stille seg i en slik posisjon.

5.2 Referansepunkt og sammenlikningsgrunnlag for ex post-evaluering

Gitt de hensikter med ex post-evaluering som ble diskutert i forrige seksjon, spesielt i forhold til å vurdere effektene oppnådd med prosjektet, er et sentralt spørsmål hva det sammenliknes mot. Her finnes det flere muligheter, som illustrert i Figur 10, og vi anbefaler en kombinasjon av de følgende:

- Det ideelle er at situasjonen etter at prosjektet er gjennomført samholdes med situasjonen man ville ha hatt på det samme tidspunktet uten at prosjektet hadde blitt gjennomført, det såkalte nullalternativet (dette er vist i figuren med situasjon C). Dette evaluerer de reelle effektene av prosjektet, men en slik kontrafaktisk analyse er komplisert, spesielt å predikere situasjonen uten prosjektet, og man kan derfor kanskje ikke kreve for høy presisjon i dette. I nullalternativet må det heller ikke sees utelukkende på situasjonen ved beslutningstidspunktet, men også tas med de fremtidige effektene av allerede vedtatte prosjekter og annen utvikling som ville ha kommet inn i bildet uansett. Selve nullpunktet for ”før-situasjonen” kan ofte være uklart å definere, men vi mener dette må basere seg på tidspunktet som det foreligger data for i Stortingsproposisjonen hvor prosjektet ble lagt frem for beslutning
- Også for spørsmålet om riktig konsept ble valgt, vil det mest riktige være å gjennomføre en kontrafaktisk analyse opp mot hva situasjonen ville blitt om man hadde valgt ett av de andre alternative konseptene, i tillegg til nullalternativet, som allerede nevnt.
- Videre kan man samholde oppnådde effekter med den estimerte nytten slik den ble fremsatt før prosjektet ble vedtatt og igangsatt (vist med situasjon B i figuren). Dette evaluerer hovedsakelig kvaliteten på planleggingsarbeidet.
- Samholdes oppnådd endring av tilstand mot før-situasjonen før prosjektet (situasjon A i figuren), kan dette gi erfaringsdata om typiske endringer som kan oppnås ved denne typen prosjekter. Dette vil være nyttig kunnskap for fremtidig estimeringsarbeid i etterfølgende prosjekter.

- Endelig kan man sammenlikne oppnådde effekter i prosjektet som evalueres med hva andre, liknende prosjekter har oppnådd. Dette sikrer at evalueringen ikke bare forholder seg til den interne estimeringen i dette ene prosjektet, men gir en ekstern referanse gjennom en form for benchmarking. Dette kan dels benyttes ex ante for å vurdere realismen i estimater for effekter og nytte, dels kan det ex post gjennomføres sammenlikning av prosjektet som vurderes mot både andre norske prosjekter og et relevant internasjonalt utvalg. Dette kan avdekke ulikheter i nøkkelparametre (eksempelvis sammenlikne krone per operaforestilling mellom de nye operaene i Oslo, Göteborg og Helsingfors), skape innsikt i prognosefeil som ofte gjøres og selvsagt lære av disse avvikene.



Figur 10 Referansegrunnlag for ex post-evaluering

Vi mener altså at alle disse alternativene bør hensyntas i ex post-evalueringer, om enn i ulik grad avhengig av prosjekt og situasjonen rundt dette.

5.3 Tidspunkt for ex post-evaluering

Vi har så langt diskutert hva man skal sammenlikne med når det gjennomføres ex post-evaluering, men ikke noe om når ex post-evalueringer bør finne sted. Figur 10 viste et mulig forløp for nytteeffektene av et prosjekt i tiden etter dets ferdigstillelse. Ofte ser man faktisk en nedgang i opplevd nytte rett etter at et prosjekt er fullført og ny infrastruktur er tatt i bruk, en fase som kan benevnes ”start-up blur”. Etter en tid oppnås mer av den nytten prosjektet ble etablert for å realisere og nettogevinsten er positiv. Om nytten eller de positive effektene fortsetter å øke i lang tid etter prosjektet, kan det såes tvil om dette skyldes selve prosjektet eller andre forhold.

Dette taler for et kompromiss med hensyn til evalueringstidspunkt:

- For tidlig (innen ett år etter avslutning) vil kunne risikere å evaluere prosjektet mens det enda ikke har realisert nytten.
- For sent (mer enn kanskje 4-5 år etter avslutning) vil evaluere prosjektet på et tidspunkt der man vil kunne oppleve ”effektutvanning” og der attribusjon av de observerte effekter til prosjektet blir vanskelig. Videre kan prosjektorganisasjonen på dette tidspunktet være så spredd at det kan være vanskelig å innhente data og informasjon samtidig som prosjektet selv ikke vil kunne lære noe av evalueringen.
- I tillegg kommer det faktum at prosjekter er svært ulike når det gjelder hvor raskt nytteeffekter materialiserer seg og ikke minst utvikler seg over tid. For noen prosjekter, for eksempel bygging av en fastlandsforbindelse for et øysamfunn, kan den første primæreffekten være bedre tilgang til arbeidsplasser og servicetilbud for beboerne. På noe lengre sikt kan utbygging av turisttilbud komme som en følge av enklere adkomst til øyen, mens i et langt perspektiv kan man finne at øyen ikke lenger har fastboende, bare hytteturister. Man kan altså observere en fasemessig utvikling av nytten av prosjektet, der nytten ikke bare fortsetter å vokse, men endrer karakter og kanskje etter hvert mister mye av sin verdi.
- Et annet moment er levetiden på det objektet/infrastrukturen prosjektet har levert, spesielt i forhold til estimert levetid. Dersom levetiden blir kortere enn forventet, forringes selvsagt den totale nytten fra prosjektet, og vice versa om den blir lengre. Dette vil man faktisk ikke kunne vite før man er klar for utrangering eller oppgradering, og som da kanskje tilsier en siste sluttevaluering på dette tidspunktet.

Alt dette taler egentlig for at tidspunktet for å gjennomføre en ex post-evaluering bør vurderes individuelt for hvert prosjekt, gjerne også om det bør gjennomføres flere evalueringer for prosjekter der nytten varierer betydelig over tid. Samtidig er det ønskelig at en eventuelt standardisert ordning for systematisk effektvurdering av offentlige prosjekter gjør prosjektene sammenliknbare, noe som dels kan oppnås gjennom et fast definert evalueringstidspunkt.

Vårt forslag til kompromiss er derfor at et anbefalt standard tidspunkt for ex post-evaluering settes til å være ved utgangen av andre driftsår for objektet levert av prosjektet. Man har da lagt bak seg eventuelle garantiperioder og oppstartsproblemer, samtidig som effektene ikke burde ha blitt særlig påvirket av annen samfunnsutvikling eller nye prosjekter. Vi anmoder dog også om å utvise nødvendig skjønn, slik at det ikke settes i gang evalueringer etter to år

som man skjønner enten vil gi unøyaktig resultat eller må gjentas på et senere tidspunkt. ”Enklere” prosjekter der nytteeffekter må forventes å inntreffe relativt raskt, kan kanskje evalueres allerede ett år etter ferdigstilling. Andre prosjekter burde kanskje vente så mye som 4-5 år, eller kanskje evalueres etter to år og deretter på nytt etter ti.

I tillegg må det nevnes at kostnadssiden burde være kjent tidligere, i hvertfall innen ett år etter avslutning. Disse data kan gjerne innhentes så snart de foreligger, spesielt siden dette er et viktig datapunkt for analyse av kostnadssiden opp mot anbefalinger ved KS2.

5.4 Metode for ex post-evaluering

Med metode mener vi her mer spesifikt hva som skal evalueres når det gjennomføres etterkantsvurdering av et prosjekt. Det betyr både hvilke parametre som skal vurderes og hvilke momenter det bør tas hensyn til, og et viktig spørsmål blir hvor generell en slik metode kan være. Prosjektene som omfattes av KS-ordningen varierer sterkt, fra veiprosjekter til innkjøp av utstyr til Forsvaret til utvikling av IT-systemer. Vår oppfatning av generaliserbarheten av metoden er at den overordnede metodikken bør kunne være felles for alle sektorer og deres prosjekter. Når det gjelder de mer spesifikke parametre som vurderes i hvert enkelt prosjekt, tror vi disse vil måtte være sektortilpassede.

Videre er vi av en klar oppfatning at evalueringsmetodikken bør tilstrebe å fange opp så brede konsekvenskategorier at den kan dekke sterkt ulike alternative konsepter (for eksempel kan et foreslått prosjekt for utbedring av en vei ha som alternativt konsept en rørledning for transport av bensin, og disse bør kunne vurderes med samme metodikk). Det er også en selvfølge at de konsekvenser som vurderes, må være både positive og negative, slik at det totale bildet for hvert prosjekt etableres. Og som ellers i evalueringer vil denne typen ex post-vurderinger måtte inkludere fordelingskjevheter prosjektet måtte føre til. Til en viss grad vil ex post-evalueringen bestå i å repetere ex ante-analysen med reelle tall, for derigjennom å vurdere arbeidet som ble gjort i prosjektets tidlige fase. I prosjekter der konsepter endres underveis, bør evalueringen vurdere det realiserede konseptet, men ved sammenlikning av konseptalternativer må også det opprinnelige konseptet vurderes.

Et annet sentralt spørsmål er hvilket datamateriale som må være tilgjengelig for de som skal gjennomføre ex post-evalueringen. Ett ytterpunkt er at det ikke stilles materiale tilgjengelig i forkant, men evaluatorene må selv innhente all nødvendig

informasjon og data. Dette vil gjøre evalueringsjobben mer omfattende og komplisert, spesielt siden prosjektorganisasjonen som har gjennomført prosjektet og etaten denne tilhører har mye bedre tilgang på data om og fra prosjektet.

Vår anbefaling er tredelt:

- *For alle prosjekter som omfattes av denne ordningen bør det gjøres løpende målinger fra prosjektavslutning av et fast sett med utvalgte parametre.* De data dette genererer vil danne en viktig del av datagrunnlaget for ex post-evalueringen. Parametrene vil nok måtte være spesifikke for hver sektor/type prosjekt, men bør kunne være felles for et større antall prosjekter innenfor den enkelte sektor. Slike parametre må bygge på nøkkeltall som vil være relativt enkelt tilgjengelige. Dette gjøres for øvrig til en viss grad allerede i dag, for eksempel for veiprosjekter, jernbaneprosjekter, osv. Slike målinger systematisk gjennomført for alle relevante prosjekter vil for øvrig gi et langt bedre datagrunnlag for å studere hvilke effekter som oppnås, når de inntreffer, avvik fra prognoser, osv. Et siste moment er at ved løpende innsamling av slike data opparbeides det også et grunnlag for både benchmarking-evaluering av ferdigstilte prosjekter og ”reference forecasting” ved målsetning av fremtidige prosjekter.
- *I tillegg vil det være en rekke bredere og prosjektspesifikke forhold en ex post-evalueringen vil måtte adressere.* Disse omfatter i større grad spesielt nytte- og effektsiden av prosjektet og vil normalt peke tilbake til de definerte målene med prosjektet, behandling i Stortinget, osv. For disse vil neppe de standard parametre som løpende samles inn være tilstrekkelig som grunnlag for evalueringen. Snarere må det nok innhentes data og informasjon av både kvantitativ og kvalitativ natur i forbindelse med gjennomføring av selve evalueringen.
- *De allerede gjennomførte gjennomgangene av prosjektet ved KS1 og KS2 bør følges mer spesifikt opp.* I den grad KS1- og/eller KS2-gjennomgangene har identifisert spesielle hensyn prosjektet burde forholde seg til, gitt relevante anbefalinger, osv., bør ex post-evalueringen sees på som en form for KS3 der sløyfen lukkes og slike forhold avsjekkes.

Vi har utviklet et forslag til vurderingsparametre for ulike typer prosjekter. Disse er kun forslag til datapunkter det er naturlig å vurdere prosjekter opp mot. Så langt det er mulig bør slike standardparametre være basert på eksisterende parametre i ulike veiledere, der disse holder god nok kvalitet. Vi har forsøkt å gruppere parametrene i noen standard kategorier:

- Investeringskostnader
- Driftskostnader

- Konsekvenser av prosjektet, forhold som bruk/produksjonsdata for kjernevirksomheten i anlegget, direkte nytte for brukere, sikkerhet og pålitelighet, miljøforhold, sosiale virkninger og regionale virkninger.

Videre har vi inndelt prosjektene som typisk omfattes av Finansdepartementets kvalitetssikringsordning i fire hovedkategorier:

- Samferdselsprosjekter, all bygging av infrastruktur for samferdsel, både innen vei, jernbane, luftfart og sjøtrafikk. Her vil metodikken bli basert på dem som etatene har utarbeidet i forbindelse med tiltaksanalyser under Nasjonal Transportplan.
- Byggeprosjekter, alle prosjekter som omfatter bygninger og som ikke faller inn under samferdsel eller forsvarsprosjekter, for eksempel sykehus, undervisningsbygg, opera, osv.
- Forsvarsprosjekter, bygging av infrastruktur eller innkjøp av materiell til forsvarsaktivitet.
- Øvrige prosjekter, prosjekter som ikke dekkes av de ovennevnte kategoriene. Disse vil nok hovedsakelig være IT-prosjekter.

I Tabell 3 presenteres vurderingsparametre for hver av disse typene prosjekter. Noen av disse bør kunne inngå i de løpende registreringer som bør gjøres for alle prosjekter, andre er mer spesielle og må innhentes ved evaluering av det enkelte prosjekt. Typisk vil det være slik at jo lengre ned i tabellen en parameter er plassert, jo mer sannsynlig er det at innsamling av data vil måtte skje ”manuelt” og i forbindelse med selve ex post-evalueringen.

Samferdsel	Bygg	Forsvar	Øvrige
<i>Investeringskostnader</i>			
Investeringskostnader for hele og ulike komponenter/deler av prosjektet, gjerne beregnet som forholdstall per enhet (løpemeter vei, transportert km, o.l.)	Investeringskostnader for hele og ulike komponenter/deler av prosjektet, gjerne beregnet som forholdstall per enhet (kvadratmeter bygg, produsert enhet, o.l.)	Investeringskostnader for hele eller ulike komponenter/deler av prosjektet, gjerne beregnet som forholdstall per enhet (kvadratmeter bygg, produsert enhet, o.l.)	Investeringskostnader for hele eller ulike komponenter/deler av prosjektet, gjerne beregnet som forholdstall per enhet (per bruker, per MB data, o.l.)

Samferdsel	Bygg	Forsvar	Øvrige
<i>Driftskostnader</i>			
Driftskostnader, også sett i relasjon til investeringskostnader i en form for livssyklus kostnadsanalyse	Driftskostnader, også sett i relasjon til investeringskostnader i en form for livssyklus kostnadsanalyse	Driftskostnader, også sett i relasjon til investeringskostnader i en form for livssyklus kostnadsanalyse	Driftskostnader, også sett i relasjon til investeringskostnader i en form for livssyklus kostnadsanalyse
<i>Konsekvenser av prosjektet</i>			
Bruk; trafikkvolum, både i selve det bygde anlegget og rundt det	Bruksdata for kjernevirksomheten i bygget; bruksvolum, både i selve bygget og rundt det, produksjonsdata (saksbehandling, pasientbehandling, vektallsproduksjons, osv.), arbeidsmiljø-/betingelser for ansatte i bygget, endringer i transportmiddelfordeling, osv.	Data for den militære kjernevirksomheten; bruksvolum, både i/for prosjektets leveranser og rundt dem, "forsvarsevne" for ulike typer operasjoner (militært forsvar eller angrep, fredsbevarende operasjoner, miljøoppgaver, ressursforvaltning, osv.), sivil og militær beredskap, arbeidsmiljø-/betingelser for ansatte i bygget, osv.	Bruksdata for kjernevirksomhet i systemet; bruksvolum, både for systemet og systemer rundt det, produksjonsdata, produktivitet i forhold til hensikten med systemet, brukertilfredshet (internt og eksternt)
Reisetidsendringer			
Sikkerhet; ulykkesstatistikk			
Pålitelighet; "oppetid", avbruddsrisiko, sårbarhet i forhold til naturkatastrofer, osv.	Pålitelighet; "oppetid", avbruddsrisiko, sårbarhet i forhold til naturkatastrofer, osv.	Pålitelighet; "oppetid", avbruddsrisiko, sårbarhet i forhold til naturkatastrofer, osv.	Pålitelighet; "oppetid", avbruddsrisiko, sårbarhet i forhold til naturkatastrofer/terror, osv.
Miljøforhold; endringer i utslippsforhold, lokal luftkvalitet, støy, inngrep i naturen, landskapsestetikk, osv.	Miljøforhold; endringer i utslippsforhold, energibruk, inngrep i naturen, landskapsestetikk, innpass i eksisterende bebyggelse, osv.	Miljøforhold; endringer i utslippsforhold, energibruk, inngrep i naturen, landskapsestetikk, innpass i nærmiljøet, osv.	
Sosiale virkninger; nærhet til tilbud, bedret livskvalitet, osv.	Sosiale virkninger; nærhet til tilbud, bedret livskvalitet, osv.	Sivile sosiale virkninger; nærhet til tilbud, bedret livskvalitet, økt sikkerhet, osv.	Sosiale virkninger; nærhet til tilbud, bedret livskvalitet, økt sikkerhet, osv.

Samferdsel	Bygg	Forsvar	Øvrige
Regional virkning/effekt av lokalisering; arealbruksvirkninger, regional utvikling som følge av prosjektet, samsvar med politiske strategier og vedtak, sysselsetting, markedsmessige effekter, andre transportbehov, osv.	Regional virkning/effekt av lokalisering; arealbruksvirkninger, regional utvikling som følge av prosjektet, samsvar med politiske/regionalpolitiske strategier og vedtak, sysselsetting, markedsmessige effekter, transportbehov, bidrag til byutvikling, osv.	Sivil regional virkning/effekt av lokalisering; arealbruksvirkninger, regional utvikling som følge av prosjektet, samsvar med politiske/regionalpolitiske strategier og vedtak, sysselsetting, markedsmessige effekter, transportbehov, bidrag til byutvikling, osv.	Regional virkning/effekt av prosjektet; industriell utvikling avhengig av hyllevarer versus nyutvikling, av om leverandører er regionale/nasjonale/internasjonale, samsvar med politiske/regionalpolitiske strategier og vedtak, sysselsetting, markedsmessige effekter, osv.

Tabell 3 Vurderingsparametre for ulike typer prosjekter

5.5 Ansvar for og roller i gjennomføring av ex post-evaluering

Dersom det skal etableres en ordning med systematiske ex post-evalueringer av offentlige prosjekter, er et viktig spørsmål hvem som skal ha ansvaret for disse evalueringene. En generell betraktning er at den omtalte løpende registreringen av standardparametre nok bør være den relevante etatens ansvar, i motsetning til prosjektet, som avvikles for raskt for å ha dette ansvaret. Dette vil for veiprosjekter si Statens Vegvesen, for jernbaneprosjekter Jernbaneverket, for byggeprosjekter utført av Statsbygg faller ansvaret på Statsbygg, osv. Selve de innsamlede data/målinger og adgangen til andre for å bruke disse bør kanskje derimot forvaltes av Finansdepartementet eller noen bemyndiget til dette av Finansdepartementet, for eksempel forskningsprogrammet Concept.

Når den gjelder den mer omfattende, helhetlige og prosjektspesifikke evalueringen vil denne måtte gjennomføres av en uavhengig instans. Dette utelukker selvfølgelig både etaten som har hatt ansvaret for prosjektet og andre fagmiljøer som har vært involvert i prosjektet på ett eller annet tidspunkt. Samtidig må det sikres at det finnes nok uavhengige, faglige miljøer til å få utført vurderingene fra gang til gang. Det finnes også andre roller som må bekles i en slik ordning:

- Bestiller av evalueringsoppdrag, dette bør normalt være det relevante fagdepartementet som for KS1/KS2, men i konfliktstilfeller kan det være et interdepartementalt utvalg eller Finansdepartementet.

