

Gro Holst Volden og Knut Samset

## Etterevaluering av statlige investeringsprosjekter

Konklusjoner, erfaringer og råd basert på pilotevaluering av fire prosjekter

Concept rapport Nr 30





Gro Holst Volden og Knut Samset

# Etterevaluering av statlige investeringsprosjekter

Konklusjoner, erfaringer og råd basert på pilotevaluering av fire prosjekter

Concept rapport Nr 30

© Concept-programmet 2013

Concept rapport nr. 30

## **Etterevaluering av statlige investeringsprosjekter**

### **Konklusjoner, erfaringer og råd basert på pilotevaluering av fire prosjekter**

*Forfattere Gro Holst Volden og Knut Samset*

*ISSN: 0803-9763 (papirversjon)*

*ISSN: 0804-5585 (nettversjon)*

*ISBN: 978-82-93253-09-9 (papirversjon)*

*ISBN: 978-82-93253-10-5 (nettversjon)*

*Sammendrag: Rapporten sammenfatter konklusjonene og erfaringene fra fire pilotevalueringer og gir på dette grunnlaget råd om metodikk og prosedyre for evaluering av store offentlige prosjekter som kommer inn under Finansdepartementets ordning for kvalitetssikring av slike prosjekter.*

*Dato: 2. januar 2013*

*Utgiver: Concept-programmet  
Institutt for bygg, anlegg og transport  
Norges teknisk- naturvitenskapelige universitet  
Høgskoleringen 7A  
7491 NTNU – Trondheim*

[www.concept.ntnu.no](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 syn.

Concept rapport nr. 30

## Forord

Concept-programmet har i 10 år drevet systematisk innsamling av data om store statlige investeringer som er underlagt ordningen med ekstern kvalitetssikring av beslutningsunderlaget (KS-ordningen). For å lære av det som gjøres er det nødvendig å følge prosjektene over tid, fra tidligfasen via gjennomføring og videre ut i driftsfasen. Helt sentralt er evalueringen i etterkant for å fastslå i hvilken grad en har oppnådd de tiltenkte effekter for brukere og samfunn. På grunn av den lange tiden som går fra tidligfasen og frem til et prosjekt er i driftsfasen, normalt 5-10 år, har det først nå vært mulig å evaluere effekter og måloppnåelse i bred forstand av de første prosjektene som gjennomgikk ekstern kvalitetssikring.

Concept tok høsten 2011 initiativ til å få gjennomført fire evalueringer. Disse er å betrakte som ”pilotevalueringer”, der hensikten har vært å bygge kompetanse og erfaring med evaluering og teste ut et generisk opplegg for evaluering av statlige investeringsprosjekter under KS-ordningen i tiden som kommer. Denne rapporten oppsummerer funnene fra de fire evalueringene, og presenterer erfaringer og anbefalinger basert på disse.

Arbeidet har vært ledet av forskningssjef Gro Holst Volden ved Concept-programmet. Evalueringene er gjennomført av tre leverandører på oppdrag fra Concept. Disse er VTI (Statens väg- og transportforskningsinstitut) ved Jan Eric Nilsson, Johan Nyström og Roger Pyddoke, SINTEF Teknologi og Samfunn ved Håkon Finne og Heidi Bull-Berg, og rådgivningsfirmaet Scanteam ved Erik Whist. Én evaluering ble i sin helhet gjennomført av Concept internt, ved Torbjørn Aass og Morten Welde. For å sikre læringseffekten har minst en forsker fra Concept også vært involvert i hvert av de eksterne evalueringsteamene (Gro Holst Volden, Knut Samset, Bjørn Andersen, Nils Olsson og Torbjørn Aass). Berørte departementer og etater har vært involvert ved at det tidlig ble oppnevnt kontaktpersoner, som har bistått med informasjon, har fulgt prosessen og bidratt med kommentarer til utkast. Også Direktoratet for økonomistyring har kommet med nyttige innspill til denne hovedrapporten. De fire evalueringene er dokumentert i separate rapporter.