Organisasjon/ kriterium	Legitimitet og uavhengighet	Fagkompetanse	Kapasitet
Riksrevisjonen	Vil nok oppfattes som uavhengig, men sees neppe på som en naturlig aktør for slike oppdrag	Har primært kompetanse om offentlige regelverk, mindre om prosjektledelse	Har kapasitet eller denne kan økes om en slik permanent ordning tillegges etaten
Senter for Statlig Økonomistyring	Vil nok oppfattes som uavhengig, men sees neppe på som en naturlig aktør for slike oppdrag	Har primært kompetanse om offentlig økonomistyring, mindre om prosjektledelse	Har kapasitet eller denne kan økes om en slik permanent ordning tillegges etaten
Statistisk Sentralbyrå	Vil nok oppfattes som uavhengig, men sees neppe på som en naturlig aktør for slike oppdrag	Har primært kompetanse om statistikk og dataanalyse, mindre om prosjektledelse	Har kapasitet eller denne kan økes om en slik permanent ordning tillegges etaten
Konsulentfirmaer/ forskningsmiljøer med rammeavtale	Om disse utvelges med omhu og med bakgrunn i tradisjon etablert for KS1/KS2 burde slike aktører oppfattes som uavhengig og passende	Riktige miljøer eller grupperinger av miljøer bør besitte all nødvendig kompetanse	En utfordring kan være at mange av de mest kvalifiserte miljøene i Norge allerede er involvert i KS1/KS2, men det finnes noen flere og ved å supplere med andre skandinaviske miljøer burde det finnes nok kapasitet, evt. kan dette tillegges Concept-programmet
Utlysning fra gang til gang	Om leverandører utvelges med omhu og med bakgrunn i tradisjon etablert for KS1/KS2 burde slike aktører oppfattes som uavhengig og passende	Riktige miljøer eller grupperinger av miljøer bør besitte all nødvendig kompetanse	I tillegg til kapasitet for å gjennomføre selve oppdraget krever en slik tilnærming at det finnes miljøer med kapasitet til å levere tilbud, noe som kan være en utfordring avhengig av oppdragssituasjon for relevante miljøer

Tabell 4 Vurdering av aktører for gjennomføring av ex post-evalueringer

- Finansiering av evalueringene, her ser vi to muligheter. Dersom det er praktisk mulig i forhold til avslutning av prosjektene, kan gjerne kostnadene til innsamling av data og andre oppgaver i forbindelse med forberedelse for

og gjennomføring av ex post-evaluering belastes prosjektets budsjett. Samtidig vil nok i mange tilfeller disse oppdragene komme for sent i tid til at prosjektens budsjett kan dekke dem, slik det gjøres for KS1/KS2. I så fall mener vi det bør finnes en egen budsjettpost for disse utgiftene, sannsynligvis i Finansdepartementet eller gjennom Concept-programmet.

- Innsamling av data, håndteres av de relevante etatene med bruk av standard indikatorer i spesielt samferdselsprosjekter. For andre typer prosjekter er det vanskeligere å peke på en generell regel, normalt vil data komme fra både drifter og den som driver kjernevirksomheten i anlegget, for eksempel Statsbygg, forsvaret, Den Norske Opera, osv.
- Tolke data og gjennomføre ex post-evalueringen samt skrive rapport, her finnes det flere alternativer; Riksrevisjonen, Senter for Statlig Økonomistyring, Statistisk Sentralbyrå, konsulentfirmaer med rammeavtale for slike oppdrag, forskermiljøer og at det gjennomføres en åpen utlysning fra gang til gang. Vi anbefaler at det etableres en egen gruppe av ekspertmiljøer (en annen enn den som allerede har avtaler for KS1/KS2) som gis en form for rammeavtale for å utføre ex post-evalueringer. Enten bør det åpnes for at disse miljøene kan omfatte internasjonale aktører, da primært fra Skandinavia grunnet språkhensyn, eller dette ansvaret kan tillegges Concept-programmet så lenge det består. Denne anbefalingen er basert på vurderingene gjengitt i Tabell 4.

5.6 Konsekvenser av den foreslåtte tilnærmingen og forutsetninger for gode ex post-evalueringer

I dette siste delkapittelet om den foreslåtte tilnærmingen til ex post-evalueringer vil vi vurdere noen sannsynlige konsekvenser av ordningen. Disse representerer i stor grad antakelser og spekulasjoner og må vurderes deretter, her finnes det lite konkret underlag å basere diskusjonen på. Når det gjelder omfang for det enkelte oppdrag med å gjennomføre en ex post-evaluering av et prosjekt, antar vi at dette vil være sammenliknbart med KS1-/KS2-analysene. Det betyr omtrentlig samme varighet og samme kostnad som for disse oppdragene.

For å sikre at ex post-evalueringene kan gjennomføres under egnede betingelser, vil det måtte stilles noen krav til både KS1-/KS2-analysene og annet datagrunnlag. For det første vil gode ex post-evalueringer, som tar for seg prosjektspesifikke forhold, avhenge av at det i KS1-analysene for prosjektet legges til rette for, men ikke begrense omfanget/bredden av, den etterfølgende ex post-evalueringen. Dette betyr

spesielt at det ved KS1 må defineres mer spesifikke kriterier og måleparametre som det er mulig å benytte for etterprøving i en ex post-evaluering. Det må også kunne kreves at de analysene som er gjort i forbindelse med KS1 er transparente slik at det er mulig å sjekke hvilke parametre som har endret seg som følge av prosjektet. En kan nemlig tenke seg at de samfunnsmessige resultatene av prosjektet blir et annet enn det som KS1-analysene ga som resultat, men som følge av forhold som lå utenfor prosjekteiers kontroll og ikke som følge av utilstrekkelige eller feilaktige analyser.

For eksempel kan et samferdselsprosjekt tenke seg at i ettertid vil ligge utenfor eller i ytterkant av det intervallet som ble lagt til grunn ved beregning av forventningsverdi på årlig trafikkendring, eller at det ble gjennomført endring i tilstøtende transportnett som har gitt trafikkmessige effekter av mer tilfeldig karakter. Dersom en endrer rammebetingelser for en av transportmåtene som er i en konkurransesituasjon, kan dette betraktes som forhold som er eksogene i forhold til beregningene, dersom dette ikke kunne sies å være tilgjengelig informasjon under KS1-prosessen. Etterprøvbare mål og transparente ex ante-analyser er derfor viktig underlagsmateriale for gjennomføring av analyser ex post. Derfor er det naturlig at det er en kobling mellom metodikken i KS1 og ex post i den forstand at en i KS1 formulerer mål og gjennomfører analysene på en slik måte at den ”store tilbakekoblingen” mellom (a) samfunnsmessige mål, effektmål, analyser ex ante og (b) analyser ex post opp mot disse målene lar seg gjennomføre. I den sammenheng bør avvik fra analyseresultatene ex ante la seg identifisere for å finne ut om avvikene skyldes data, analysemodell eller eksogene tilfeldigheter. Dersom det er en god, operasjonell målformulering i KS1 som ex ante-analysene er gjennomført opp mot for å sikre at tiltaket er analysert i henhold til å identifisere grad av måloppnåelse, så vil viktige premisser for at analysene ex post skal bli av god kvalitet, være lagt. Kvaliteten på analyser av store transportprosjekter hviler for en stor del på hvordan prognosearbeidet er utført i tidligfasen. Flyvbjerg, Holm og Buhl (2005) fokuserer på forekomst av skjevheter i trafikkberegninger sammenlignet med faktisk utvikling. De har særlig vært opptatt av å unngå ”optimism bias” ved at det legges for optimistisk trafikkutvikling til grunn i tidligfasevurderingene. Vi vil imidlertid ikke gå videre inn på denne problematikken her, men heller henviser til Flyvbjerg, Holm og Buhl for en dypere behandling, også om referanseprognosemetoden, som de peker på som et mulig svar på unøyaktige ex ante-analyser.

I neste kapittel vil vi forsøke illustrere videre hvordan den anbefalte metoden for ex post-evaluering kan ta form ved å demonstrere bruk av den på noen konkrete prosjekter.

6 Eksempler på gjennomføring av ex post-evaluering med ny metode

I dette kapitlet presenterer vi noen konkrete eksempler på hvordan ex post-evalueringer kan gjennomføres i tråd med anbefalingene fra foregående kapittel. Relatert til den tredelte anbefalingen i kapittel 5.4, kan eksemplene sees på som illustrasjoner av disse punktene:

- Prosjektet for bygging av nytt rikshospital er illustrert evaluert med vår metode i kapittel 6.1. Dette viser hvordan en bredere og prosjektspesifikk evaluering av kan gjennomføres.
- Oppfølging av jernbane- og veiprosjekter i kapitlene 6.2 og 6.3 er eksempler på hvordan en løpende oppfølging av mer konkrete parametre kan gjøres. I tillegg viser evalueringen av veiprosjektet hvordan en bredere nytte-kostnadsanalyse kan gjennomføres.

6.1 Evaluering av Nytt Rikshospital

Det nye Rikshospitalet på Gaustad i Oslo er et universitetssykehus med 585 senger. Byggeprosjektet ble avsluttet vinteren 2001, 2 ½ år senere enn planlagt, og med kostnadsoverskridelser på om lag 30 %. På grunn av de tids- og kostnadmessige overskridelsene tok Stortinget i 2000 initiativ til en evaluering av prosjektet. Tre år etter at det nye sykehusanlegget var blitt tatt i bruk, publiserte Rikshospitalet dessuten selv en sammenstilling av utviklingen på utvalgte måleparametre for behandlingmessig og vitenskapelig produktivitet og oppfyllelse av eksterne pålegg.

6.1.1 Foreliggende evalueringer

Den førstnevnte evalueringen ble, i samsvar med Regelverk for Statens anskaffelsesvirksomhet, gjennomført i 2000/2001 på oppdrag av det daværende Arbeids- og administrasjonsdepartementet (Whist et al., 2001). Denne vurderingen fokuserte på *gjennomføringen* av prosjektet, og ikke på de samfunnsmessige effektene

(bortsett fra ressursbruken målt som rene byggekostnader). De spørsmålene Arbeids- og administrasjonsdepartementets evaluering fokuserte på, var disse:

- 1) Har forutsetningene for prosjektet vært realistiske?
- 2) Har planleggingen av prosjektet vært god nok?
- 3) Har kostnadsoverslaget vært realistisk og grundig nok?
- 4) Har prosjekteringen vært realistisk og grundig nok?
- 5) Har styringen med prosjektet vært god nok?
- 6) Har ansvarsforholdene i prosjektet vært godt nok avklart?
- 7) Har organiseringen og faseinndelingen av prosjektet vært god nok?
- 8) Har inndeling av entrepriser og valg av entreprisformer, styring og samordning av entrepriser, byggeledelse og ledelse av utstyrsprosjektet vært gode nok?
- 9) I hvilken grad har grensesnittet mellom bygg og utstyr vært godt nok?
- 10) Har informasjonsflyten i prosjektet vært god nok?
- 11) I hvilken grad har man vært bevisst, og hvordan har man avvendt konflikterende mål som kvalitet, økonomi (kostnader og kostnadseffektivitet) og overholdelse av tidsfrister?
- 12) I hvilken grad har endringer underveis i prosjektet vært fornuftige ut fra hensyn til mål, kvalitet og økonomi?
- 13) Vurdering av om rolleavklaring mellom de ulike aktørene har vært god nok
- 14) Har løsninger og prosess bidratt til oppfylling av mål for NHR?

De sammenfattende konklusjonene tok dels utgangspunkt i et prosjektorganisasjonsperspektiv og dels i utbyggerens perspektiv. Et bredere samfunnsmessig (samfunnsøkonomi og samfunnsnytte) perspektiv inngikk imidlertid ikke i evalueringen.

Av punktene ovenfor inneholder svarene på pkt. 3 (realismen i kostnadsoverslaget) opplysninger som er relevante i forhold til de faktorene og parametrene som er nevnt i vårt forlag til vurderingsmetodikk for byggeprosjekter, men de faktiske byggekostnadene sammenliknes i evalueringen verken med andre tilsvarende prosjekter eller med den oppnådde nytten, men med de kostnadsoverslagene som var utarbeidet på forhånd. Det er derfor kvaliteten på budsjetteringen som evalueres, og ikke hvor godt økonomisk prosjektet har vært.

Også svarene på pkt. 11, 12 og 14 inneholder opplysninger som er relevante i forhold til vår foreslåtte vurderingsmetode. Vurderingsgrunnlaget er imidlertid kun

prosjektets egne mål, og ikke mål i et bredere samfunnsmessig perspektiv. Mål som gjelder sideeffekter av prosjektet (å oppnå positive sideeffekter og unngå/reducere negative sideeffekter) var f. eks. ikke med i vurderingsgrunnlaget.

Vår foreslåtte metode innebærer at også deler av planleggings- og beslutningsprosessen skal gjøres til gjenstand for evaluering. Seriositeten i utredningen av alternative konsepter inngår som en sentral del her. Arbeids- og administrasjonsdepartementets evaluering av Prosjekt Nytt Rikshospital dreide seg imidlertid kun om det realiserte prosjektet. Hvorvidt alternative konseptløsninger (herunder sammenslåing med andre sykehus, ulike grader av desentralisering av Rikshospitalets funksjoner til andre sykehus, og alternative lokaliseringer av nytt Rikshospital) ble behandlet seriøst, inngikk ikke i evalueringen. Heller ikke vurderingen av nullalternativet (fortsatt sykehusdrift i de gamle rikshospitalsbygningene i Pilestredet) ble evaluert.

I tillegg til Arbeids- og administrasjonsdepartementets evaluering, publiserte Rikshospitalet selv en sammenstilling av utviklingen for utvalgte effektindikatorer de tre første årene etter at det nye sykehuset ble tatt i bruk (Rikshospitalet, 2004). De utvalgte indikatorene var:

- Antall vitenskapelige artikler, doktorgrader og patenter
- Antall behandlinger
- Antall dager på ventelister, og
- Oppfyllelse av pålegg fra helsetilsynet.

Disse indikatorene forteller noe om utviklingen innenfor enkelte, relevante aspekter ved sykehusets ”performance”, men de sier ikke i seg selv noe om endringene skyldes selve byggeprosjektet (mer hensiktsmessige lokaler og utstyr), stordriftsfordeler pga. sammenslåing av tidligere separate administrative enheter, eller forhold som ikke har noe å gjøre med byggeprosjektet og de administrative forutsetningene som lå til grunn for dette (f. eks. generelt skjerpete krav til doktorgradskompetanse for å få vitenskapelige stillinger, eller til at forskere publiserer vitenskapelige artikler). Indikatorene er dessuten begrenset til Rikshospitalets egen ”produksjon”, og sier ikke noe om hvorvidt økt aktivitet på Rikshospitalet helt eller delvis er en følge av overføring av oppgaver fra andre sykehus til Rikshospitalet. Indikatorene er med andre ord i seg selv ikke i stand til å si om behovsdekningen på samfunnsnivå innenfor de områdene indikatorene gjelder, er blitt styrket som følge av byggeprosjektet.

Rikshospitalets sammenstilling av effekter omfatter heller ikke sidevirkninger av prosjektet, f. eks. miljøkonsekvenser.

6.1.2 Anvendelse av vår foreslåtte metode på Rikshospitalprosjektet

Vår foreslåtte metode legger stor vekt på produktevaluering, men vi anbefaler at det også inngår elementer av en *prosessevaluering* av framleggings- og beslutningsprosessen. Som det framgår annet sted i denne rapporten (kapittel- og delkapittel angis her), anbefaler vi at disse prosessrelaterte temaene tas med i evalueringen:

- 1) Hvilke føringer ble satt for ex ante-analysene og konseptvalg?
- 2) Har det skjedd utvikling/endringer i nytte-kostnadsestimater utover i prosessen eller i design/utforming av konsept, eller avbøtende miljøtiltak?
- 3) Grad av rimelighet i ulike beregninger (nytte, kostnad, osv.) og diskonteringsfaktorer for realisert og alternative konsepter
- 4) Seriøsitet i utredningen av alternative konsepter.

Arbeids- og administrasjonsdepartementets evaluering fokuserer først og fremst på punktene 2 og 3 ovenfor. Evalueringen konkluderer bl.a. med at budsjetttrammer og kostnadsoverslag ikke var realistiske, særlig fordi de bygde på en rekke delforutsetninger som hver for seg var oppnåelige, men vanskelige å nå, og som det samlet sett var nærmest umulig å oppfylle. Det gjøres videre detaljert rede for hvordan konseptet er blitt endret gjennom en rekke utvidelser underveis i prosessen, og hvordan disse utvidelsene i sin tur medført nye kostnadsoverslag. Materialet vi har hatt tilgang til, tyder ikke på at det er blitt utført formelle nytte-kostnadsanalyser av prosjektet, og grad av rimelighet i diskonteringsfaktorer og nytte-kostnadsestimater vil derfor ikke være aktuelle evalueringstemaer for dette prosjektet.

Vi kjenner ikke til hvilke føringer som ble satt for ex ante-analysene og konseptvalg, men dette er ting som ifølge vår foreslåtte metode ville måtte undersøkes. Det samme gjelder seriøsiteten i behandlingen av alternative konsepter. Ut fra det foreliggende materialet ser det ikke ut til å ha vært gjort særlig omfattende arbeid med vurdering av alternative konsepter. Ifølge Statsbygg la Sosial- og helsedepartementet i St.meld. 38 (1987-88) fram en analyse av flere lokaliseringalternativer, hvorefter stortinget 8. oktober 1988 konkluderte med at det skulle startes planlegging av et nytt Rikshospital på Gaustad som erstatning for fire

sykehus i Oslo: Det gamle Rikshospitalet, Revmatismesykehuset, Sophies Minde og Kronprinsesse Märthas Institutt. Vurderingen av alternative lokaliseringkonsepter begrenser seg altså til de overveielsene som ble gjort i den nevnte stortingsmeldingen.

Ifølge Arbeids- og administrasjonsdepartementets evalueringsrapport (Whist et al., 2001), var aktivitetsforutsetningene fra 1992 et kompromiss mellom de kreftene som ønsket et stort og de som ønsket et lite rikshospital. I den politiske prosessen i siste halvdel av 1980 og begynnelsen av 1990-årene gikk et mindretall av Stortingets sosialkomité (medlemmene fra Senterpartiet) inn for en samordning av Ullevål sykehus og Rikshospitalet, lokalisert til Ullevål (Innst.S.nr.193 (1993-1994)). Det ser imidlertid ikke ut til å ha vært foretatt noen formell vurdering av dette eller andre alternative konsepter.

Når det gjelder *produktrelaterte* tema, innebærer vår anbefalte metode at effektivvurderingen omfatter følgende parametere:

- Investeringskostnader, i ulike komponenter/deler av prosjektet, evt. beregnet som forholdstall per enhet
- Driftskostnader, også sett i relasjon til investeringskostnader i en form for LCC-analyse
- Konsekvenser:
- Bruk; bruksvolum, både i selve bygget og rundt det
- Pålitelighet; ”oppetid”, avbruddsrisiko, sårbarhet i forhold til naturkatastrofer, osv.
- Miljøforhold; endringer i utslippsforhold, energibruk, inngrep i naturen, landskapsestetikk, innpass i eksisterende bebyggelse, osv.
- Sosiale virkninger; nærhet til tilbud, bedret livskvalitet, osv.
- Regional virkning/effekt av lokalisering; arealbruksvirkninger, regional utvikling som følge av prosjektet, samsvar med politiske/regionalpolitiske strategier og vedtak, sysselsetting, markedsmessige effekter, transportbehov, bidrag til byutvikling, osv.

- Kjernevirksomheten i bygget; produksjonsdata (saksbehandling, pasientbehandling, vektallsproduksjon osv.), arbeidsmiljø/-betingelser for ansatte i bygget, osv.

De faktiske investeringskostnadene framgår av Arbeids- og administrasjonsdepartementets evalueringsrapport og beløp seg til 5944 millioner kr (Whist et al., 2001). Basert på opplysninger hos Whist et al (ibid.) og Rikshospitalet (2004), kan investeringskostnadene omregnes til følgende forholdstall per enhet:

- Per m² golvareal (BTA): ca. 31.500 kr
- Per ansatt: ca. 1,6 mill. kr.
- Per inneliggende pasient i 2003: ca. 102.000 kr
- Per poliklinisk behandling i 2003: ca. 36.000 kr

Ifølge Arbeids- og administrasjonsdepartementets evaluering (Whist et al, 2001) var de årlige driftskostnadene for Rikshospitalet i 2001 ca. 2,6 milliarder kr. Dette tallet utgjør en noe høyere andel av investeringskostnadene enn det som er vanlig, ettersom investeringskostnadene ved sykehusutbygging normalt tilsvarer driftskostnadene i 2,5 – 3 år (Whist et al., ibid.). For Rikshospitalet er det tilsvarende forholdstallet litt under 2,3. Arbeids- og administrasjonsdepartementets evalueringsrapport peker imidlertid på at Rikshospitalet har en mye større andel av høykostnadsbehandling enn et gjennomsnitts sykehus.

6.1.3 Konsekvenser

Bruk/bruksvolum. Ifølge Rikshospitalet (2004) hadde sykehuset i 2003 58.000 inneliggende pasienter og 165.000 polikliniske pasienter. Disse tallene er betydelig høyere enn det behandlingsnivået Rikshospitalet ble dimensjonert for (henholdsvis 33.000 og 133.000 pasienter). Ut fra disse tallene alene kan man likevel ikke konkludere sikkert om hvordan prosjektet har bidratt til den samlede behovsdekningen mht. pasientbehandling. Hvis kostnadsoverskridelsene f. eks. har medført at andre sykehus eller sykehusavdelinger er blitt nedlagt eller har fått redusert tilbudene, må det korrigeres for slike effekter før konklusjonen trekkes om Rikshospitalets netto bidrag til dekning av behovene for pasientbehandling.

Pålitelighet. Vårt foreliggende materiale inneholder ikke opplysninger om driftssikkerhet, ”oppetid” mv. Slike opplysninger vil det trolig ikke være så vanskelig å innhente, og tallene kan så sammenliknes med tilsvarende tall fra sammenliknbare sykehus og sykehusavdelinger andre steder. Rikshospitalet er med sin beliggenhet på Gaustad neppe særlig utsatt for naturkatastrofer, ettersom området verken ligger i risikosone for flom, skred, orkaner eller andre naturkatastrofe-typer som forekommer i Norge.

Miljøforhold. Det ble gjennomført en ex ante-vurdering av miljøkonsekvenser etter bestemmelsene i plan- og bygningslovens kapittel VII-a. Verken Arbeids- og administrasjonsdepartementets ex post-evaluering eller Rikshospitalets egen erfaringsoppsamling nevner imidlertid slike konsekvenser. Det var imidlertid betydelige konflikter omkring riving av Gaustad gård, som mange mente var verneverdig, og omlegging av Gaustadbekken. Begge disse inngrepene var konsekvenser av det valgte utbyggingskonseptet. En miljøorganisasjon (Natur og Ungdom) stevnet staten for behandlingen av reguleringssaken og gjennomførte aksjoner for å prøve å hindre rivingen (Statsbygg, 2001). Det var også betydelig strid omkring trafikkspørsmål som bilveg og parkering kontra kollektivtrafikk, og framføring og finansiering av trikk. Det ferdigstilte prosjektet omfatter en forlengelse av sporvognslinjen fra Ullevål Hageby og en parkeringsdekning på ca. 0,35 parkeringsplasser per ansatt.

Ifølge Tennøy (2004) økte andelen av de ansatte som reiste til jobben med bil, fra 19 til 39 % etter flyttingen av Rikshospitalet fra Pilestredet til Gaustad, mens andelen som gikk eller syklet sank fra 28 til 19 % og kollektivtrafikk-andelen gikk ned fra 52 til 42 %. Gaustad ligger mer perifert i bystrukturen og i et område med lavere befolkningstetthet enn Pilestredet, og andelen av befolkningen (og dermed også sannsynligvis av de ansatte) som kan nå sykehuset innenfor akseptabel gang- eller sykkelavstand er derfor lavere på Gaustad enn i Pilestredet. Tilgjengeligheten med kollektiv transport er også dårligere, trass i den nyetablerte trikkeforbindelsen. På forhånd hadde Oslo kommune satt som forutsetning/krav/målsetting at kollektivandelen skulle være minst 50 %. Sammenliknet med den tidligere lokaliseringen i Pilestredet er transportmiddelfordelingen for de ansattes reiser til og fra jobben endret i en retning som gir høyere energiforbruk og utslipp. Det er også sannsynlig at bilandelen er noe høyere enn den ville vært ved samlokalisering med Ullevål sykehus, som er den eneste alternative lokaliseringen som nevnes i de dokumentene vi har hatt tilgang til.

På den annen side har flyttingen av Rikshospitalet åpnet for byomdanning av den tidligere lokaliteten i Pilestredet, med etablering av ca. 1400 sentralt beliggende

boliger og et betydelig antall arbeidsplasser og høy vektlegging av energisparende og økologisk gunstige byggeprinsipper. Disse boligene og arbeidsplassene har fått en lokalisering som legger til rette for en høy andel ikke-motorisert og kollektiv transport og et lavt transportbehov i hverdagen. For å kunne anslå de samlede (både direkte og indirekte) transporteffektene og transportrelaterte miljøkonsekvensene av byggingen av Rikshospitalet på Gaustad, måtte man anslå hvilken alternativ lokalisering og utforming disse boligene og arbeidsplassene sannsynligvis ville fått hvis tomten i Pilestredet Park ikke var blitt ledig.

I tillegg til å vurdere transportrelaterte miljøkonsekvenser, vil en ex post-evaluering av miljøkonsekvenser etter vår anbefalte metode innebære en vurdering av energibruk i det nye bygget (per kvadratmeter golvareal og per ansatt), samt konsekvenser i form av nedbygging av natur- og landbruksarealer, inngrep i særlige økosystemer, eventuell blokkering eller nyetablering av stier eller turveger, samt landskapsmessige virkninger. Materialet vi har hatt tilgang til gir imidlertid ikke grunnlag for å foreta slike vurderinger her.

Sosiale virkninger. En analyse av hvordan befolkningens nærhet til behandlingstilbud har endret seg på grunn av prosjektet, må ta hensyn både til befolkningsmengden innenfor ulike avstandssoner fra Rikshospitalet og befolkningsmengden innenfor tilsvarende avstandssoner fra eventuelle behandlingstilbud som er nedlagt som følge av etableringen av nytt Rikshospital. Hvis analysen i tillegg skal omfatte bedret livskvalitet (bør vi virkelig anbefale dette – det er ikke lett å måle!), vil vel dette forutsette en bredt anlagt spørreskjemaundersøkelse blant et representativt utvalg av befolkningen i hele anleggets influensområde, dvs. i hele Norge.

Regionale og arealbruksmessige virkninger. Siden både det gamle Rikshospitalet og de tre øvrige tidligere sykehusene som inngår i det nye Rikshospitalet lå i (Stor-)Oslo, må de regionalpolitiske virkningene av etableringen på Gaustad anses som minimale. Sammenliknet med tenkelige alternative lokaliseringer vil det imidlertid kunne forekomme regionalpolitiske virkninger. Hvis ett av lokaliseringalternativene hadde vært f. eks. i en by eller region der det i de senere årene er nedlagt mange industriarbeidsplasser, ville etablering av Rikshospitalet i denne byen eller regionen utvilsomt fått betydelige regionale konsekvenser. Den regionalpolitisk motiverte utflyttingen av statlige institusjoner fra Oslo i 1970- og 1980-årene viser at staten ikke har vært fremmed for å bruke lokaliseringen av statlige institusjoner som et regionalpolitisk virkemiddel. Vi kjenner imidlertid ikke til hvilke lokaliseringskonsepter som forelå i den tidlige fasen av planleggingen av det nye rikshospitalet, og vi har derfor ikke grunnlag for å si noe om sannsynlige regionale

virksomheter av den valgte lokaliseringen kontra alternative lokaliseringer og/eller dimensjoneringer av Rikshospitalet.

Når det gjelder arealbruksmessige virkninger, representerer byggingen på Gaustad en av de få utvidelsene av byggesonen i Oslo som har skjedd i løpet av de siste par tiårene, hvor det alt dominerende av utbyggingen er skjedd som fortetting og bytransformasjon. Som byutvidelse ligger imidlertid Gaustadområdet i direkte forlengelse av et eksisterende område med kunnskapsbedrifter (Blindern/Forskningsparken) og i direkte tilknytning til Oslo Universitets allerede eksisterende prekliniske institutt og det psykiatriske sykehuset på Gaustad. Selv om utbyggingen medførte bygging på inntil da ubebygde jorder, hadde byggetomten i lang tid vært satt av i kommuneplan/generalplan til utbyggingsformål, og utbyggingen medførte ikke brudd på markagrensen. Byplanmessig må lokaliseringen på Gaustad karakteriseres som rimelig gunstig, både i forhold til mulige agglomerasjonsfordeler ved å være lokalisert i et område med mange forskere på felles og tilgrensende fagområder, mht. atkomst og tilgjengelighet, og mht. områdets lokale kvaliteter (solforhold, utsikt osv.).

Kjernevirksomheten. Rikshospitalets erfaringsoppsamling (2004) angir som nevnt at antall inneliggende pasienter i 2003 var 58.000 og antall polikliniske pasienter 165.000. Per ansatt innebærer dette henholdsvis ca. 15 og litt over 40 pasienter årlig. For å sette disse tallene i perspektiv, burde det gjøres en sammenlikning med tilsvarende tall fra andre sykehus med tilsvarende behandlingstilbud. Det framgår videre at utbyggingen har ført til et dramatisk fall i antall saker på venteliste for behandling på Rikshospitalet (fra 450 i 2001 til 125 i 2004). Rikshospitalets erfaringsoppsamlingsnotat peker vider på at det er skjedd en vesentlig framgang fra 2001 til 2003 i oppfyllelsen av pålegg fra Helsetilsynet. Det framgår imidlertid ikke om den lave graden av oppfyllelse i 2001 skyldtes problemer i forbindelse med flytteprosessen eller avspeilet et stabilt nivå før flyttingen. (Det samme kan også innvendes i forhold til venteliste-tallene ovenfor.) For å gi et realistisk bilde av effekten av prosjektet i forhold til disse parametrene, bør det derfor skaffes data som går lenger tilbake i tiden.

Rikshospitalets erfaringsoppsamling (2004) framholder videre at pasienttilfredsheten (inkludert tilfredshet med standard på bygninger så vel som erfaringer med lege og sykepleietjenesten, og organiseringen og informasjonen omkring egen sykdom og behandling) er høy. Dette er dokumentert både gjennom sykehusets egne undersøkelser og gjennom den nasjonale undersøkelsen Stiftelse for Helsetjenesteforskning har utført. I den sistnevnte undersøkelsen kom Rikshospitalet best ut blant de store sykehusene. Det framgår imidlertid ikke om den

høye pasienttilfredsheten er kommet etter at det nye rikshospitalet ble etablert på Gaustad, eller om Rikshospitalet også før flyttingen skåret høyt på pasienttilfredshet. For å kunne vurdere hvordan byggeprosjektet har påvirket pasienttilfredsheten, må dette bringes på det rene.

Ifølge Rikshospitalets erfaringsoppsamling (2004) har alle relevante Helse-, miljø- og sikkerhetstilsyn undersøkt Rikshospitalet etter flyttingen og konkludert med at det ikke forelå alvorlige mangler eller avvik. HMS-arbeidet ble også funnet å være i god gjenge. Samme rapport oppgir også at sykefraværet er gått ned og er lavere enn ved de fleste sammenliknbare sykehus. Det framgår imidlertid ikke i hvor lang tid sykefraværet har vist nedadgående tendens. Det er ikke utenkelig at det kan ha vært et forhøyet sykefravær i perioden opp mot og i flytteperioden, som ofte vil oppleves som stressende. For å få et reelt bilde av utviklingen mht. arbeidsmiljø, stress og sykefravær bør det derfor skaffes data som går tilstrekkelig langt tilbake (2-3 år?) i tiden før flytting.

6.2 Effektvurdering av fire jernbaneprosjekter

I det følgende illustreres en måte å følge opp infrastrukturprosjekter. Oppfølgingen er fokusert på bruk av et begrenset antall indikatorer, som sammenstilles over en forholdsvis lang tidsperiode. I dette tilfelle er data for punktlighet, reisetid, togfrekvens og antall reisende brukt. Fremstillingen er basert på Olsson (2005)

Data om reisetid er basert på publikumsruter og grafiske ruter. Antall avganger er enten beregnet fra visuell inspeksjon av grafiske ruter, eller fra Jernbaneverkets publikasjon Jernbanestatistikk. Antall reisende i perioden 1987 til 2004 er basert på På Sporet (1987-2005). Kostnadstall er basert på opplysninger i statsbudsjett (St.prp 1) dersom ikke annet er angitt. Punktlighet er basert på en kombinasjon av opplysninger i Jernbaneverkets årlige punktlighetsstatistikk og kompletterende opplysninger fra Jernbaneverkets punktlighetsoppfølgere.

6.2.1 Vestfoldbanen og Sandeparsellene

I hovedrapporten fra utredningen av modernisert Vestfoldbane mellom Oslo og Skien presenterte NSB (1992a) hovedtrekkene i prosjektet. Når prosjektet blir omtalt i Norsk Jernbaneplan 1994-97 (St.meld nr. 35, 1992-93) legges det opp til å begynne med to parseller (Kobbervikdalen-Skoger og Åshaugen-Sande-Holm). Det er også disse to parsellene som først får bevilgninger (St.prp. nr 1 1992-93 og St.prp.

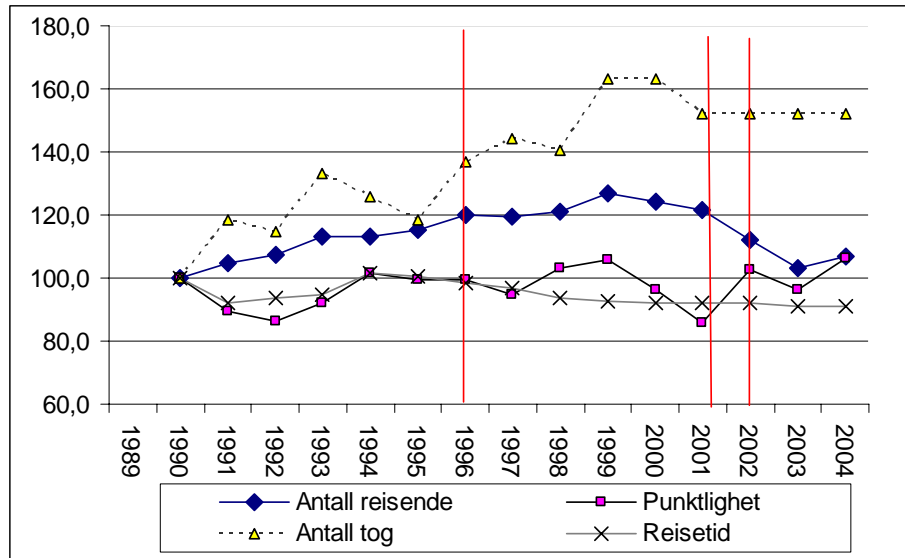
nr 1 1993-94). Tabell 5 viser de parseller som er bygget hittil, når de ble åpnet for trafikk og kostnader for utbyggingene.

Parsell	Åpnet for trafikk ¹⁶	Første estimat			Endelig sum		
		NOK	År	2002 NOK	NOK	År	2002 NOK
Kobbervikdalen-Skoger (parsell 2)	12 aug 1995	133	1994	159,3	162,4	1995	189,8
Skoger-Åshaugen (parsell 3)	5 okt 2001	344	1995	402,1	448,3	2002	448,3
Åshaugen-Sande-Holm (parsell 4)	5 okt 2001	224	1994	268,4	512,6	2002	512,6
Nykirke kryssingsspor	11 juni 2002	120	2000	125,2	122,9	2002	122,9
Totalt				955			1274

Tabell 5 Utbygde nye parseller på Vestfoldbanen, med åpningsdato og kostnad

Figur 11 viser utviklingen år for år for punktlighet, antall reisende, reisetid og antall avganger. Antall avganger var vært vesentlig øket siden planene for utbygging av Vestfoldbanen ble laget. I tillegg har antall reisende økt. Punktlighet og reisetid har kun i mindre omfang blitt endret i løpet av den studerte perioden. Utbyggingen har i liten grad gitt reduksjon i reisetid. Økningen i frekvens som var planlagt i NSB (1992a) er stort sett gjennomført, selv om kun cirka 12 % av den da planlagte dobbeltsporstrekningen er utbygget. Det er dokumentert at økninger i antall tog på en banestrekning og økning i antall reisende ofte gir lavere punktlighet (se Olsson og Haugland, 2004). Man skulle derfor ha forventet en reduksjon i punktligheten som en følge av økning i antall tog og passasjerer. Analysen tyder på at den utførte utbyggingen har bidratt til at antallet tog kunne økes på Vestfoldbanen uten at punktligheten er blitt redusert. Det hører til saken at det ikke går godstog på Vestfoldbanen siden 1999. I tillegg er fremføringstiden for persontogene mer enhetlig. En mer enhetlig kombinasjon av tog kan også ha bidratt til det har vært mulig å øke frekvensen uten store punktlighetsproblemer.

¹⁶ Basert på Bjerke og Holom (2004)



Figur 11 Utviklingen i antall reisende¹⁷, antall avganger, punktlighet og reisetid. Situasjonen i 1987 er satt til indeks 100 for alle disse forholdene. Antall avganger er telt som ankomster til Oslo S fra Vestfoldbanen, med varierende utgangspunkt (Skien, Larvik, Sandefjord eller Tønsberg)¹⁸

Dersom punktligheten analyseres med høyere oppløsning fremkommer at det har vært store variasjoner fra måned til måned i perioden siden 1987. Punktligheten på Vestfoldbanen avhenger til stor del av situasjonen i Oslo-området og strekningen Oslo-Lillehammer. Siden 1991 går de fleste avganger på Vestfoldbanen også til Lillehammer/Hamar. Eksempelvis henger dårlig punktlighet på Vestfoldbanen i slutten av 2001 sammen med driftsproblemer på mellom Oslo og Lillehammer. JBV Utredning (2004:35) viser at punktligheten for regiontogene på Vestfoldbanen og til Lillehammer henger tett sammen.

Når utbyggingen har et så begrenset omfang som er tilfelle for Vestfoldbanen så er det sannsynligvis andre faktorer som spiller større rolle enn selve jernbaneutbyggingen. For Vestfoldbanen påvirkes antall avganger i stor grad av trafikkomleggingen i tilknytning til åpningen av Gardermobanen.

¹⁷ Tellepunkt Drammen

¹⁸ Før april 1995 ble punktlighet målt som antall tog som er mindre enn fem minutter forsinket. Etter april 1995 er grensen for hva som defineres som forsinkede tog lagt ved tre minutter. Data i figuren er omregnet for å vise andel tog som er mindre enn 3 minutter forsinket i hele tidsperioden. Basert på data fra 2004 for Vestfoldbanen så er omregningsfaktoren satt til 0,967 ved omregning fra punktlighet basert på 5 minutters forsinkelse til 3 minutter. Punktligheten vises som gjennomsnitt for hvert år. Reisetid er målt som gjennomsnitt for alle tog hverdager fra Sandefjord til Oslo.

I Tabell 6 vises en sammenligning av situasjonen når Hovedrapporten NSB (1992a) ble laget, forventet utvikling og faktisk situasjon tre år etter åpning av den senest ferdigstilte parsellen.

Dagens situasjon er at bare deler av moderniseringen er foretatt. Verdiene som beskriver dagens situasjon i Tabell 6 ligger som forventet mellom verdiene fra før utbyggingen og de estimerte verdiene for situasjonen etter fullstendig utbygging.

	Før utbygging (1990/91)¹⁹	Forventet etter²⁰	Faktisk høsten 2004 (13.6.2004-8.1.2005)
Reisetid hele strekningen (Oslo-Larvik) ²¹	2t 14 min	1t 48 min	2t 07 min
Reisetid over ny del av banen (Skoppum-Drammen)	42 min	(35 min) 7 min reduksjon ²²	38 min
Antall tog pr. dag (en retning)	27	40	41
Punktlighet	72% ²³	95%	89% ²⁴
Lengde nye spor		140 km ²⁵	17 km ²⁶
Kostnad		6047 mill NOK (2001)	1 274 mill NOK (2002) ²⁷
Antall reisende	1 478 617	1 708 064 ²⁸ - 3 196 278 ²⁹	1 510 000

Tabell 6 Sammenligning mellom utgangspunktet for utbygging, utvalgte forventede effekter av utbyggingen og situasjonen i 2004 for Vestfoldbanen

¹⁹ Basert på referansealternativet i NSB (1992)

²⁰ "Alternativ A" i NSB (1992)

²¹ Gjennomsnitt antall reisende

²² NSB (1997)

²³ Punktighet innenfor 5 minutter

²⁴ Vanlig punktighetsstatistikk viser 86% punktighet, men da basert på punktighet innenfor 3 minutter. Frem til 1992 ble punktighet beregnet innenfor 5 minutter. Punktighetstallene er derfor konvertert til punktighet innenfor 5 minutter, slik at de er sammenlignbare.