Prosjektet har hatt en redaktørgruppe bestående av Ola Lædre, Nils Olsson Gro Holst Volden og Knut Samset, som har hatt ansvar for metaevalueringen og de samlede anbefalingene. De to sistnevnte har forfattet denne rapporten.

Trondheim, januar 2013

Knut Samset, Programansvarlig,  
Concept-programmet, NTNU

# Innhold

<b>SAMMENDRAG</b> .....	4
<b>SUMMARY</b> .....	12
<b>1. INNLEDNING OG FORMÅL MED STUDIEN</b> .....	16
1.1.    KVALITETSSIKRING, LÆRING OG LÆRINGSSLØYFER .....	16
1.2.    EX POST EVALUERING SOM LÆRINGSARENA.....	18
1.3.    FIRE PILOTEVALUERINGER OG FORSLAG TIL ET ENKELT OPPLEGG FOR EVALUERING .....	20
<b>2. METODIKK FOR PROSJEKTEVALUERING</b> .....	24
2.1.    DEFINISJON AV "VELLYKKET" PROSJEKT .....	25
2.2.    VALG AV OVERORDNET EVALUERINGSMODELL .....	30
2.3.    DATAINNSAMLING OG ANALYSE.....	39
<b>3. EVALUERING AV FIRE PROSJEKTER. SAMMENFATNING</b> .....	47
3.1.    SVINESUND KONTROLLOMRÅDE .....	47
3.2.    NYTT DOBBELTSPOR ASKER-SANDVIKA .....	56
3.3.    VEIPARSELL MOMARKEN-SEKKELSTEN .....	64
3.4.    BYGGING AV SKJOLD-KLASSE MTB-ER TIL SJØFORSVARET.....	71
<b>4. DRØFTING AV LÆRDOMMER FRA EVALUERINGENE</b> .....	79
4.1.    GENERELLE ERFARINGER FRA PILOTEVALUERINGENE .....	79
4.2.    TIDSPUNKT FOR EX POST EVALUERING.....	81
4.3.    AVGRENSNING – AVHENGIGHET MELLOM PROSJEKTER .....	82
4.4.    METODIKKEN: MÅLORIENTERT EVALUERING OG SAMFUNNSØKONOMISK ANALYSE .....	84
4.5.    MANGLER VED PROSJEKTLOGIKK OG MÅLSTRUKTUR .....	89
4.6.    INFORMASJONSGRUNNLAGET OG BEHOVET FOR TRIANGULERING .....	89
4.7.    TILGANG TIL EKSISTERENDE DATA.....	91
4.8.    KOMPETANSE OG TEAMSAMMENSETNING .....	93
4.9.    KS-ORDNINGENS BETYDNING.....	96
<b>5. RÅD OG ANBEFALINGER OM EVALUERING AV INVESTERINGSPROSJEKTER</b> .....	98
<b>LITTERATUR</b> .....	105

## Sammendrag

Det investeres hvert år milliardbeløp i store investeringsprosjekter over statlige budsjetter. Eksempelene spenner fra veier og jernbaneanlegg, til signalbygg i kultursektor, nytt forsvarsmateriell og store IKT-prosjekter. Å vurdere om disse prosjektene er vellykkede og samfunnsnyttige, kan bare gjøres etter at de er avsluttet og kommet i driftsfasen. Slike evalueringer gjennomføres imidlertid i liten grad i dag. Vi har behov for bedre kunnskap om hvor godt statlige investeringsmidler omsettes til samfunnsnytte i Norge. Det bør derfor i langt større grad enn i dag gjennomføres systematiske etterevalueringer av statlige investeringsprosjekter. Den primære hensikten må være læring – både internt i statlige virksomheter og på tvers i staten, for å medvirke til forbedringer i fremtidige investeringsprosjekter.

Concept-programmet driver følgeforskning på store investeringer som er underlagt krav om ekstern kvalitetssikring (KS-ordningen) og disse prosjektene er derfor et naturlig utgangspunkt for denne studien. For KS-prosjektene ligger det særlig godt til rette for systematisk læring ettersom resultatene av etterevalueringen (ex post) kan sammenholdes mot etatenes konseptvalgutredning og kvalitetssikrers vurderinger av prosjektene i forkant (ex ante). Vurderingene og anbefalingene i denne rapporten er likevel generelle og ment å være gyldig også for andre, mindre investeringsprosjekter.

### Metodikk for overordnet prosjektevaluering

Ex post evaluering av prosjekter handler om å vurdere i hvilken grad prosjektet har vært ”vellykket”. Vurderingen omfatter både et operasjonelt perspektiv (gjennomførte man det man skulle på en effektiv måte), et taktisk perspektiv (realiserte man de tiltenkte gevinster for målgruppene) og et strategisk perspektiv (var tiltaket relevant og nyttig i samfunnsmessig forstand). Tidligere erfaringer viser at det er det strategiske aspektet man er notorisk dårligst på å vurdere, fordi det er krevende og det utfordrer måten man tenker på. Ofte grenser det også mot det politiske. Ikke minst derfor er det viktig å ha et sterkt fokus på denne typen vurderinger.

Vi legger til grunn at en modell for evaluering av store statlige investeringsprosjekter bør være *målorientert*. Vi har testet ut en slik evalueringsmodell, ofte omtalt som OECD-modellen, som er internasjonalt anerkjent og anbefalt blant annet av FN og EU. Modellen spesifiserer fem overordnede evalueringskriterier; *produktivitet, måloppnåelse, virkninger, relevans og levedyktighet*. Videre supplerer vi denne modellen med en *samfunnsøkonomisk analyse*, som innebærer å sammenstille alle samfunnsøkonomiske nytte- og

kostnadsvirkninger av prosjektet, fortrinnsvis målt i penger. Samfunnsøkonomisk analyse skal normalt være en viktig del av beslutningsunderlaget for store investeringsprosjekter, og evalueringen kan da ta form av en etterprøving av analysen som ble gjort ex ante.

Selve evalueringsarbeidet innebærer en prosess med systematisk disaggregering og aggregering av informasjon. Med andre ord, man bryter ned de overordnede evalueringskriteriene i mer konkrete evaluerings spørsmål og indikatorer som er relevante i forhold til evalueringsobjektet. Deretter innhenter man data, analyserer disse og foretar en samlet vurdering for hvert evalueringskriterium og for prosjektet samlet. I rapporten diskuteres momenter som er viktige når en skal velge indikatorer og tilhørende datainnsamlings- og analysemetoder. Informasjonens kvalitet er et slikt moment som må vektlegges. Samtidig vet vi at mange evalueringer gjøres under tidspress og med begrenset budsjett, og en må derfor finne en balanse som sikrer at evalueringen både er nyttig og har tilstrekkelig legitimitet.

## Resultater fra fire pilotevalueringer

I denne studien har tre eksterne fagmiljøer vært engasjert for å evaluere hvert sitt prosjekt etter denne felles, overordnede evalueringsmodellen. For å sikre læringsutbyttet har Concept hatt ansvar for én evaluering i tillegg og hatt med minst en forsker i hvert av teamene. Prosjektene er valgt ut i samråd med Finansdepartementet. De har alle vært om lag fem år inne i driftsfasen, men er for øvrig helt ulike. Prosjektene er:

1. Svinesund kontrollområde
2. Nytt dobbeltspor Asker-Sandvika
3. Veiparsell E 18 Momarken-Sekkelsten
4. Bygging av Skjold-klasse MTBer til Sjøforsvaret

Ingen av prosjektene hadde gjennomgått kvalitetssikring av konseptvalget (KS1) men alle hadde gjennomgått kvalitetssikring av styringsunderlag og kostnadsestimat (KS2). Det innebar at en hadde et godt sammenlikningsgrunnlag for vurderingen av resultatene i et operasjonelt og til dels taktisk perspektiv.

Derimot var sammenlikningsgrunnlaget for de strategiske vurderingene og den samfunnsøkonomiske analysen til dels manglende eller i alle fall vanskeligere tilgjengelig. Ved senere evalueringer av prosjekter som har gjennomgått KS1 antar man at denne situasjonen vil være bedre.