²⁵ Basert på Bruer (1989) og NSB (1992)

²⁶ Basert på www.jernbaneverket.no per 11.11.2004

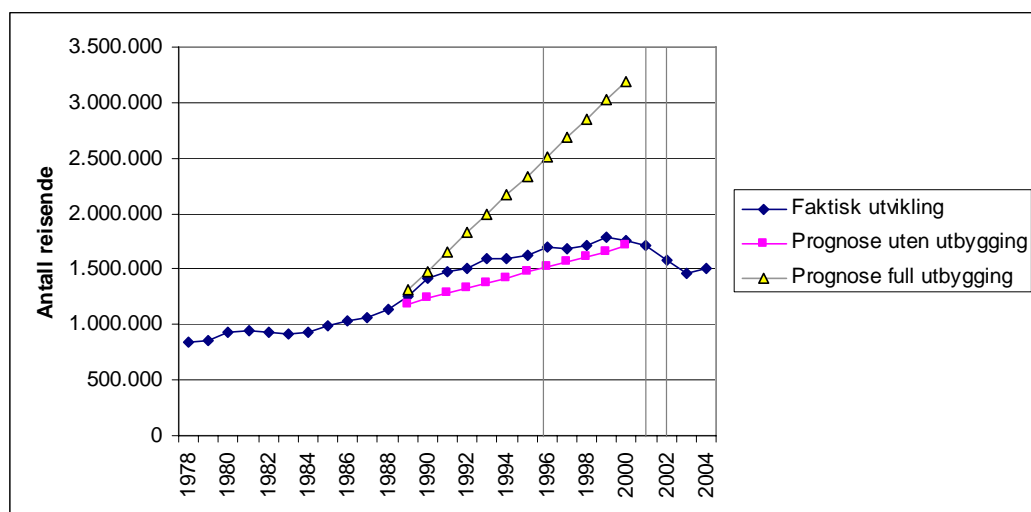
²⁷ Basert på tabell 2

²⁸ Estimert for referansealternativet i år 2000, uten noen utbygging

²⁹ Estimert for år 2000, som er siste året i prognosen. Basert på full utbygging.

Figur 12 viser utviklingen i antall reisende på Vestfoldbanen før, under og etter utbyggingen. Sammenlignet med Alt. A i Figur 12 så er det kun bygget ca 12 % av det planlagte dobbeltsporet. Man kunne derfor forvente at utviklingen i antall reisende skulle være vesentlig mindre enn det som er beregnet for Alt. A, men noe mer enn for basis-alternativet. Figur 12 viser at den faktiske utviklingen i antall reisende til og med 2000 er noe høyere det var som forventet uten utbygging.

Etter 2000 skjer en forholdsvis kraftig nedgang i antall reisende. Årsakene til dette er analysert i JBV Utredning (2004). Som hovedårsaker oppgis utbygging av ny E18, forbedret ekspressbusstilbud, økte billettpriser, redusert tilgjengelighet av billettsalg og forholdsvis stort omfang av forsinkelser.



Figur 12 Utviklingen i antall reisende på Vestfoldbanen. Kilder: 1986-2004: På Sporet³⁰ (1987-2005). Faktisk 1978-1988: Bruer³¹ (1989:12)

6.2.2 Gardermobanen

Gardermobanen ble besluttet som en del av etableringen av ny hovedflyplass på Gardermoen. Kostnadsutviklingen for Gardermobanen er fulgt opp av NOU 1999:28, hvor Tabell 7 er hentet fra. Overskridelsene på utbyggingen er fremst relatert til ekstrakostnader til Romeriksporten (1 300 mill), og til å binde sammen Gardermobanen med Hovedbanen (443 mill).

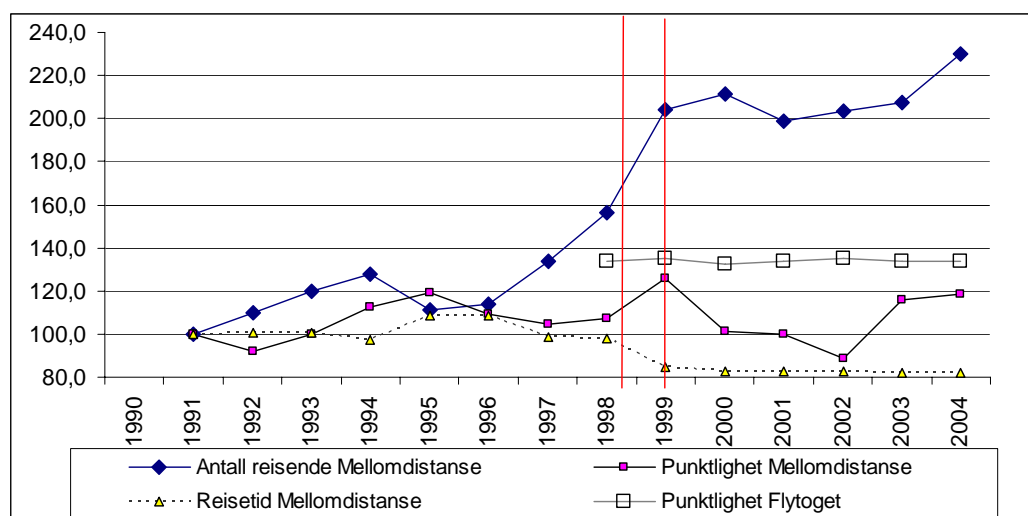
³⁰ Tellepunkt Drammen

³¹ Ikke angitt tellepunkt, men der det er overlapp mellom Bruer og På Sporet så er tallene identiske, noe som tyder på at det også her er tellepunkt Drammen.

Strekning	Åpnet for trafikk	Opprinnelig estimat			Endelig kostnad	
		NOK	År	1998 NOK	NOK	År
Oslo - Gardermoen	1. aug. 1999 ³²	3300	1992	3587	6243	1998
Gardermoen – Eidsvoll	27 aug. 1998	1300	1992	1413	1500	1998
Total				5000	7743	

Tabell 7 Utviklingen i kostnad for Gardermobanen (basert på NOU 1999:28, side 97-99)

Figur 13 viser utviklingen fra 1991 til 2004 for den strekning der Gardermobanen går. Sammenligningen er fremst basert på InterCity/mellomdistanse-togene mellom Oslo S (Skien) og Lillehammer (Hamar) fordi dette er det togproduktet der før-og ettersituasjonen er mest sammenlignbar. I tillegg er punktligheten for Flytoget lagt inn.



Figur 13 Utviklingen i antall reisende³³, punktlighet, og reisetid fra 1991 til 2004 på strekningen der Gardermobanen går. Situasjonen i 1991 er indeksert til 100

³² Lillestrøm-Gardermoen åpnet 27 aug 1998

³³ Tellepunkt Hamar for mellomdistanse

Tabell 8 viser en sammenligning mellom situasjonen før tiltaket, forventet etter, og faktisk situasjon tre år etter åpning.

	Før utbygging (1991)	Forventet etter	Faktisk 2002
Flytoget:			
Reisetid Flytoget		19 min ³⁴	19 min
Antall reisende Flytoget		6 360 ³⁵	3 870
Punktlighet Flytoget		95 % (2 min)	95 % (3 min)
Frekvens		6 tog / time	6 tog / time
IC/Mellomdistanse:			
Reisetid hele strekningen (Oslo-Hamar)	1t 41 min	(1t 21min) 20 min reduksjon	1t 23 min
Reisetid over ny del av banen (Oslo-Eidsvoll)	59 min	(40) 20 min reduksjon	38 min
Antall tog pr. dag (en retning)	11	(11) Ikke angitt endring	21
Punktlighet IC	71 %	Bedring	82 %
Antall reisende ³⁶	410 286	Ikke oppgitt	851 000 ³⁷
Hele banen			
Antall tog (Lillestrøm)	262	460	509
Lengde nye spor		64 km	64 km
Kostnader		5 000 mill NOK	7 743 mill NOK

Tabell 8 Nøkkeltall for Gardermobanen, før tiltaket, estimert og faktisk utfall

³⁴ NSB (1991)

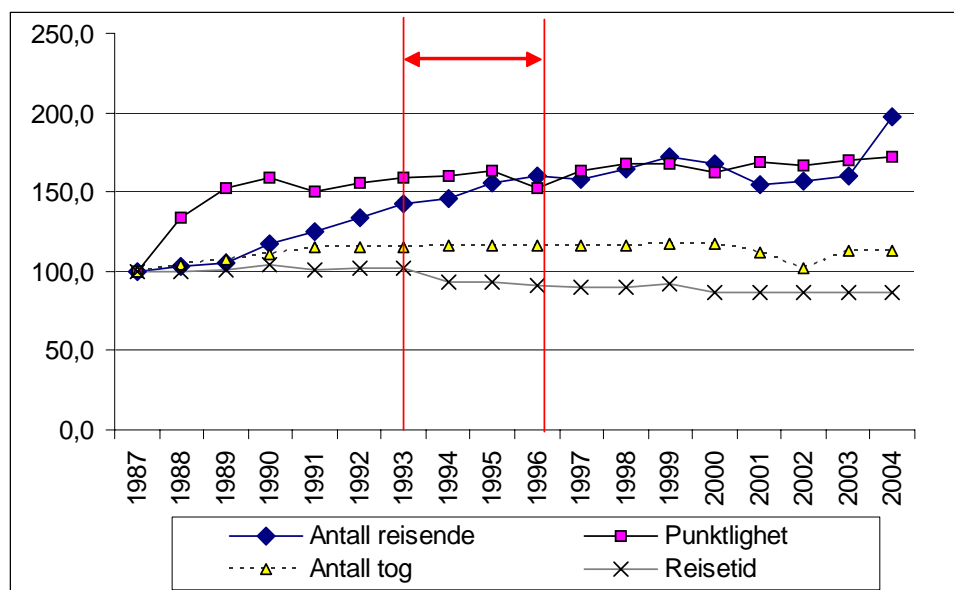
³⁵ St prp 90 (1991-92)

³⁶ Tellepunkt Hamar

³⁷ 2003

6.2.3 Østfoldbanen, Ski-Sandbukta

Dobbeltspor Ski-Sandbukta (Moss) ble bygget mellom 1992 og 1995. De nye parsellene ble tatt i bruk skrittvis fra juli 1993 til september 1996. I tillegg ble det i 2000 åpnet 7 km dobbeltspor sør for Moss, mellom Rygge og Haug, men denne utbyggingen er ikke gjenstand for spesiell analyse i denne rapporten. Figur 14 viser utviklingen i antall reisende, punktlighet, og antall avganger fra 1987 til 2004 på Østfoldbanen.



Figur 14 Utviklingen i antall reisende³⁸, punktlighet, og antall avganger fra 1987 til 2004 på Østfoldbanen. Utviklingen gjelder InterCity/mellomdistanse, med unntak av antall avganger, som er for alle typer tog. Situasjonen i 1987 er indeksert til 100

Berg et al. (1998) inneholder en sammenstilling av den økonomiske utviklingen av prosjektet. Sluttregningen ble på 1 708 millioner NOK (1986). I følge Berg et al. (1998) ble prosjektet igangsatt når Stortinget etterspurte en utredning av dobbeltspor mellom Ski og Moss. Det pågikk da planlegging av et nytt kryssingsspor på strekningen. Dette arbeidet ble derved utvidet til å omfatte hele strekningen. Kostnadsestimatet som ble brukt var meget foreløpig.

Den følgende sammenstillingen av de forventede effektene av utbyggingen er basert på NSB (1986). Utredningen i NSB (1986) omfatter til dels hele dobbeltsporstrekningen mellom Ski og Moss, men ble laget på et tidspunkt hvor det ikke var endelig besluttet om bygging av dobbeltspor på hele strekningen.

³⁸ Tellepunkt Moss for mellomdistanse

Utredningen er derfor fokusert på å identifisere hvilke delstrekninger som bør prioriteres. Da sammenligningsgrunnlaget er mindre så er oppfølgingen av utbyggingen på Østfoldbanen som presenteres i denne rapporten av mindre omfang enn hva som er tilfelle for Vestfoldbanen.

	Før utbygging (1986)	Forventet etter³⁹	Faktisk våren 1998
Reisetid hele strekningen ⁴⁰ (Oslo-Halden)	2t 01 min	(1t 46 min) Reduksjon med 15 min	1t 49 min
Reisetid over ny del av banen (Ski-Moss)	32 min	17 min 40 sek	19 min 45 sek
Antall tog pr. dag (en retning)	88	Økning	102
Punktlighet	50%		88%
Lengde nye spor		35,3 km	32,69 km (dobbelspor Ski- Sandbukta)
Kostnader		300 mill NOK (1986) ⁴¹	1 558 mill NOK ⁴² (1996)
Antall reisende ⁴³	458 724	Ikke angitt	783 066

Tabell 9 Nøkkeltall for Østfoldbanen, før utbygging av dobbeltspor Ski-Sandbukta, estimert etter utbygging og faktisk utfall

6.2.4 Bergensbanen, Finsetunellen og Finse-Gråskallen

Utbyggingen på Bergensbanen skiller seg noe fra de andre prosjektene, da reisetidsreduksjon og kapasitetsutvidelse ikke var de eneste begynnelse for utbygging. Regularitet og vedlikeholdskostnader var også viktige målsettinger med prosjektet.

³⁹Basert på NSB (1986)

⁴⁰ IC/Mellomdistanse

⁴¹ Berg et al. (1998)

⁴² Basert på Hoff og Moen (2000). Deres oppfølgende kost-nytte-analyse er basert på denne verdien. I følge Berg et al. (1998) er 1.563 mill NOK (1995-kr) tildelt prosjektet.

⁴³ Tellepunkt Moss, gjelder mellomdistanse

Det fremgår av dokumentasjonen rundt prosjektene at det hele tiden har vært planlagt med en trinnvis utbygging av strekningene, se eksempelvis (NSB, 1992b). Åpningsdato og kostnadsutvikling for de bygde parsellene fremgår av Tabell 10.

Parsell	Åpnet for trafikk	Første estimat		Endelig kostnad			
		NOK	År	2000 NOK	NOK	År	2000 NOK
Finsetunnelen	16. mai 1993	290	1989	380	294	1993	342
Tunga-Finse	4. okt 1998	296	1994	340	188	2000	188
Gråskallen	19.sep.1999	71	1994	82	281	2000	281
Total		657		801	763		811

Tabell 10 Oversikt over de analyserte byggetrinnene bygde. Kostnadene viser til henholdsvis første og siste gang prosjektene er omtalt enten i statsbudsjett⁴⁴

Tabell 11 viser at reduksjonen i reisetid i 2002 til stor grad er oppfyllet, men punktligheten er kun delvis bedret. Utilfredsstillende punktlighet bidro til at reisetiden er økt i 2003 og 2004., spesielt for noen tog. Dette har gitt bedring i punktligheten. Utviklingen i antall reisende er diskutert av JBV Utredning (2004).

	Før utbygging (1989)	Forventet etter ⁴⁵	Faktisk 2002
Reisetid hele strekningen (Oslo-Bergen) ⁴⁶	6t 23 min	(6t 10min) Reduksjon med 13 minutter	6t 24 min
Reisetid over ny del av banen (Ustaoset-Myrdal)	63 min	(50) Reduksjon med 13 minutter	55 min
Antall tog pr. dag (en retning)	8	(8) Ikke angitt endring	8
Punktlighet	73	78	75
Kostnader		801	811
Antall reisende ⁴⁷	684 227	799 897	623 386

Tabell 11 Nøkkeltall for Bergensbanen, før utbygging av Finsetunnelen og Finse-Gråskallen, estimert etter utbygging og faktisk utfall

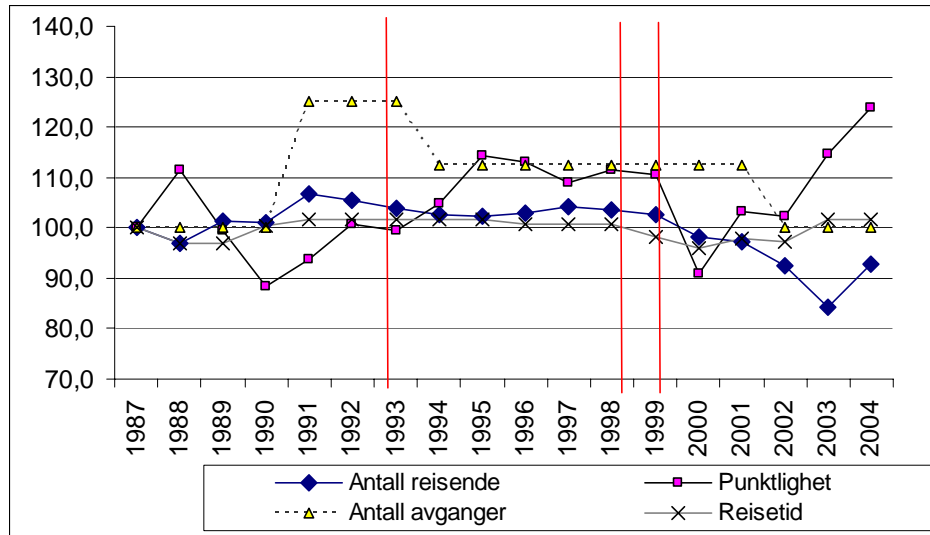
⁴⁴ Første estimat for Finsetunnelen kommer fra NJP 90-93

⁴⁵ Basert på NSB (1991)

⁴⁶ Tog 61

⁴⁷ Tellepunkt Finse

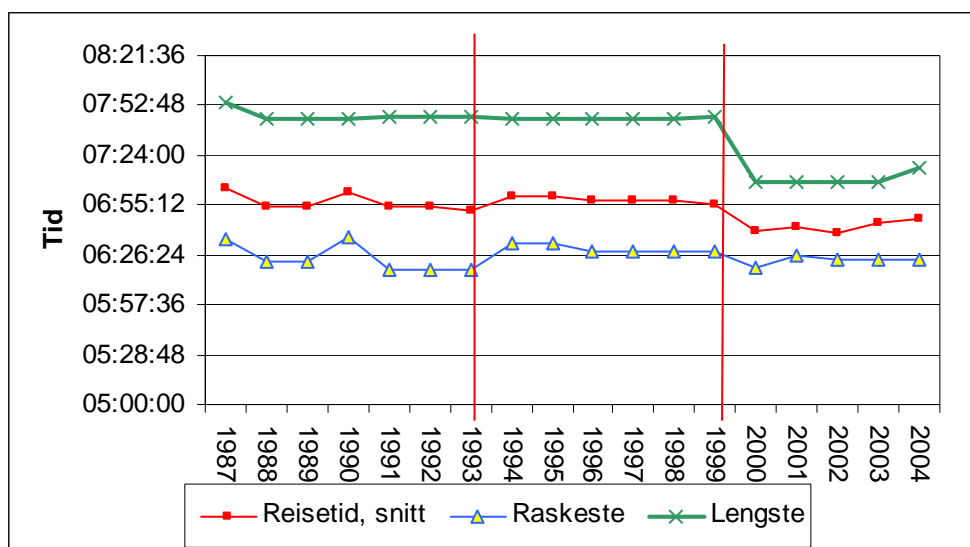
En viktig målsetting med de nye parsellene var å minimere snømengdene på sporet. Som en konsekvens av redusert behov for snørydding er vedlikeholdsbasen på Finse lagt ned (Aspengren, 1999).



Figur 15 Utviklingen i antall reisende⁴⁸, punktlighet, og antall avganger fra 1987 til 2004 på Bergensbanen. Situasjonen i 1987 er indeksert til 100

For Bergensbanen så er måling av reisetid til stor del avhengig av om man tar utgangspunkt i gjennomsnittlig reisetid for alle persontog, eller fokuserer på de raskeste togene. Sammenstillingene i Tabell 11 og Figur 15 er baserte på oppfølging av et tog gjennom hele tidsperioden. Figur 16 viser at de raskeste togene har hatt en marginal reduksjon i reisetid i perioden. Den gjennomsnittlige reisetiden er redusert mer, fordi de langsomme ”dagtogene” har fått kortere reisetid. Utbyggingen ser derfor ut å i større grad ha gitt mer enhetlige reisetider for ulike tog, enn å redusere reisetiden for de raske togene.