Evalueringsteamene møtte en rekke utfordringer, både knyttet til manglende sammenlikningsgrunnlag, men også uklar avgrensning, problemer med tilgang på

data, metodiske problemer og uklarheter etc. En del oppgaver som vi i utgangspunktet antok ville være forholdsvis greie å gjennomføre, viste seg å være langt mer utfordrende og ga resultater som var mer variable enn forventet. Utfordringene drøftes i denne rapporten. Nedenfor gis først en kort sammenfatning av prosjektene og det de fire teamene er kommet frem til.

### ***Svinesund kontrollområde***

Prosjektet omfatter byggingen av et nytt felles kontrollområde for Tollvesenet, Statens vegvesen og Politiet, ved landets mest trafikkerte grenseovergang mot Sverige. Som en følge av sterk trafikkøkning over grensen gjennom mange år ble det planlagt ny vei og bru over Svinesund. Det lå i kortene at når vegens trasé ble flyttet, måtte kontrollstasjonene flytte etter. Da dette skjedde valgte en å samlokalisere kontrollstatene og således samtidig legge til rette for et tettere samarbeid mellom dem.

Evalueringsteamet mener at gjennomføringen av prosjektet fremstår som god og effektiv. Mål knyttet til kontrolleffektivitet og smidigere trafikkavvikling er i overveiende grad oppnådd, selv om potensialet for samarbeid mellom etatene trolig bør utnyttes videre. Det er også oppstått problemer med fremkommeligheten for tungtrafikken de siste årene som må adresseres. Det er ellers få negative sidevirkninger, og prosjektet anses som relevant og levedyktig gitt det tollpolitiske regimet som gjelder i dag. Den samfunnsøkonomiske lønnsomheten har det imidlertid ikke vært mulig å tallfeste. Prosjektet anses samlet sett som rimelig vellykket og samfunnsnyttig, både i operasjonelt, taktisk og strategisk perspektiv.

### ***Nytt dobbeltspor Asker-Sandvika***

Prosjektet omfatter utbygging av 9,5 km nye dobbeltspor for jernbane på strekningen Asker-Sandvika (mesteparten i tunnel) samt arbeid på begge stasjoner og nytt signalanlegg. Prosjektet var første byggetrinn i et større utbyggingsprosjekt for nye dobbeltspor helt fra Asker til Skøyen. Bakgrunnen var et økende kapasitetsproblem da en hadde bare ett dobbeltspor og betydelig trafikk med både lokaltog, regiontog, gods og etter hvert flytog på det samme sporet.

Evalueringsteamet finner at produktiviteten i prosjektet var (med ett unntak) tilfredsstillende. Både måloppnåelsen og den samfunnsøkonomiske lønnsomheten vurderes derimot å være begrenset, og var trolig overvurdert i forkant. En ser bl.a. at togtilbudet kun er marginalt bedret, og utbyggingen synes ikke å ha gitt noen stor vekst i passasjertallene. Tvert imot ser en at ruteplaner ikke er lagt om som forutsatt, og at det fremdeles er flaskehalsen lenger inn mot Oslo som begrenser gevinstuttaket. En mener at prosjektet i et strategisk perspektiv *kan* vise seg å være



fornuftig, men dette kan ligge langt frem i tid og ifølge evaluatør synes det ikke å ha vært godt nok utredet.

### ***Veiparsell Momarken-Sekkelsten***

Prosjektet omfatter en 6,2 kilometer lang veiparsell på E18 mellom Mysen og Askim i Østfold. Veien er del av en større utbygging av ny E18 gjennom hele Østfold. Bakgrunnen for dette var at den gamle veien gikk innom alle tettstedene og at strekningen var trafikkert med bl.a. mye godstrafikk, noe som ga lav fremkommelighet, høy ulykkesrisiko og stor miljøbelastning.

I et operasjonelt perspektiv, og vurdert kun for denne parsellen isolert, finner evalueringsteamet at resultatet er relativt godt til tross for betydelig kostnadsoverskridelse. En mener at denne type parsellvise utbygging fremstår som en lite kostnadseffektiv måte å bygge veg på. Trafikk- og ulykkesdata tyder på god måloppnåelse og en mener å kunne godtgjøre høy samfunnsøkonomisk lønnsomhet. Blant negative sidevirkninger er redusert kollektivandel da bilen som fremkomstmiddel er blitt mer attraktiv og toget har mistet passasjerer. Evaluatør mener at relevans og levedyktighet trolig er positiv, dog er dette usikkert i lys av at alternative tiltak aldri ble utredet i forkant. Teamet konkluderer med at prosjektet alt i alt fremstår som rimelig vellykket og samfunnsnyttig med den store trafikkveksten som har skjedd og som trolig vil fortsette.

### ***Bygging av Skjold-klasse MTB-er til Sjøforsvaret***

Prosjektet omfatter bygging av fem missiltorpedobåter i Skjoldklassen. Dette er fartøy som skal utføre overflateoperasjoner som ledd i Sjøforsvarets utøvelse av sjøkontroll. Den nye serien etterfølger tidligere MTB-serie som etter hvert nådde sin tekniske og operative levealder. Prosjektet bygger på tidligere utvikling av en prototyp, og må også ses i sammenheng med utviklingen av nye sjømissiler som er utviklet parallelt i et annet prosjekt.

Evalueringsteamet finner at det har vært store forsinkelser, og at kontraktsprisen ble unødig høy på grunn av sen avgjørelse. Dette kan likevel ikke den gjennomførende organisasjonen lastes for. I et operasjonelt perspektiv fremstår prosjektet derfor som middels vellykket. De forsvarsmessige effektmålene knyttet til disse fartøyene vil trolig nås isolert sett, men innenfor gitte budsjettammer for Forsvaret synes anskaffelsen ikke å være særlig relevant, og kan f.eks. innebære at fregattene, som har et større bruksområde, må redusere sin aktivitet. Det er en betydelig risiko for at fartøyene tas ut av tjeneste lenge før tiltenkt levetid utløper. Det lå ikke bare militære men også næringspolitiske hensyn bak investeringen. Det norske verftet som fikk kontrakten ble reddet der og da, men på sikt er

syssesttingsgevinstene høyst usikre. Teamet konkluderer med at prosjektet alt i alt synes å være lite vellykket.

## Lærdommer fra evalueringene

Intensjonen for Concept-programmet som oppdragsgiver har vært å gjennomføre en metaevaluering av de fire evalueringene som trekker ut viktige erfaringer, og også vurderer godheten og svakhetene ved evalueringene og det grunnlag de bygger sine konklusjoner på.

Generelt er det å si at evalueringene av «Kontrollområde Svinesund» og «Skjold MTB» er de to evalueringene som i størst grad lykkes i å følge den målorienterte evalueringsmodellen. Begge er brede evalueringer som bygger på omfattende informasjonsunderlag, befaring og intervjuer med alle sentrale parter. De to andre er mer avgrensede deskstudier med hovedvekt på analyse av produktivitet og samfunnsøkonomisk lønnsomhet, der informasjonsunderlaget til dels er mangelfullt og slik sett ikke gir tilstrekkelig grunnlag for enkelte av de konklusjonene som trekkes.

Det må presiseres at ingen av disse evalueringene skal brukes som modell for hvordan evalueringer av statlige prosjekter bør gjennomføres, men den foreliggende Svinesund-evalueringen og i noen grad Skjold-evalueringen, ligger nærmest å kunne fungere som dette.

Gjennomgangen av de fire evalueringene ga derimot, nettopp fordi kvaliteten var varierende, mange nyttige erfaringer og lærdommer som i sin tur kan brukes til å forbedre opplegget for nye etterevalueringer av store investeringsprosjekter.

Alle evalueringsteamene møtte den utfordring at prosjektet de skulle evaluere ikke kunne ses isolert men var avhengig av andre tiltak. De måtte gjøre en vurdering av hva som er en hensiktsmessig avgrensning av evalueringsobjektet og dermed i hvilken grad tidspunktet for evalueringen var optimalt. Alle teamene valgte en relativt smal prosjektavgrensning. Det kan kanskje kritiseres, men samtidig ble evalueringen på den måten håndterbar og det ble mulig å evaluere investeringen etter få år i drift. Alle fire hadde uansett et sterkt fokus på å se prosjektet i en større helhet.