⁴⁸ Tellepunkt Finse



Figur 16 Utviklingen i reisetid på Bergensbanens persontog 1987 – 2004, fra Oslo til Bergen

6.2.5 Sammenstilling

For å illustrere effekten av de studerte prosjektene er det laget sammenstillinger, dels av forholdet mellom faktisk og planlagt utfall, dels mellom situasjonen før og etter tiltakene. Tabell 12 viser forskjellen mellom den estimerte og den faktiske utviklingen tre år etter ferdigstilling av prosjektene. Negativ verdi for reisetid indikerer at reisetiden er kortere enn planlagt. Negative verdier for punktlighet innebærer at punktligheten er lavere (færre tog i rute) enn planlagt. Estimerte verdier for Vestfoldbanen refererer til de opprinnelige planene for en samlet utbygging, ikke senere planer for parsellvis utbygging. Sammenlignet med kostnadsestimatene for hver parsell har Vestfoldbanen ca 33% overskridelser.

Prosjekt	Reisetid, ny strekning	Reisetid, hele strekningen	Antall tog	Punktlighet	Antall reisende	Kostnad
Vestfold	9 %	18 %	3 %	-6 %	-53 %	-79 %
Bergen	10 %	4 %	0 %	-4 %	-22 %	1 %
Gardermoen	-3 %	2 %	11 %	0 %	-38 %	55 %
Østfold	12 %	3 %	*	2 %	*	262 %

Tabell 12 Forskjell mellom faktisk utvikling og forventet situasjon, målt i prosent. Forskjellen = (faktisk – forventet) / forventet. * = Forventet antall er ikke oppgitt i den tilgjengelige dokumentasjonen

På lignende måte viser Tabell 13 forskjellen mellom situasjonen før og etter utbyggingene. Hensyn er her ikke tatt til de presenterte estimatene. Negative verdier for reisetid viser kortere reisetid etter enn før utbyggingen. Positive verdier for antall tog indikerer en trafikkøkning. Positive verdier for punktlighet innebærer færre forsinkede tog, det vil si en bedring av punktligheten. For antall reisende innebærer negative verdier at det er færre reisende etter investeringen enn før. Det er ikke funnet estimat over forventet antall tog eller reisende på Østfoldbanen fra tidlige faser av prosjektet, slik at en sammenligning har ikke kunnet gjøres.

Planlagt	Faktisk	Prosjekt	Reisetid, ny strekning	Reisetid, hele strekningen	Antall tog	Punktlig-het	Antall reisende
Samlet		Vestfold	-10 %	-5 %	52 %	24 %	4 %
Parsellvis	Parsellvis	Bergen	-13 %	0 %	0 %	3 %	-9 %
Samlet	Samlet	Gardermoen	-36 %	-18 %	94 %	15 %	108 %
Parsellvis		Østfold	-41 %	-10 %	16 %	76 %	71 %

Tabell 13 Forskjell mellom faktisk utvikling før og etter investeringene, målt i prosent. Forskjellen = (etter - før) / før

6.3 Evaluering av prosjektet for bygging av Frøyatunnelen

6.3.1 Innledning

Dette kapitlet beskriver i korte trekk et eksempel på en konkret analyse av et større samferdselsprosjekt, nemlig Frøyatunnelen i Sør-Trøndelag. Eksempelet sammenfatter etterprøvingen av den undersjøiske tunnelen mellom Hitra og Frøya på Rv 714 i Sør-Trøndelag. Etterprøvingen er gjort etter metodikken i *Veileder for etterprøving av prissatte konsekvenser av store vegprosjekt* (Statens vegvesen, Utbyggingsavdelingen, 2006), som igjen ligger til grunn for tilråkningene om metodevalg i denne rapporten. Eksempelet er basert på rapporten *Etterprøving av prissatte konsekvenser for Frøyatunnelen* (Fordal, 2007), og teksten som omhandler grunnlagsmateriale og beregninger er gjengitt noe forkortet med tillatelse fra Statens Vegvesen, Vegdirektoratet. Noen mer prinsipielle betraktninger er føyd til.

I etterprøvingen av store vegprosjekt skal grunnlag og forutsetninger ved bevilgning etterprøves og ses i forhold til situasjonen som man i dette reelle eksempelet har ca 6 år etter åpning. Normalt vil en etterprøving av de samfunnsøkonomiske analysene finne sted etter 2-3 år. Det kan imidlertid tenkes at prosjektet er gjennomført ut fra

hensyn til næringsliv, sysselsetting og bosetting i prosjektets influensområde. Dette er forhold som det generelt sett er vanskelig å måle opp mot den kontrafaktiske situasjonen (som er utviklingen uten at prosjektet hadde blitt realisert), og hvor det i alle tilfeller er snakk om virkninger som det kan ta vesentlig lenger tid enn 2-3 år å få manifestert. I vårt forslag til metodikk har vi derfor i begrenset grad gått inn på å tallfeste slike virkninger av prosjektet, selv om de kan ha betydelig interesse. En annen grunn for å ikke ta dette med er at de samfunnsøkonomiske virkningene (målt som tilgang og bruk av realressurser i en nasjonaløkonomi) i vesentlig grad blir fanget opp i godt utførte nytte-kostnadsanalyser. Vi mener likevel at der hvor det er en klar målsetting om at prosjektet er motivert ut fra f.eks. regional balanse (som kan bety at økonomisk aktivitet blir overført fra en region til en annen) så bør dette etterprøves, men da må høyst sannsynlig etterprøvingen skje tidligst 5 år etter ferdigstillelse.

Den nye vegforbindelsen mellom de to øyene Hitra og Frøya i Sør-Trøndelag (Rv 714) ble åpnet 23. juni 2000. Det var tidligere en ferjeforbindelse her, Kjerringvåg (Hitra) – Flatval (Frøya), med 343 ÅDT i siste hele år (1999) med ferje. Ferjeforbindelsen (ekskl. tilførselsveger) var 5 km lang. Vegforbindelsen består av en 5,3 km lang undersjøisk tunnel under sundet mellom Hitra og Frøya, den såkalte Frøyatunnelen, samt tilførselsveger fra eksisterende hovedveger på Dolmøya (Vågen) på Hitrasiden og Hammarvika på Frøya.

Hitra–Frøya-forbindelsen var det siste av tre prosjekter som i løpet av 1990-tallet ble realisert i dette området. De to andre prosjektene var tunnelen mellom Hitra og fastlandet (åpnet 1994), og bru til Fjellvørsøya (1992). Hovedmålet med vegprosjektene var å bedre framkommeligheten ved å erstatte ferjeforbindelser med faste vegforbindelser, og at dette skulle gi gunstige regionale virkninger i dette kystområdet i Sør-Trøndelag.



Figur 17 Oversiktskart

Prosjektene ble finansiert gjennom et felles bompengeprojekt, i form av forskuddsinnkreving på de tre ferjeforbindelsene som nå har fått vegforbindelse, og etterskuddsinnkreving. Bomstasjonen ved Frøyatunnelen var i drift i ca 2,5 år, fram til 1. februar 2003 da man vedtok å overføre videre nedbetaling av anlegget til bomstasjonen ved den andre tunnelen, Hitratunnelen som forbinder Hitra med fastlandet. Pr. i dag ser bompengerperioden ut til å vare til ut 2008 eller 2009, da anlegget (ekskl. bru over Dolmsundet) er nedbetalt. Endringen av innkrevingsopplegget innebar altså trafikken mellom Hitra og Frøya etter 1. februar 2003 slapp å betale bompenger, og at det kun var trafikk til/fra fastlandet som fortsatt måtte betale bompenger (med uforandrede takster, men med en forlengelse av bompengerperioden som man i 2002 antok ville bli 3 – 4 år).

6.3.2 Stortingsvedtak om bygging, prosjektomtale i Stortingsproposisjoner

Endelig vedtak om byggingen av prosjektet ble gjort av Stortinget den 4. desember 1997, og omfattet følgende:

5,3 km lang undersjøisk tunnel, mellom Norddolm på Hitra og Hammarvika på Frøya, samt 2,7 km riksveg på Hitra og Frøya inkl. ny bru over Dolmsundet, utbyggingen delfinansieres med bompenger.

Byggingen startet høsten 1997 (forberedende arbeider), og arbeidet med tunnelen startet i januar 1998. På bakgrunn av den omfattende kartleggingen av geologiske

forhold i Frøyfjorden, og den komplekse geologien som ble påvist, ble det funnet at det kunne være en stor risiko for kostnadsoverskridelser i prosjektet. Stortinget ble informert om risikoen for kostnadsoverskridelser. Stortingets vedtok at beslutning om tidspunktet for bygging av brua over Dolmsundet (med tilhørende riksveg) skulle utsettes til etter at man hadde gjort gjennomslaget for tunnelen under Frøyfjorden.

Med bakgrunn i Stortingsproposisjon 76 for 2000-2001 ble det i mai 2001 endelig bestemt å utsette byggingen av bru over Dolmsundet til etter 2002. Begrunnelsen for dette var de uavklarte planene vedrørende kryssingen av sundet. Riksvegmidlene som var forespeilet prosjektet i 2001 (59,3 mill. kr.) ble bestemt å skulle omfordes til andre riksveger i Sør-Trøndelag.

Grunnlaget for Stortingets vedtak om å bygge Frøyaforbindelsen var Stortingsproposisjoner fra 1989 til 2002 (sistnevnte omhandler endring av bompengerekravet).

6.3.3 Planer, historikk og omfang

I 1989 ble hovedplan for "Fastlandsforbindelsen til Frøya" godkjent. I hovedplanen var prosjektstrekningen delt i to byggetrinn (I og II). Byggetrinn I (senere kalt alternativ 1 i kommunedelplanen) omfattet at eksisterende Vettastraumen bru fortsatt skulle benyttes, at det skulle bygges ca 3 km ny riksveg fra denne brua og til tunnelpåhugget på Vågen, deretter tunnel til Frøya og tilknytning til Rv 716 i et nytt kryss ved Hammarvika. Byggetrinn II (ny kryssing av Dolmsundet på Hitra) er pr. i dag ikke realisert. Vi går ikke nærmere inn på en detaljert beskrivelse av utviklingen i selve planverket. Generelt skal beskrivelsen romme alle de elementer som har betydning for å kunne gjennomføre ex post-evalueringen i henhold til vedtatt og gjennomførte planer.

Prosjektet fikk første bevilgning i 1997, 2,7 mill. kroner i statlige midler og ca 24 mill. kr. bompenger/ penger fra fylke og kommune. I 1998, da arbeidet tok til på ordentlig, ble i alt 141 millioner fra fylket og bompenger stilt til disposisjon. Første store statlige bevilgning var i 1999, med 36 mill. kr. I omtalen av prosjektet i Stortingsproposisjon 1 1997/98 ble det presisert at Dolmsundet bru var en del av prosjektet og at brua også var med i det samlede kostnadsoverslaget for prosjektet (540 mill. kr. i 1998).

I stortingsproposisjonen fra 1997/98 ble det sagt at beslutningen om tidspunkt for start av bygging av Dolmsundet bru skulle utsettes til etter at man hadde fått gjennomslag i tunnelen. På den måten ville man ha noe å gå på kostnadmessig i tilfelle en kostnadsoverskridelse (som man fryktet pga. usikker geologi). I etterprøvingssammenheng skal prosjektet slik det foreligger ved første bevilgning legges til grunn. I begge stortingsproposisjonene (1996/97 og) var det forutsatt at ny kryssing av sundet skulle være med i prosjektet, selv om prosjektets total kostnad var usikker. Pga. denne usikkerheten valgte man i 1997 å utsette (men ikke ta ut) byggingen av bru over Dolmsundet og tilstøtende veg.

Ut fra denne historikken legger vi derfor til grunn at kryssing av Dolmsundet er en del av prosjektet som etterprøves, dvs. at prosjektet starter ved Storstvatnet/ Melkvika på Hitra. Da Frøyatunnelen, dvs. fra Vågen på Dolmøya og til Hammarvika på Frøya åpnet sommeren 2000, så man for seg at arbeidet med bygging av riksveg og bru over Dolmsundet kunne fortsette og fullføres tidligst i 2002. Prosjektet som er bygd pr. 2007, omfatter som tidligere nevnt ikke kryssing av Dolmsundet/ Vettastraumen. Følgende er bygd (lengder er oppgitt i etterprøvingsskjemaet):

- Ny riksveg fra Vågen på Dolmøya og til Hammarvika (Rv 716) på Frøya, inkl. 5,3 km tunnel.
- Utbedring av Fv 385, inkl. et nytt kryss, mellom Kjerringvåg (på Dolmøya) og Vågen (ved start på den nye tunnelen) slik at denne kan fungere som midlertidig riksveg (i dag er dette Rv 714) inntil permanent løsning avklares og blir bygd.
- Utbedring av 7 km kommunale og private veger (dette i følge sluttrapporten for anlegget, vi har ikke oversikt over hvilke veger dette gjelder).
- Gang- og sykkelveg langs Rv 716 på Frøya (Hammarvika – Nabeita)

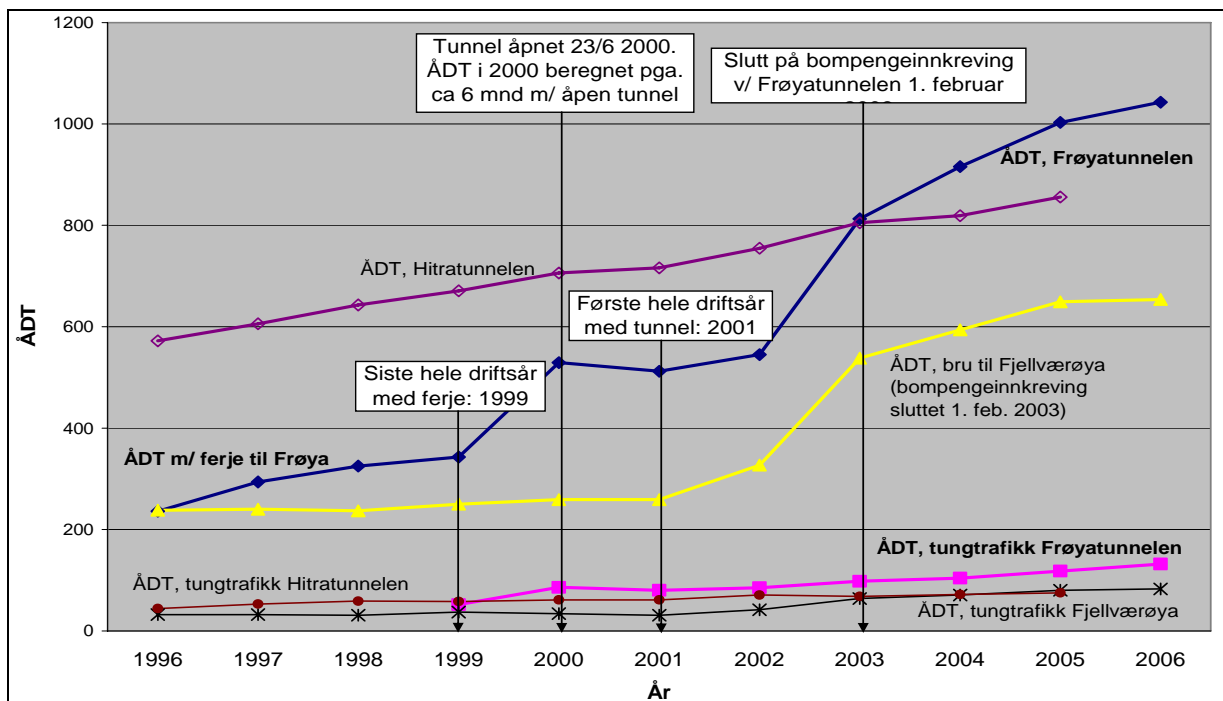
6.3.4 Trafikkgrunnlag og nyskapt trafikk

Denne delen av analysen er helt sentral for å kunne etablere et sammenligningsgrunnlag for å kunne vurdere situasjonen med og uten prosjektet. Derfor har vi vist framgangsmåten relativt detaljert. Metodikken for trafikkanalysen kan imidlertid variere. I framtidige prosjekter av denne størrelsen er det grunn til å tro at transportmodellene som skal ligge til grunn for utarbeidelse av Handlingsprogrammet i NTP vil bli benyttet. Da bør samme modellkonsept bli lagt til grunn for ex post-analysene, der en også kan fange opp og korrigere for effekter av tiltak i tilstøtende transportnett som kan påvirke trafikkvolumene.

Trafikktallene fra bompengeselskapet, Hitra Frøya Fastlandssamband AS (HFF), viste at trafikken siste hele år med ferje (i 1999) var ÅDT 343 kjøretøy pr. døgn. Tungtrafikkandelen samme år var ca 15 %.

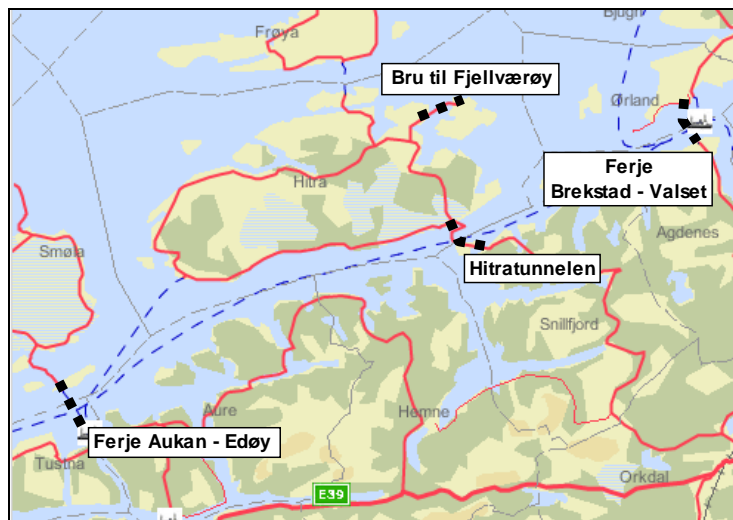
Trafikken i Frøyatunnelen telles kontinuerlig, og gir pålitelige tall for trafikkutviklingen i tunnelen siden åpningen 23. juni 2000. Trafikken i de ca 6 månedene av 2000 med tunnelen i bruk, er omregnet til ÅDT 529 kjøretøy pr. døgn. I første hele driftsår, i 2001, var ÅDT 512 kjøretøy pr. døgn. Etter at bomstasjonen ble fjernet tidlig i 2003 økte trafikken i løpet av kort tid, fra ÅDT 545 i 2002 til 813 ÅDT i 2003, ÅDT 916 i 2004, 1003 i 2005, til foreløpig ÅDT i 2006 på 1043 kjøretøy pr. døgn. Det er viktig å ha en kritisk holdning til holdbarheten i tellingene. I dette prosjektet foreligger det telledata som har begrenset verdi på grunn av varigheten på telleperiodene, og som derved er utelatt fra analysen.

Tellepunkter som gir kontinuerlige data finnes ved Fjellværøya (Krabbsundbrua) og i Hitratunnelen. Krabbsundbrua åpnet i 1992, Hitratunnelen i 1994. På Krabbsundbrua var det også en bomstasjon, denne ble tatt ned samme dag som bommen ved Frøyatunnelen. Trafikkutviklingen på de tre stedene fra og med 1996 er vist i figuren nedenfor.



Figur 18 Trafikkutviklingen i Frøyatunnelen og på bru til Fjellværøya og i Hitratunnelen

I en ex post-evaluering utgjør den kontrafaktiske situasjonen (situasjonen uten prosjektet) sammenligningsgrunnlaget. Mye av dokumentasjonen nedenfor er knyttet til å etablere et slikt sammenligningsgrunnlag. For å se trafikkvekst som ville ha kommet uavhengig av tunnelen, er trafikken i Frøyatunnelen sammenlignet med Hitratunnelen og to ferjeforbindelser⁴⁹ i nærheten av Hitra og Frøya med ca samme trafikkgrunnlag som Frøyaferja. Den ene ferja, Aukan – Edøy⁵⁰ i Møre og Romsdal, har ÅDT 150 – 200 kjøretøy pr. døgn. Den andre, Brekstad – Valset⁵¹ i Sør-Trøndelag, har ÅDT ca 300.