Et viktig mål med studien var å teste ut evalueringsmodellen, det vil si målorientert evaluering med fem evalueringskriterier, i kombinasjon med samfunnsøkonomisk lønnsomhet. Modellene ble nok i starten oppfattet til dels som konkurrerende fremstillinger av det samme, og som nevnt la de fire teamene ulik vekt på de to modellene. Alle evalueringsteamene kom likevel til at begge modellene er nyttige sammen, og at en bør tilstrebe å kombinere dem. Modellene kan betraktes som

---

ulike måter å sammenstille informasjon på og representerer derfor en form for metodemessig triangulering. Samtidig må en sørge for å unngå at det gjøres dobbeltarbeid.

Det var også noen utfordringer med hver av modellene. Ved bruk av den målorienterte modellen møtte teamene utfordringer med urealistiske mål, overlapp og sorteringsproblemer mellom kriteriene, samt usikkerhet og vanskeligheter knyttet til det å skulle vurdere samfunnsfaglige spørsmål med en grense mot politikk. Ved gjennomføring av den samfunnsøkonomiske analysen hadde teamene betydelige utfordringer knyttet til mangler ved referansegrunnlaget, mangelen på gode data og problemer med å kvantifisere virkningene.

Dels bør disse utfordringene kunne løses med opplæring og veiledning, og med bedre data og større åpenhet om prognoser etc. som er brukt i tidligfasen. Når det gjelder den samfunnsøkonomiske analysen bør en trolig også redusere ambisjonsnivået med hensyn til tallfesting og verdsetting, og heller sikte mot å gjøre en grov og overordnet vurdering av den samfunnsøkonomiske lønnsomheten på samme måte som for de fem andre kriteriene.

Selv om oppgaven var å benytte et felles overordnet evalueringssopplegg, valgte de fire teamene noe ulike tilnæringer. Dette skyldtes dels at prosjektene var ulike, noe som krever ulike valg av indikatorer og datainnsamlings- og analysemetoder for å besvare de seks evalueringskriteriene. Slik bør det være og det er viktig at evalueringssopplegget tillater slik fleksibilitet. Men i tillegg så vi at balansen mellom evalueringskriteriene, og valget av tilnærming hva gjelder datainnsamling og analyse, ble noe forskjellig på grunn av ulik faglig og erfaringsmessig bakgrunn hos evaluatorene. Spesielt var evalueringsteamene for de to samferdselsprosjektene dominert av samfunnsøkonomer med sektorkompetanse. En ser at disse to evalueringene preges av en mer tradisjonell transportøkonomisk tilnærming, med fokus på kvantitative indikatorer. Dette var ikke tilsiktet da studien startet, og skyldes altså trolig at teamene var for homogene. Triangulering av både metoder, kilder, kompetanse etc. er svært viktig – dette visste vi i forkant, men en bør i senere evalueringer ha et enda større fokus på dette. Ikke minst er triangulering mellom kvalitativ og kvantitativ tilnærming viktig. Faktabaserte data, statistikk etc. bør ligge i bunn der slike kan innhentes med rimelig ressursbruk, men intervju med sentrale informanter samt egen observasjon og befaring er også kritisk viktig for en god evaluering.

Det sterke fokuset på kvantitative indikatorer ga imidlertid et viktig funn. Det ble brukt mye tid og ressurser på å forsøke å få ut sentrale nøkkeltall som en vet finnes, eller burde finnes, men som evaluator ikke lyktes i å få tilgang til. Det gjaldt blant annet trafikk tall før-etter på jernbane. Det vil være en viktig utfordring i fremtidige evalueringer å sikre at eksterne evaluatorene får enkel tilgang til slike

nøkkeltall. På flere sektorområder kan det være behov for klarere regler som støtter opp om innsamling, lagring og tilgjengeliggjøring av data.