Figur 19 Strekninger som er benyttet som referansestrekninger for trafikkutvikling

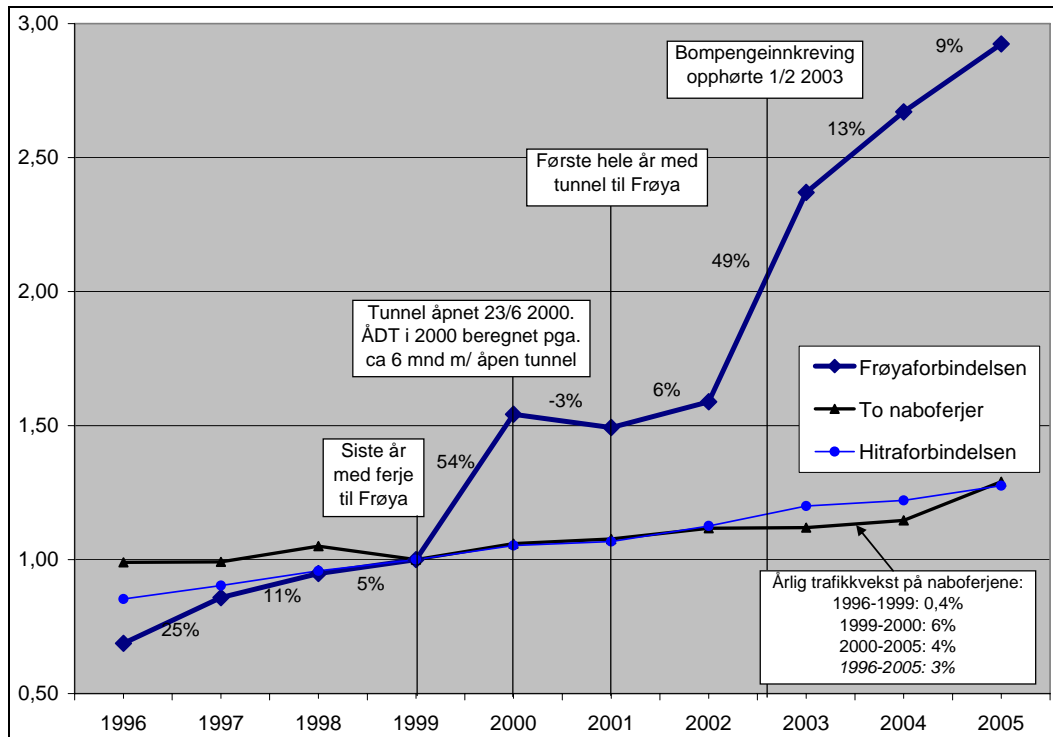
Trafikkutviklingen er vist nedenfor. Tallene for de to ferjene er slått sammen, og benyttet som referanse for generell trafikkvekst for ferjetrafikken. Trafikkutviklingen er vist i forhold til referanseåret 1999 (1999 = 1,0). Figuren viser også årlig trafikkutvikling målt i % pr. år.

Trafikken på Frøyaforbindelsen økte med 54 % i løpet av første år med tunnel. På de to referansestrekningene med ferjer var økningen 6 %, mens Hitratunnelen økte med 5 %. Hvis vi legger generell trafikkvekst for de to ferjene til grunn (6 %), og antar at denne veksten også hadde kommet på Frøyaforbindelsen, vil nyskapt trafikk første år med tunnel være $54 \% - 6 \% = 48 \%$ av trafikken i 1999.

⁴⁹ Kilde: Vegvesenets ferjestatistikk for 1999 – 2005.

⁵⁰ Ferjestrekningen (Rv 669) er 8 km, ferja er den ene av to ferjer som må benyttes for å komme fra fastlandet til Smøla

⁵¹ 5,7 km lang ferjestrekning (Rv 710) ytterst i Trondheimsfjorden



Figur 20 Trafikkutvikling på Frøyaforbindelsen, Hitratunnelen og to andre ferjeforbindelser i samme område (referanseår 1999 = 1,0)

Fra 2000 til og med 2002 var trafikkveksten i Frøyatunnelen ca 3 %, dvs. litt under veksten på de to ferjene. I de to første⁵² årene uten bompenger, i 2003 og 2004, økte trafikken betydelig mer i Frøyatunnelen enn på de to ferjene. Veksten på Frøya var 49 % første år uten bom, og 13 % andre år uten bom (2004). Totalt ga dette en trafikkvekst⁵³ på 68 % disse to årene, mens referanseferjene hadde trafikkvekst på ca 8 % (4 % pr. år) i samme periode. Dette betyr at bortfallet av bompenger ga en nyskapt trafikk på ca 60 % i forhold til trafikken i 2002.

Trafikken i Hitratunnelen fra 1996 til 1999 hadde større vekst enn de to ferjene. Trolig skyldes dette at det er relativt kort tid etter Hitratunnelen åpnet (1994), og at man derfor hadde en trafikkvekst som var større enn generell trafikkvekst. At ferja til Frøya hadde sterkere vekst enn de to naboferjene i samme perioden, skyldes trolig trafikkveksten fra den nye Hitratunnelen.

⁵² Bompenggeinnkrevingen sluttet 1. februar i 2003. Våre trafikk tall for 2003 er ikke korrigert for dette, noe som betyr at ÅDT for første år uten bompenger er litt lavere enn hva det virkelig skal være.

⁵³ Trafikkvekst på 49% det første året og 13% det påfølgende året gir at veksten i løpet av de to årene blir $1,49 \times 1,13 = 1,68$, dvs. 68%

I nytteberegninger av ferjeavløsningsprosjekter er nytten av nyskapt trafikk vanligvis en betydelig del av den totale nytten. Ut fra trafikkutviklingen i figuren i forrige kapittel har vi valgt å definere den nyskapte trafikken i to komponenter:

- Nyskapt trafikk 1: Kommer som følge av at fastlandsforbindelsen blir etablert, og i løpet av det første året⁵⁴ (2000) etter at tunnelen er åpnet.
- Nyskapt trafikk 2: Kommer som følge av at bompengeneinnkrevningen ved Frøyatunnelen opphører, og i løpet av de første to årene⁵⁵ (2003 og 2004) etter dette.

Kvantifiseringen av den nyskapte trafikken blir da:

Nyskapt trafikk 1: Trafikkøkningen fra siste driftsår med ferje til første driftsår med tunnelen (54 %), minus generell trafikkvekst (6 %). Dette gir: $48 \% \cdot \text{ÅDT}_{1999} = 0,48 \cdot 343 = \text{ÅDT } 165$.

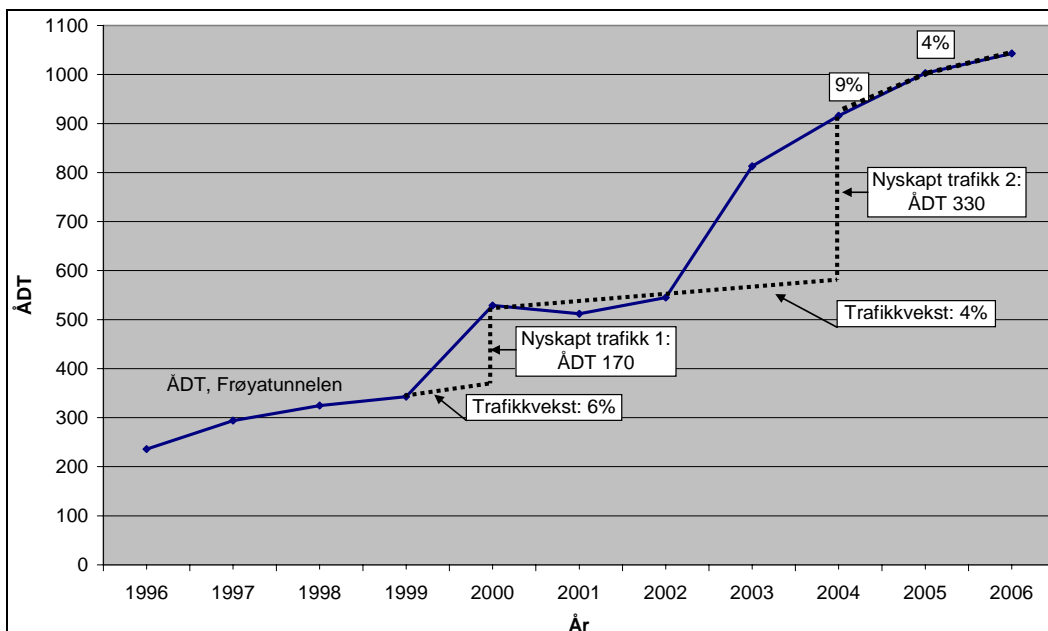
Nyskapt trafikk 2: Trafikkøkning fra siste driftsår med bomstasjon (2002) og i løpet av de to første år uten bompenger (2003 og 2004). Total vekst i denne perioden var 68 %, mens tilsvarende vekst på de to ferjene var 8 %. Dette gir da nyskapt trafikk på 60 % av ÅDT i 2002: $0,6 \cdot \text{ÅDT}_{2002} = 0,6 \cdot 545 = \text{ÅDT } 326$.

Ut fra denne analysen blir trafikkutviklingen som vist i neste figur. Figuren viser årlig trafikkvekst, samt de to store sprangene i form av nyskapt trafikk i 2000 og i 2004. Den årlige generelle trafikkveksten er mellom 4 % og 9 % pr. år.

Dersom vi antar at tunnelen ikke hadde bompenger, og at all nyskapt trafikk 1 og 2 kom samtidig, ville samlet nyskapt trafikk i så fall ha vært $\text{ÅDT} \text{ (avrundet)} = 165 + 326 = \text{ÅDT } 491$.

⁵⁴ 2000 er benyttet år som referanseår, selv om tunnelen var bare åpen for trafikk fra 23. juni (ca 6 mnd) dette året. ÅDT for 2000 er tall for 6 mnd, omregnet til ÅDT.

⁵⁵ 2003 er betraktet som et "helt" år uten bompenger, noe som betyr at vi underestimerer ÅDT i 2003 med ca ÅDT 20-30 (og dermed den nyskapte trafikken tilsvarende)



Figur 21 Trafikkutviklingen 1999 – 2006 som skal legges inn i nytteberegningene. Figuren viser avrundede trafikktall

I tabellen under er nyskapt trafikk skilt ut fra den generelle trafikkveksten. Dette er gjort for lette og tunge kjøretøy. I løpet av første år med tunnel (her er 2000 regnet som første år med tunnel) er nyskapt trafikk ÅDT 165, derav 31 tunge. I 2004 (to år etter at bom ble fjernet) er total nyskapt trafikk ÅDT 491, derav 42 tunge. Ut fra dette ser man at bortfall av bompenger ga en mindre vekst i nyskapt tungtrafikk (fra 31 til 42 ÅDT) enn for total nyskapt trafikk som vokste fra ÅDT 165 til ÅDT 491.

ÅR	ÅDT Frøya	ÅDT u/nyskapt	Nyskapt, totalt	ÅDT tunge Frøya	ÅDT-t, u/nyskapt	Nyskapt ÅDT-t
1999	343			52		
2000	529	364	165	86	55	31
2001	512	378	134	80	57	23
2002	545	393	152	85	58	27
2003	813	409	404	98	60	38
2004	916	425	491	104	62	42
2005	1003	464	539	118	68	50
2006	1043	491	552	132	72	60

Tabell 14 Trafikkutviklingen for tunge og lette kjøretøy og for nyskapt/ ikke-nyskapt trafikk

Det er flere faktorer som bidrar til at det er vanskelig å beregne nyskapt trafikk. Ikke minst gjelder dette hvordan man skiller mellom nyskapt trafikk og generell trafikkvekst.

Forhåndsinnkreving av bompenger på ferjene (takstøkning 40 % på ferjebillettene fra 1. januar 1991) før tunnelen åpnet kan i teorien ha bidratt til å dempe ferjetrafikken. Dette betyr altså at man kan ha hatt en ”unaturlig” lav ferjetrafikk. Dermed ble spranget fra siste år med ferjetrafikk og de første årene biltrafikk i tunnelen tilsvarende ”unaturlig” høy. Vi har ikke undersøkt dette nærmere, og det er vanskelig å anta hvilken betydning dette har hatt.

Åpningen av Hitratunnelen i 1994 antas også å ha hatt betydning for trafikken til/fra Frøya, mest sannsynlig i form av økt trafikk både før og etter åpning av Frøyatunnelen. Men det er vanskelig å tallfeste denne trafikken. Trafikk mellom Frøya og fastlandet må i dag betale bompenger på Sandstad (ved Hitratunnelen). Når denne innkrevingen opphører i løpet av kort tid vil man trolig få en ny trafikkvekst både i Hitratunnelen og Frøyatunnelen. Dette vil mest sannsynlig føre til at trafikkveksten de første årene etter at bomstasjonen på Sandstad fjernes blir større enn det vi har lagt til grunn i etterprøvingen.

6.3.5 Nyttekostnadsanalyser

Analyseopplegget er basert på gjeldende metodikk for samfunnsøkonomiske analyser. Det er viktig å påse at metodikken for ex ante og ex post-vurderinger er konsistent, slik at utfallene lar seg sammenligne. I de fleste tilfeller betyr dette at ex post-evalueringene må kjøres med samme metodikk som ex ante-analysene. Dette kan innebære noen utfordringer, fordi en bør ha mulighet for å sjekke at ex ante-analysene er kjørt med et velbegrunnet datagrunnlag. Det er imidlertid grunn til å tro at dette problemet vil bli mindre i kommende ex post-evalueringer på grunn av de kravene som nå settes til ex ante-analysene. Vi anbefaler for ordens skyld at man i framtidige ex ante-analyser spesifiserer og arkiverer underlagsmaterialet slik at man har mulighet for å revidere analysene med ny metodikk når ex post-evalueringen skal gjennomføres.

Beskrivelsen av framgangsmåten er, i likhet med beskrivelsen av trafikkgrunnlaget, gitt i noen detalj. Eksempelet er som nevnt hentet fra ”virkeligheten”, og det framgår hvilke problemer som kan oppstå dersom ex ante-analysene ikke er fullstendig dokumentert.

Nyttekostnadsanalyse ved valg av løsning

Nyttekostnadsanalysen er gjort i 1987 med EFFEKT4. Vi har ikke hatt tilgang til denne beregningen eller forutsetningene bak den. Følgende framgår om nyttekostnadsanalysen i hovedplanen:

Nyttekostnadstallet, N/K, for alternativ ”øst” er 3,2. Utbyggingskostnad var ca 270 mill. kr. Denne omfatter trolig ikke kryssing av Dolmsundet, men vi har ingen sikker dokumentasjon på dette. Det var forutsatt 3 % trafikkvekst (antar at dette gjelder beregningsperioden på 25 år i nytteberegningen). I beregningene antok man at trafikken i åpningsåret var ÅDT 600 (vi er usikre på åpningsåret, men 1993 er nevnt som åpningsår i hovedplanen). Hvis trafikkprognosen på ÅDT 600 er for 1993, vil dette bety at trafikken ble forutsatt tredoblet (gitt at ferjetrafikken i 1993 var ca ÅDT 200) i åpningsåret.

Resultatet, $N/K=3,2$, fra den første nyttekostnadsberegningen viste et svært lønnsomt prosjekt, dvs. at man for hver investert krone ville få 3,2 tilbake.

Nyttekostnadsanalyse ved bevilgning, 1997

Beregningen er gjort med utgangspunkt i database fra EFFEKT5.25 som ble forelå ved bevilgning i 1997. Beregningsår og åpningsår er 2001 (i opprinnelig beregning: 2002).

Utbyggingskostnad: 533,7 mill. kr (prisnivå 1997), budsjettandel 51%. Det ser ut som om man i opprinnelig beregningen kun la inn statsandelen, og utelot andelene fra fylkeskommune/ kommune (som utgjorde 11%, dvs. at offentlig andel i sum ble 62%).

Trafikkvekst: Ved bevilgning var det forutsatt 1,2% årlig vekst for lette og 1,4% for tunge i hele perioden

Tungtrafikkandel: Forutsatt tungtrafikkandel 10%

Nyskapt trafikk

Forutsatt trafikk i førsituasjonen, dvs. på ferjestrekningen (vegnett 0) mellom Hitra og Frøya i 2002 var:

ÅDT-lette=272, ÅDT-tunge=30, dvs. ÅDT⁵⁶ 302 kjøretøy pr. døgn (10% tunge)

Nyskapt trafikk ble anslått ut fra trafikkutviklingen i Hitratunnelen som åpnet i 1994. Her ble det observert ca 35 % vekst fra siste ferjeår til første år med tunnel. Den nyskapte trafikken ble fordelt på to strekningene, Frøya – Hitra og Frøya – Trondheim. Følgende nyskaping av trafikk ble forutsatt:

- Frøya – Hitra 30 % trafikkøkning, både for lette og tunge kjøretøy
- Frøya – Trondheim 20 % trafikkøkning, for lette og tunge kjøretøy

Dette ga, fordelt på reiserelasjoner og forutsatt tungtrafikkandel på 10 %:

	Kjøretøy pr. døgn	Lette kjt. pr døgn	Tunge kjt. pr døgn	Nyskapt (%)	Nyskapt (ÅDT-lette)	Nyskapt (ÅDT-tunge)	Nyskapt (ÅDT)
ÅDT-Frøya-Trheim	272	245	27	20 %	49	5	54
ÅDT-Frøya-Hitra	30	27	3	30 %	8	1	9
ÅDT-tot	302	272	30		57	6	63

Tabell 15 Trafikkgrunnlag for beregning ved bevilgning, dvs. trafikk på ferjestrekning og nyskapt trafikk fordelt på reiserelasjon og på tunge/lette kjøretøy. ÅDT 2001

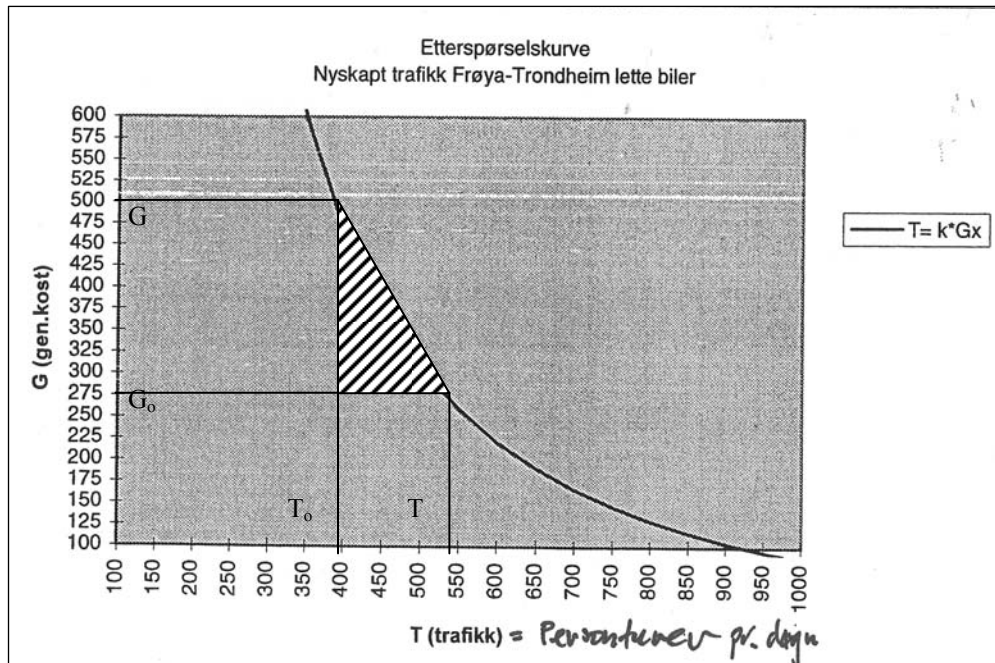
Innkrevingen av bompenger ble 1. februar 2003 flyttet til bomstasjonen ved Hitratunnelen. I 1997 var man ikke klar over at dette ville skje (vedtaket ble gjort i november 2002). Opprinnelig var bompengeperioden forutsatt å gå til 2011 eller 2012. Nyskapt trafikk som følge av at bommen ble tatt ned er altså ikke med i beregningene som forelå ved bevilgning i 1997.

Nytten av nyskapt trafikk ble i 1997 gjort beregnet ved bruk av etterspørselskurver som ble ”konstruert” for hhv. tunge og lette kjøretøy for de to reiserelasjonene. Utgangspunktet for å lage kurvene var at man forutsatte en trafikkvekst tilsvarende det man hadde for Hitratunnelen som åpnet i 1994 (dvs. ca 30% engangsvekst). For dokumentasjon av etterspørselskurvene vises til dokumentasjonsrapporten fra Fjellanger Widerøe fra 1997 for nyttekostnadsberegningene for Frøyaforbindelsen.

⁵⁶ Trafikkberegningen i EFFEKT5 for 2002 viser at ÅDT er 292, mens man som grunnlag for beregning av nyskapt trafikk opererer med ÅDT 302 i 2002, det er uklart hva dette skyldes (mest trolig en feil)

Nytte av nyskapt trafikk er vist som skravert område i figuren nedenfor, dvs.:

Nytte av nyskapt trafikk pr. døgn: $NN \text{ døgn} \approx \frac{1}{2} * (G_0 - G) * (T - T_0)$



Figur 22 Etterspørselskurve for personbiltrafikk (antall personturer). Skravert område er nytte av 140 nyskapte personturer, gitt en vekst fra 400 til 540 personturer pr. døgn

I kurven er det lagt inn et eksempel for å vise hvordan nytten kan beregnes. Dersom trafikk i førsituasjonen (T_0) er 400 personturer pr. døgn, og denne vokser til $T = 540$ personturer pr. døgn etter utbygging, er nyskapt trafikk 140 personturer pr. døgn. Veksten har kommet som følge av at generalisert reisekostnad pr persontur for denne reisetrekningen har gått ned fra 500 kr. pr. døgn (G_0) til 275 kr. pr. døgn (G), dvs. en besparelse på 225 kr. Nytt av nyskapt trafikk pr. år blir i dette beregningseksempelet blir da:

$$NN \text{ år} \approx \frac{1}{2} * (500 - 275) * (540 - 400) * 365 = \frac{1}{2} * 225 * 140 * 365 = 5,75 \text{ mill. kr. år.}$$

Med denne metoden er årlig nytte av nyskapt trafikk på Frøyaforbindelsen beregnet:

- Frøya – Hitra, lette kjøretøy: 2376,1 kr/døgn
- Frøya – Hitra, tunge kjøretøy: 3395,8 kr/døgn
- Frøya – Trondheim, lette kjøretøy: 258,8 kr/døgn

- Frøya – Trondheim, tunge kjøretøy: 810,6 kr/døgn

Sum nytte pr. dag (1997-kr) 6840,6 kr/døgn

Sum nytte pr år: 2,5 mill. kr/år

Lagt inn i EFFEKT som engangsnytte 2002 (2001-kr): 3,05 mill kr/år

Nytte av nyskapt trafikk utenfor prosjektområdet

I 1997 forutsatt man at all nyskapt trafikk (ÅDT 65) er mellom grensepunkt 1 og 4 (1 er Hammarvika på Frøya, 4 er ved prosjektets start på Hitra), og at denne trafikken kommer fra Trondheim som er 148 km unna. Dette gir en lang kjøredistanse utenfor prosjektområdet, noe som innebærer økt ulykkesrisiko og økt vegvedlikehold. EFFEKT kan beregne kostnadene for dette (antatt ulykkesfrekvens er da 0,19).

Nyttekostnadsanalyse ettersituasjon, 2006

Nyttekostnadsanalysen er gjennomført med samme metodikk og programversjon (EFFEKT5.63) som for bevilgningssituasjonen. For våre beregninger har vi følgende endringer i forhold til bevilgningssituasjonen:

Utbyggingskostnad: I sluttrapporten⁵⁷ for anlegget (19. mars 2003) er utbyggingskostnaden 459,6 mill. kr (2001-kr). Vi har lagt denne til grunn i etterprøvingen. Da prosjektet (bompengeinnkreving og ferdigstilling av veg/bru over Dolmsundet) ennå ikke er avsluttet er ikke endelig oppgjør i forhold til bomselskapet gjort. I etterprøvingen er det forutsatt at bompengeandelen blir som angitt i stortingsprop. 76 fra 2001, dvs. 50 % fra offentlige budsjett, dvs. at sum andeler fra stat, fylkeskommune og kommuner er 50 %.

Sammenligningsår/åpningår: Anlegget ble åpnet ca 1 år før det som var forutsatt ved bevilgning (2002). Faktisk åpningsår = sammenligningsår = første hele kalenderår etter åpning: 2001.

Trafikkvekst: I ettersituasjonen er det lagt til grunn observert trafikkutvikling for lette og tunge til og med 2006. Trafikkutviklingen, dvs. antatt trafikkutvikling gitt at man ikke har nyskapt trafikk, er 6 % vekst fra 1999 til 2000, 4 % årlig vekst fra 2000 til og med 2004, deretter 9 % vekst fra 2004 til 2005, og 6 % vekst fra 2005 til 2006.

⁵⁷ Kostnaden i sluttrapporten er oppgitt i 2000-kr, 445,8 mill.kr.

Fra 2006 og ut beregningsperioden er det lagt inn samme vekst som i beregning ved bevilgning, dvs. 1,2 % årlig vekst for lette og 1,4 % for tunge.

Tungtrafikkandel: Ved bevilgning var det forutsatt tungtrafikkandel 10 %, mens faktisk tungtrafikkandel var 15,5% de første tre årene (2000-2002), deretter ca 12 %. Vi har lagt til grunn 12 % tungtrafikkandel i ettersituasjonen.

Virkelig kjørehastighet: Kjøretidsregistreringen er fra prosjektets start (ved Melkvika) på Hitra og til prosjektstrekningens slutt (krysset Rv 714/716) på Frøya. Strekningen er 13,7 km. Hastighet fra EFFEKT er 65,7 km/t (for lette kjøt), noe som ga kjøretid 13 min og 2 sek. Kjøretidsregistreringene ga et snitt på 12 min og 19 sek. Hastigheten på lenke 133-10 ble etter dette korrigert til 68km/t, og dette ble lagt inn som konstant hastighet for lette kjøretøy i EFFEKT. For tunge kjøretøy og på øvrige veglenker er beregnet hastighet fra EFFEKT benyttet.

Ulykker: På lenke 133-10 har det skjedd 5 ulykker i fra 1. juli 2000 til ut 2005. Dette gir en ulykkesfrekvens på 0,25. Det er da forutsatt en gjennomsnittlig trafikkmengde på 800 ÅDT i denne perioden, og at observert ulykkesfrekvens (0,34) er korrigert mot en antatt normalfrekvens på 0,17.

Bompenger: Det er lagt inn 7 mill. kr i innkrevingskostnad for bompenge, dvs. 2,8 mill kr/år (beløpet er hentet fra bompengeselskapets nettside) og 2,5 års drift.

Nytte av nyskapt trafikk

Faktisk nyskapt trafikk er beregnet ovenfor. Følgende er lagt inn i nytteberegningen: ÅDT 165 første år etter åpning (dvs. ÅDT i 2001) og deretter ÅDT 326 i løpet av første to år uten bompenge (ÅDT 2004).

Nytten er først beregnet for trafikkveksten for første år med tunnel (nyskapt trafikk 1). Etterspørselskurvene som ble laget for bevilgningssituasjonen er brukt for å finne nytten for nyskapt trafikk 1 (ÅDT 165 i 2000). Det er ellers benyttet samme forutsetninger for trafikkfordeling og antall personer pr. kjøretøy som i beregningen fra 1997.

Ved bruk av etterspørselskurvene regnes først ÅDT om til personturer. Lette kjøretøyer er forutsatt å ha 1,8 person pr. kjøretøy, tunge biler er forutsatt å ha 1,3 personer pr. kjøretøy. Man går deretter inn i etterspørselskurven ved T0, dvs. persontrafikk siste år på ferja, og ved T0+nyskapt trafikk (=Te). På denne måten

”plukkes” årlig nytte fra kurvene for hhv. tunge og lette for hver av de to reiserelasjonene Hitra-Frøya og Trondheim–Frøya.

Nytten for *nyskapt trafikk 1 blir da: 3,2 mill. (2002-kr).*

For nyskapt trafikk som følge av bompengene (nyskapt trafikk 2, ÅDT 326) er nytten forutsatt å komme i 2004. Nyttens av nyskapt trafikk 2 kommer som følge av at bompengene forsvinner. Siste hele år med bompenger var i 2002, bruttoinntektene fra bompengene var da 16,1 mill. kr. ÅDT i 2002 var 545. Dette gir at hvert kjøretøy har betalt følgende bomavgift:

Gjennomsnittlig bomavgift pr. kjøretøy: $16,1 \text{ mill. kr.} / (365 * 545 \text{ kjt}) = 81 \text{ kr/kjt.}$
Med tungtrafikkandel 15,6 % har vi gått ut i fra at de tunge kjøretøyene betalte 240 kr. pr. passering, mens de lette betalte 52 kr pr. passering.

Med disse bomtakstene kan nytte av nyskapt trafikk 2 når nyskapt trafikk er ÅDT 326 (dvs. ÅDT-lette er 315 og ÅDT-tunge er 11):

ÅDT-nyskapt 2003-2004 = $(315 \text{ kjt/døgn} * 365 \text{ døgn/år} * 52 \text{ kr/kjt.} + 11 \text{ kjt/døgn} * 365 \text{ døgn/år} * 240 \text{ kr/kjt}) / 2. = 3,45 \text{ mill. kroner (i 2002-kr).}$

Beregnet nytte er lagt inn som nytte for ett år for lette og tunge kjøretøy, dvs. totalt 3,2+3,45 mill. kr = 6,65 mill. kr. Denne nytten er over det dobbelte av det som var lagt inn ved bevilgning.

EFFEKT regner automatisk at nytten av nyskapt trafikk kommer i første år av beregningsperioden, mens virkelig situasjon er at nyskapt trafikk kommer i to ”doser” ett år og tre år etter åpning. Det hadde vært mulig å ta hensyn til dette ved å lage en utbyggingsplan med to utbygde vegnett (i tillegg til vegnett 0) i EFFEKT5 hvor første utbygde vegnettet gjelder perioden 2000 t.o.m. 2002, og at det andre vegnettet åpner i 2003. Men feilen denne tidsforskyvingen gir i diskontert nytte av nyskapt trafikk antas å bety lite for beregningsresultatet.

Nytte av nyskapt trafikk utenfor prosjektområdet

Ved bevilgning var det forutsatt at trafikken mellom grensepunkt 1 og 4 (Hammarvika på Frøya til Hitra) gikk til/fra Trondheim, 148 km unna. Denne trafikken får dermed en lang distanse med kjøring utenfor prosjektområdet, noe som innebærer økte ulykkes- og vedlikeholdskostnader.

I ettersituasjonen vet vi at nyskapt trafikk er større enn forutsatt ved bevilgning. Det er imidlertid ukjent hvor den nyskapte trafikken kommer fra. Men det virker urimelig at all nyskapt trafikk skal til/fra Trondheim. I ettersituasjonen har vi forutsatt at det kun er de to trafikkstrømmene mellom Hitra (grensepkt. 4) og Frøya (grensepkt. 1 og 2) som ”belaster” vegnettet utenfor beregningsområdet. Vi har videre antatt at disse to trafikkstrømmene utgjør 78 % av totaltrafikken, dvs. samme fordeling som forutsatt for den øvrige, ikke-nyskapte trafikken. Nyskapt trafikk for strømmene 1-4 og 2-4 er lagt inn i EFFEKT og er som vist i tabellen nedenfor:

	Lette	Tunge	Busser	Sum ÅDT
Nyskapt ÅDT i 2001	135	27	4	166
Andel i 2-4 og 1-4:	106	21	3	130
Nyskapt ÅDT i 2004	449	35	6	490
Andel i 2-4 og 1-4:	351	27	5	383

Tabell 16 Nyskapt trafikk for strømmene 1-4 og 2-4

Kjørelengde for nyskapt trafikk utenfor beregningsområdet er satt til 50 km. Dette er gjort ut i fra en antagelse om at de to nærmeste tettstedene på Hitra er Fillan og Sandstad som er hhv. ca 10 og 25 km unna prosjektområdet (Storvatnet), mens de to største ”fjernmålene” er antatt å være Orkanger (ca 100 km) og Trondheim som er ca 145 km unna. Hvis vi antar at 70 % er lokaltrafikk med kjørelengde 20 km (utenfor prosjektområdet), og 30 % har kjørelengde 120 km, gir dette en midlere kjørelengde på 50 km. Dette er lagt inn i EFFEKT, sammen med en antatt ulykkesfrekvens på 0,19 (samme som ved bevilgning).

6.3.6 Resultater og konklusjon

De viktigste dataene som har betydning for nyttekostnadsanalysen er vist i tabellen nedenfor.

<i>Etterprøving av prissatte konsekvenser for:</i>		<i>Rv 714 Frøyatunnelen m/ tilstøtende vegnett</i>		
Nøkkeldata for de prissatte konsekvensene:		Ved bevilgning (1997)	Ettersituasjon (2005)	Merknad
1.	Anleggskostnad (Prisnivå 1999)	589,0 mkr	459,6 mkr	Avvik pga at Dolmsundet bru + tilstøtende veg ikke er bygd, bedre geologi enn først antatt, kortere anleggstid en opprinnelig forutsatt.
2.	Anleggsperiode	4 år (2001)	2 år 8 mnd (2000)	Samme som ovenfor.
3.	Trafikkmengde v/ åpning av ny veg	353	512	
4.	Trafikk på avlastet (gammel) veg	-	-	Avlastet veg er nedlagt ferje
5.	Tungtrafikkandel (på ny veg)	10%	12%	
6.	Årlig trafikkvekst i beregningsperioden, 25 år:	1,2% (f.o.m. 2001)	18,3%	For ettersituasjon: 18,3% for perioden 2001 til og med 2005. 5% vekst i 2006. Fylkesprognose 1,2% deretter benyttet for ettersituasjon, 2006-2025.
7.	Kjørehastighet (beregnet og obs.)	74 km/t	68km/t	Hastighet i ettersituasjonen blir lavere fordi man benytter "gammel" rv 714 på en strekning fordi Dolmsundet bru ikke er bygd.
8.	Ulykkesfrekvens, ny E10	0,17	0,25	For ny veg: 5 ul. m/ personskade 2000-2005, som gir Uf på 0,25 e. korreksjon.
9.	Ulykkesfrekvens avlastet veg	-	-	Avlastet veg er nedlagt ferje
10.	Antall svært støyplagede personer	-	-	Ikke beregnet

Tabell 17 Nøkkeldata for nyttekostnadsanalysene

<i>Etterprøving av prissatte konsekvenser for:</i>		<i>Rv 714 Frøyatunnelen m/tilstøtende vegnett</i>	
	Ved bevilgning (1997)	Etter-situasjon(2005)	Merknad (avvik er kommentert i forhold til situasjon ved bevilgning)
Utbyggingskostnad (2001-kr)	589,0	459,6	Tunnel ble billigere enn forutsatt + Dolmsundet bru med tilførselsveger er ikke bygd
Nytte-kostnadsanalyse (2001-kr):	Kalkulasjonsrente 7% er benyttet i begge beregninger		
Tidskostnader	113,0	179,1	Mer trafikk (høyere trafikkvekst enn oppr. forutsatt), dette selv om trafikk i ettersituasjon har lenger kjøredistanse pga. at det ikke er bru over Dolmsundet.
Kjøretøyers driftskostnader	-3,0	-19,6	Jf. kommentar over om kjøredistanse
Nytte av nyskapt trafikk	40,5	223,3	Jf. kap. 2 vedr. nyskapt trafikk, denne er ca ÅDT 500, mot forutsatt ca ÅDT 65.
Ulempeskostnader for ferjetrafikanter	42,0	73,0	Mer trafikk enn oppr. forutsatt
Ulykkeskostnader (*)	-12,3	-41,5	Ved bevilgning: ulykkeskostnad utenfor prosjektområde pga nyskapt trafikk ser ut til å være for høy. I ettersituasjon: mer trafikk+ større kjøredistanse pga at bru over Dolmsundet mangler.
Miljøkostnader (kun luftforurensning)	9,5	8,6	Beregnete miljøkostnader er pga. global/regional luftforurensning, støy ikke vesentlig tema.
Innkrevingskostnader bompenger	0	-6,8	Ikke lagt inn i beregning ved bevilgning.
Restverdi	38,4	30,0	Avvik skyldes at faktisk utbyggingskostnad ble lavere enn oppr. forutsatt.
Sum nytte (N)	228,2	446,1	
- Sum investering og drift (eks. mva)	469,8	290,1	
= Netto nytte (NN)	-241,6	156,0	
Sum investering& drift over budsjett (inkl. mva) (K)	283,3	180,3	
Nytte-kostnadsbrøk (NN/K)	-0,85	0,87	

Tabell 18 Resultater fra nytteberegninger

Resultatene fra nyttekostnadsanalyse for situasjon på bevilgningstidspunktet og for ettersituasjon er vist i tabellen nedenfor. Resultatene bør presenteres med vekt på å synliggjøre effektene i henhold til det som var målsettingene med tiltaket. Eksempelvis kan fokus ligge på næringslivets transportbehov. Da kan i så fall posten ”tidskostnader” brytes ned på lette og tunge biler, og ”lette biler” kan ytterligere brytes ned på arbeidsrelaterte reiser og fritidsreiser. Samme nedbryting kan gjøres for posten ”nytte av nyskapt trafikk”.

Hovedmålet med vegprosjektet var å bedre framkommeligheten ved å erstatte ferjeforbindelsen med en vegforbindelse. Dette bl.a. for å få positive regionale effekter. Effektberegningen viser at framkommelighet er viktigste, positive komponent i nytteberegningen (sum tidskostnader og kjøretøyers driftskostnader), både for analysen ved bevilgningstidspunktet og i ettersituasjon.

Trafikkutviklingen siden prosjektets åpning i 2000 er betydelig sterkere enn det man forutsatte da prosjektet fikk første bevilgning i 1997. Beregnet nytte (N) i ettersituasjonen er ca dobbelt så stor som beslutningsgrunnlaget ved bevilgning skulle tilsi. Halvparten av prosjektets beregnede totale nytte kommer som følge av nyskapt trafikk

Det er registrert to ”sprang” i trafikken i løpet av tiden etter åpning. Det første kom i løpet av de første to årene etter prosjektets åpning, og det neste kom etter at bompengeneinnkrevningen ble flyttet til Hitratunnelen. Det er vanskelig å skille hva som er nyskapt trafikk, dvs. trafikk som kommer som en direkte følge av dette tunnelprosjektet, og hva som ville ha kommet uavhengig av Frøyatunnelen. Årsaken til at det er vanskelig å skille på dette er at de andre nye vegprosjektene i området også kan ha bidratt til økt reiseaktivitet og reiseomfang (Hitratunnelen og brua til Fjellværøy). I tillegg er det vanskelig å kvantifisere effekten bompengeneinnkrevningen har hatt på trafikkutviklingen. Dette fordi at det siden 1991 har vært bompengeneinnkreving i området, først i form av forhåndsinnkreving på ferjene. Deretter ble det bompengeneinnkreving i tilknytning til hvert av prosjektene etter at disse ble åpnet for trafikk. Fra 2003 ble bompengeneinnkreving endret til kun å omfatte innkreving fra en bomstasjon (ved Hitratunnelen). Når bompengeneinnkrevningen ved Hitratunnelen opphører forventes ytterligere trafikkvekst også i Frøyatunnelen

Prosjektet pr. 2006 er ikke er bygd slik det var forutsatt ved første bevilgning. Bru over Dolmsundet og tilhørende tilførselsveg er ikke bygd, noe som gjør at dagens trafikk får ekstra kjørelengde i forhold til forutsatt i 1997. Sammen med at de geologiske forholdene for tunnelbygging var bedre enn antatt i opprinnelige

kostnadsoverslag, har ufullført Dolmsund bru medført at prosjektet pr. desember 2001 var ca 130 mill. kr billigere enn forutsatt ved bevilgning. Pga. lavere utbyggingskostnad blir derfor beregnet netto nytte (NN) betydelig høyere enn beregnet i 1997, og samfunnsøkonomisk lønnsomhet beregnet i ettersituasjonen blir også positiv.

Sammenliknet med det som var beregnet/forutsatt før første bevilgning, ser ettersituasjonen ut til å få høyere ulykkeskostnad (altså en forverring i forhold til beregnet ved bevilgning) og høyere driftskostnad for kjøretøy (også en forverring). Dette skyldes at det generelt er mer vegtrafikk enn beregnet ved bevilgning, dette gjelder også nyskapt trafikk som belaster vegnettet utenfor prosjektområdet.

Totalt sett viser etterprøvingen at man har nådd hovedmålene om bedre framkommelighet (dvs. erstatte ferje med veg) og positiv utvikling/vekst i regionen. Nyttekostnadsberegningene viser at prosjektet har vært samfunnsøkonomisk lønnsomt. Lønnsomheten er et resultat av den gunstig trafikkutviklingen, dvs. den nyskapte trafikken. Prosjektet ville fått negativ nytte, og $NN/K = -0,37$ hvis man fjernet nytten som følge av nyskapt trafikk.

Det er imidlertid vanskelig å etterprøve i hvilken grad trafikkutviklingen (den nyskapte trafikken) har kommet som følge av Frøyatunnelen alene, eller om naboprojektene Hitratunnelen og bru til Fjellvørsøya (eller andre forhold uavhengig av vegprosjektene) også har hatt betydning. Mest sannsynlig har de to naboprojektene og den positive utviklingen i havbruk og turisme hatt vesentlig betydning for trafikkutviklingen i Frøyatunnelen.

I nyttekostnadsberegningen for ettersituasjonen er det derfor sannsynlig at man har regnet inn effekter (dvs. nytte av nyskapt trafikk) som like mye skyldes de to andre prosjektene. Ideelt sett burde man hatt en samfunnsøkonomisk beregning som inkluderte alle de tre prosjektene på Hitra/Fjellvørsøya/Frøya. Til tross for de nevnte usikkerheter rundt nyttekostnadsanalysene er konklusjonen at Frøyatunnelen har bidratt positiv til den regionale utviklingen i denne delen av Sør-Trøndelag, samtidig som den har medført en vesentlig forbedring av tilgjengelighet og framkommelighet for innbyggere og næringsliv på Frøya. Totalt sett er det derfor grunn til å hevde at Frøyatunnelen ble et mer samfunnsøkonomisk lønnsomt prosjekt enn det som ble forelagt beslutningstagerne ved bevilgning i 1997.

Om regionale virkninger

Nytten av nyskapt trafikk er en klar indikasjon på gunstige regionaløkonomiske effekter, fordi den gjenspeiler en økning i aktivitetsnivået som følge av prosjektet. Analysen har imidlertid ikke konkludert med hvorvidt deler av denne nyskapte trafikken er overført trafikk fra andre regioner, som følgelig kan ha kommet dårligere ut. Det er imidlertid ikke lett å se at slike effekter kan ha vært av vesentlig betydning i dette tilfellet, der et øysamfunn blir knyttet tettere mot fastlandet og der det er lite innslag av konkurrerende vegstrekninger eller transportmidler. I andre typer prosjekter kan imidlertid vegvalgseffekter (og i noen tilfeller også transportmiddelvalg) gjøre at det er grunn til å være oppmerksom på at tilsynelatende vesentlige volumeffekter kan ha klare overføringseffekter i seg, og at de samfunnsøkonomiske virkningene følgelig kan være mindre enn det en analyse av trafikkutviklingen isolert sett kan tyde på. En av de vesentlige fordelene med bruk av transportmodeller er at man kan få rendyrket nettoeffektene av tiltaket.

Som nevnt innledningsvis er det grunn til å vurdere en ex post-analyse av regionale virkninger dersom dette er et uttalt mål med prosjektet.

Miljø, og vanskelig kvantifiserbare virkninger (VKF)

Dette eksempelet inneholder ingen omtale av vanskelig kvantifiserbare sosiale virkninger. Endret reisetid til knutepunkt som f.eks. nærmeste by eller terminal kan angis. Det samme kan gjelde reduksjon i antall sosiale møteplasser (bortfall av ferjeforbindelse), eventuelle handelsvirkninger (sannsynlig langsiktig virkning ut over 2 år) ved at nærbutikk eventuelt forsvinner i konkurranse med Hitra eller fastlandet. Nærbutikk er også en sosial møteplass. Globale miljøvirkninger (utslipp til luft) er med, og beregningsopplegget for transportetatene har også metoder for verdsetting av støy og lokal luftforurensning som kan presenteres dersom slike effekter finnes.

Når det gjelder andre VKF som barrierevirkninger og landskapsmessige forhold, så kan dette ex post-evalueres ved hjelp av etatenes metodikk for slike analyser.

Konklusjoner

Prosjektgruppen har i det foregående drøftet forslag til metode for ex post-evalueringer av store offentlige prosjekter. Her oppsummeres gruppens forslag.

Hensikten med ex post-evalueringer er for det første å supplere de to etablerte punktene for kvalitetssikring, KS1 og KS2. Ex post-evalueringene skal vurdere om de opprinnelige målsettingene for prosjektene faktisk blir oppnådd. I tillegg vil ex post-evalueringene kunne gi viktige bidrag til å forebedre ex ante-analysene, til å øke bevisstheten om at ex ante-beregningene som investeringen bygger på må være realistiske eller ”edruelige”, og til å standardisere analyser og beslutningsunderlag ex ante. Gjennom ex post-evalueringer vil en også kunne bygge opp en erfaringsdatabase om effekter av store investeringer, som vil kunne være nyttig når nye prosjekter skal utformes og planlegges.

I vurderingene har en rekke hensyn blitt lagt til grunn:

- Evalueringene bør bygge på anerkjent kunnskap om og erfaringer fra evalueringer. Det er derfor i og for seg ikke snakk om ny metodeutvikling, men tilpasning av eksisterende metodeapparat til det beslutnings- og kvalitetssikringssystemet som er etablert, og videreutvikling av dette.
- Evalueringene må være gjennomføre innenfor realistiske ressursrammer.
- Evalueringen må i størst mulig grad kunne gjennomføres på grunnlag av data som enten er allment tilgjengelig eller som uansett genereres i forberedelser og gjennomføring av prosjektet.
- Evalueringen må kunne gjennomføres innen rimelig tid etter at prosjektet er gjennomført.

Vi forslår at alle ex post-evalueringer baseres på følgende hovedprinsipper:

1. Ex post-evalueringen skal ha som formål å kartlegge effekter av det gjennomførte prosjektet sammenlignet med en situasjon der prosjektet ikke ble gjennomført (null-alternativet), og sammenlignet med effekter av alternative men ikke valgte konsepter som ble utviklet og vurdert i KS1-evalueringen.

Videre bør evalueringene inneholde:

2. En vurdering av gjennomføringen av prosjektet med hensyn til effektivitet, kvalitet og resultater. Gjennomføringen bør vurderes mot et utvalg relevante referanseprosjekter.
3. En vurdering av i hvilken grad nytten og kostnadene knyttet til det realiserede prosjektet faller sammen med de ex ante-beregninger som lå til grunn for beslutningen om gjennomføringen av prosjektet, samt forklaring av eventuelle avvik
4. En analyse av prosjektets samfunnsmessige effekter i et bredt perspektiv
5. En vurdering av om riktig konsept ble valgt, herunder om et annet konsept burde ha blitt valgt i lys av den nye kunnskap man har ervervet, både om det gjennomførte prosjektet og om den utviklingen som har skjedd i de samfunnsmessige og teknologiske rammene rundt prosjektet.
6. En samlet oppsummering av erfaringer fra prosjektet, herunder en formidlings-/læringsstrategi.

Det bør bygges opp en database/et bibliotek der erfaringene fra evalueringene samles. På flest mulig av evalueringspunktene bør data standardiseres og legges til rette for framtidig benchmarking, forskning og analyse. Data bør forvaltes slik at de kan være tilgjengelige for aktuelle framtidige gjennomførere av KS1, KS2 og ex post-evalueringer, for eksempel i regi av Senter for Statlig Økonomistyring (SSØ) eller forskningsprosjektet Concept.

For at ex post-evalueringene skal bli gode forutsettes det et godt datagrunnlag, herunder nullpunktsmåliger, data om nullalternativet og alternative konsepter. Dette stiller krav til beslutningsunderlaget og tilhørende analyser. Det vil være nødvendig at det i KS1 og KS2 sikres et godt datagrunnlag for framtidig ex post-evaluering.

Det aktuelle fagdepartementet, i samarbeid med Finansdepartementet, bør ha hovedansvaret for at ex post-evalueringen gjennomføres. Det bør enten tas høyde i prosjektets budsjett for kostnadene ved evalueringen eller det må opprettes en budsjettpost for ex post-evalueringer for de tilfeller der evalueringene vil komme etter at prosjektet er fullført og ”lukket”. Tidspunktet for ex post-evalueringen vil nok måtte variere fra prosjekt til prosjekt, men dersom det lar seg gjøre anbefales to år etter at prosjektet er fullført som et standardisert tidspunkt.

7 Referanser

- Ackerman, F. og Heinzerling, L. (2004) **Priceless: On Knowing the Price of Everything and the Value of Nothing**, New Press, New York
- Andersen, Bjørn og Fagerhaug, Tom (2001) **Performance Measurement Explained: Designing and Implementing your State-of-the-Art System**, ASQ Quality Press, Milwaukee
- Andersen Bjørn og Pettersen, Per-Gaute (1995) **Benchmarking - en praktisk håndbok**, Tano Forlag, Oslo
- Ansoff, I.H. (1975) **Managing Strategic Surprise by Response to Weak Signals**, California Management Review, Vol. XVII, No. 2, Winter
- APM-konsortiet (2002) **APM (Automated Performance Measurement) Final Report**, Brussel, Belgia
- Arge, N.; Homleid, T og Stølan, A. (2000): **Modeller på randen... Bruk av transportmodeller i norske byområder. En evaluering**, LOKTRA-prosjekt, Oslo
- Aspengren, N.C. (1999) **Fra Roa til Bergen Historien om Bergensbanen**, Baneforlaget, Oslo
- Baird, M.E. og Stammer, R.R (2000) **Measuring the performance of state transportation agencies – three perspectives**, Transportation planning, public participation, and telecommuting
- Berg, P, Stene, E.A., Fon, A.M., Kvarsvik, T., Bye, B., Aksnes, L., Hallquist, Ø., og Pedersen, P. (1998) **Gjennomgang av investeringer i samferdselssektorene**, Rapport fra en arbeidsgruppe med deltakelse fra Finansdepartementet, Samferdselsdepartementet, Jernbaneverket, Vegdirektoratet og Luftfartsverket, Prosjektet for styring av statlige investeringer, Finansdepartementet, Oslo
- Bredrup, Harald (1995) **Performance Measurement in a Changing Competitive Industrial Environment: Breaking the Financial Paradigm**, Dr.ing.-avhandling, Norges Tekniske Høgskole, Trondheim
- Bruer, A. (1989) **Modernisering av vestfoldbanen**, Planutredning, Prosjekt 89125-2
- Byggeriets Evaluerings Center (2005) **Virksomhedens Karakterbog**, <http://www.byggeevaluering.dk/object.php?obj=fa0029&GSID=842594>
- Camp, Robert C. (1989) **Benchmarking: The Search for Industry Best Practices that Lead to Superior Performance**, ASQC Quality Press, Milwaukee
- Cowi (2004) **Uttesting av metodikk for etterprøving av prissatte konsekvenser av store veganlegg**, Utkast til hovedrapport
- Davis, J.G. (1975) **Keeping Project Costs in Line, Machine Design**, Vol. 47, No. 29, pp. 128-133
- De Lancer Julnes, P. (2006) **Performance Measurement**, Evaluation, Vol 12(2) pp 219-235

- Deming, W.E. (1982) **Out of the Crisis: Quality, Productivity and Competitive Position**, Cambridge University Press, Cambridge
- ENAPS-konsortiet (2002) **ENAPS (European Network of Advanced Performance Studies) Final Report**, Brüssel, Belgia
- Finansdepartementet (2005a) **Veileder til gjennomføring av evalueringer**, Finansdepartementet, Oslo
- Finansdepartementet (2005b) **Veileder i samfunnsøkonomiske analyser**, Finansdepartementet, Oslo
- Flyvbjerg, Bent, Holm, Mette Skamris og Buhl, Søren (2002) **Underestimating Costs in Public Works Projects: Error or Lie?** PA Journal, Vol. 68, No. 3.
- Flyvbjerg, Bent, Holm, Mette Skamris og Buhl, Søren (2005). **How (In)accurate Are Demand Forecasts in Public Works Projects? The Case of Transportation**, Journal of the American Planning Association, 71(2), pp. 131-146
- Flyvbjerg, B., Bruzelius, N. og Rothengatter, W. (2003) **Megaprojects and Risk. An Anatomy of Ambition**, Cambridge University Press, Cambridge
- Fordal, T V (2007) **Etterprøving av prissatte konsekvenser for Frøyatunnelen**, COWI/Statens Vegvesen
- Gosling (2000) **Aviation system performance measures for state transportation planning. Issues, problems, and performance measures in airports and airspace**, Transportation research record (1703), pp. 7-15
- Gottschalk, P. og Wenstøp, F. (1985) **Kvantitativ beslutningsanalyse for ledere og planleggere**, Universitetsforlaget, Oslo
- Grove-White, R. (1997) **The Environmental Valuation Controversy**, i Foster, J. (red.) **Valuing Nature?**, Economics, Ethics and Environment, pp. 21-31, Routledge, London
- Grünfeldt, L.A. og Jakobsen, E.W (2006) **Hvem eier Norge? Eierskap og verdiskapning i et grenseløst næringsliv**, Universitetsforlaget, Oslo
- HM Treasury (2004) **The Green Book. Appraisal and Evaluation in Central Government**, TSO, London
- Hronec, Steven M. (1993) **Vital Signs: Using Quality, Time, and Cost Performance Measurements to Chart your Company's Future**, Amacom, New York
- Independent Project Analysis, IPA (2006) **Project Evaluation System**, http://www.ipaglobal.com/inside%20pages/Prod_Serv/PES/pes.html
- JBV Utredning (2004) **Utviklingen i persontrafikken på jernbanen. En analyse av årsakene til nedgangen i trafikken i perioden 2000-2003**, JBV, Oslo
- Johansen, Leif (1977) **Samfunnsøkonomisk lønnsomhet**, Industriøkonomisk institutt
- Kaplan, R.S. og Norton, D.P. (1992) **The Balanced Scorecard – Measures that Drive Performance**, Harvard Business Review, Jan./Feb., pp. 71-79
- Kaplan, R.S. og Norton, D.P. (1996) **The Balanced Scorecard: Translating Strategy into Action**, Harvard Business School Press, Boston

- Kenworthy, J. (1990) **Don't Shoot Me – I'm Only the Transport Planner**, i Newman, P., Kenworthy, J. og Lyons, T. **Transport Energy Conservation**, Murdoch University, Perth
- Kostnadsberegningutvalget (1997) **Nytte-kostnadsanalyser: Prinsipper for lønnsomhetsvurderinger i offentlig sector**, NOU 1997:27
- Lauridsen, H. (2001) **The Scandinavian Transport Planning Experience**, Nordic Road and Transport Research, No. 2, 2001
- Lohmann, L. (1997) **Cost-Benefit Analysis: Whose Interest, Whose Rationality?**, The Corner House, London
- Miller, R. og Lessard, D.R. (2001) **The Strategic Management of Large Engineering Projects: Shaping Institutions, Risks, and Governance**, MIT Press, Boston
- Nielsen, O. A. og Fosgerau, M. (2005) **Overvurderes tidsbenefit af vejprojekter?**, Paper til Trafikdage på Aalborg Universitet
- Norsk Jernbaneklubb (1987-2005) **På Sporet**, Norsk Jernbaneklubb, Oslo
- NOU (1999:28) **Gardermoprojektet, Evaluering av planlegging og gjennomføring**, Statens trykking, Oslo
- NSB (1986) **Dobbelsporet Ski-Moss**, Utredningsoppdrag nr. 112, Oslo
- NSB (1992a) **Modernisert Vestfoldbane Jernbaneutredning Hovedrapport**, NSB Konsernstab strategi og miljø, Oslo
- NSB (1992b) **Linjeomlegging Tunga-Finse**, NSB Banedivisjonen Region Vest
- Næss, P. (2006) **Cost-benefit analyses of transportation investments: neither critical nor realistic**, Journal of Critical Realism, Vol. 5, No. 1, 2006, pp. 32-60
- Odgaard, T., Kelly, C. og Laird, J. (2005) **Current Practice in Project Appraisals in Europe**, Paper til Trafikdage på Aalborg Universitet
- OECD (2000) **OECD glossary of evaluation and results based management terms**, Organisation for Economic Co-operation and Development, Paris
- Olsson, N.O.E. (2004) **Effektivrdering**, SINTEF rapport/Concept-rapport 1050-3, Trondheim
- Olsson N.O.E. og Fagerhaug, T. (2005) **Hvordan ble det? Effektoppfølgning av store offentlige prosjekter**, NTNU/SINTEF, Trondheim
- Olsson, N.O.E. og Haugland, H. (2004) **Influencing factors on train punctuality - results from some Norwegian studies**, Transport Policy, 11(4), pp. 387-397
- Owens, S. og Covell, R. (2002) **Land and Limits: Interpreting Sustainability in the Planning Process**, Routledge, London
- Owens, S., Rayner, T. og Bina, O. (2004) **New agendas for appraisal: Reflections on theory, practice and research**, Environment and Planning A 36, 11, 1943-1959
- Oxera (2005) **How should the ex post evaluation of trunk road schemes be enhanced?**, Rapport til Department for Transport, <http://www.dft.gov.uk/pgr/roads/network/research/hsebe/howshouldtheexpostevaluation3912>
- Rikshospitalet (2004) **Erfaringer fra Gaustad**, tilgjengelig på www.sykehusplan.no

- SACTRA (1999) **Transport and the Economy**, Standing Advisory Committee on Trunk Road Appraisal, London
- Samset, K. (2003) **Project Evaluation: Making investments succeed**, Tapir Akademisk Forlag, Trondheim
- Samset, Knut, Berg, Peder og Klakegg, Ole Jonny (2006) **Front End Governance of Major Public Projects**, EURAM06 Conference, Oslo
- SIKA (2003) **Uppföljning av de transportpolitiska målen**, SIKA-rapport 2003:5, <http://www.sika-institute.se/>
- SIKA (2004a) **Uppföljning av de transportpolitiska målet och dess delmål**, SIKA-rapport 2004:3, <http://www.sika-institute.se/>
- SIKA (2004b) **Mått för måluppföljning: Tänkbara mått för de transportpolitiska delmålen om tilgjänglighet, regional utveckling och transportkvalitet**, SIKA-rapport 2004:5, <http://www.sika-institute.se/>
- Sink, D.S. og Tuttle, T.C. (1989) **Planning and Measurement in your Organization of the Future**, Industrial Engineering and Management Press
- Sound Transit Expert Review Panel (2005) **Letter from the Expert Review Panel to the Board of Sound Transit**, datert 2. Desember
- Spendolini, Michael J. (1992) **The Benchmarking Book**, Amacom, New York
- Stallworthy, E.A. (1979) **Developments in Project Cost Control**, Engineering and Process Economics, Vol. 4, No. 1, pp. 29-36
- Stame, N. (2006) **Governance, Democracy and Evaluation**, Evaluation, Vol 12(1) pp 7-16
- Statens Vegvesen, Vegdirektoratet, Utbyggingsavdelingen (2006) **Rapport UTB 2006/2**, Oslo
- Strand, A. (2006) **Flermålsanalyser i store statlige investeringsprosjekt. Kapittel med presentasjon og analyse av eksempler**, foreløpig upublisert notat, Transportøkonomisk institutt, Oslo
- Tennessee Rail Systems (2001) **Strategic Project Evaluation Protocols and Procedures**, Tennessee Department of Transportation, Chattanooga
- Tennøy, A. (2003) **Bidrar bruk av transportanalyser i byplanleggingen til vekst i biltrafikken?**, Paper til Trafikdage på Aalborg universitet
- Thanassoulis, E. (2000) **DEA and its Use in the Regulation of Water Companies**, European Journal of Operational Research, No. 127, pp. 1-13
- United Nations (2003) **Protocol on Strategic Environmental Assessment to the Convention on Environmental Impact Assessment in a Transboundary Context**, Kiev
- Vegdirektoratet (2003) **Etterprøving av store veganlegg. Forslag til metodikk og registreringsskjema**, Oslo
- Washington State Department of Transportation (2006) **State of the Practice**, Inventory, March 2004, Learning from others

Whist, E., Bergsland, K. H., Jordanger, I., Ore, K. M. og Øvrebø, T. (2001) **Evaluering av NHR-prosjektet**, Arbeids- og administrasjonsdepartementet, tilgjengelig på www.sykehusplan.no

Wikipedia (2007) www.wikipedia.org