Det var også en underliggende ambisjon å forsøke å si noe om KS-ordningens betydning for prosjektenes vellykkethet. Dette er selvsagt vanskelig med et utvalg på kun fire prosjekter (ingen kontrollgruppe) og fikk ikke mye fokus i disse evalueringene. Det vi kan si er at alle de fire prosjektene har gjennomgått KS2, og det faktum at tre av fire hadde en sluttkostnad som var innenfor kostnadsrammen, kan tyde på at rammene var realistiske.

## Råd og anbefalinger om evaluering av investeringsprosjekter

Til slutt i denne rapporten har vi listet opp 26 konkrete anbefalinger til senere evalueringer av store statlige investeringsprosjekter, basert på erfaringene fra de fire pilotevalueringene.

Rapporten anbefaler at man fortsetter å benytte den evalueringsmodellen som ble testet ut; med fem kriterier i kombinasjon med samfunnsøkonomisk analyse. En anbefaler altså å vurdere prosjektet med hensyn på: *produktivitet, måloppnåelse, virkninger, relevans, levedyktighet* samt *samfunnsøkonomisk lønnsomhet*. Et slikt minimum av standardisering av fremtidige etterevalueringer vil gi en mulighet for å trekke ut samme type informasjon fra mange evalueringer, og gi læring på tvers, etter hvert forhåpentligvis også om KS-ordningens betydning.

Det er imidlertid nødvendig å holde modellen på et generisk og fleksibelt nivå slik at konkrete indikatorer og datakilder tilpasses det enkelte investeringsprosjekt. Vi er skeptiske til å foreslå ytterligere standardisering, for eksempel i form av faste evalueringsspørsmål og indikatorer, eller faste måleskalaer for å kvantifisere de kvalitative vurderingene, selv om fremtidige evaluatorene kanskje kunne ønske dette. Evalueringer bør gjennomføres av personer med erfaring og riktig kompetanse, og et opplegg basert på standardiserte skjema eller regneark som skal fylles inn er åpenbart ingen god ide. Sentralt i en etterevaluering er å få klarlagt og vurdert alle virkninger som kan tilbakeføres til investeringen, tilsiktede og utilsiktede, kort- og langsiktige, for ulike aktører etc. Dette er viktigere enn i hvilken grad en lykkes med å sortere informasjonen mellom kriteriene og om en greier å kvantifisere den samfunnsøkonomiske lønnsomheten.

Det er viktig at evaluator/teamet er uavhengig, upartisk og har god evalueringskompetanse og erfaring fra evalueringsarbeid. Noe fag- og sektorkompetanse er en fordel, men enda viktigere er tverrfaglighet i evalueringsteamet. Teamleder bør være den som har mest evalueringskompetanse.

Generelt er det å si at evaluering ikke er forskning, og mangelen på data vil trolig *alltid* være et problem gitt de tids- og kostnadsrammene evaluator opererer under.

En god overordnet evaluering bør kunne gjennomføres intensivt på et par måneder og sluttrapporten bør være begrenset mht. sidetall. En bør fokusere på hva som er ”godt nok”, samtidig som en setter høye kvalitetskrav og redegjør for metodevalg og begrensninger. Det innebærer å ha et særlig fokus på validitet, slik at de indikatorene som velges er klart relevante for å besvare evalueringsspørsmålene. Videre peker vi på viktigheten av at nødvendig faktainformasjon fra den enkelte sektor er tilgjengelig for senere evaluatører. Triangulering er svært viktig, og en bør normalt bruke både kvantitativ og kvalitativ informasjon og analyse, flere informasjonskilder, respondenter, etc.

Den enkelte statlige virksomhet og departement må selvsagt finne et opplegg som passer inn i den øvrige monitorerings- og evalueringsvirksomheten, på en slik måte at evalueringene av de store investeringsprosjektene kommer til nytte og bidrar til forbedringer av fremtidige prosjekter.

## Summary

Annual investments in public projects in Norway amount to billions. Examples are roads and rail infrastructure, public buildings, defense acquisitions and large ICT projects. The degree of success and the benefits to society from such investments can only be determined some time after they have entered into the operational face. However, formal evaluations to this effect are seldom done. Obviously, there is a need to know more about the long-term utility of public investments, and hence for carrying out doing systematic ex-post evaluations. The purpose would be to learn from experience both within the responsible ministries and agencies and government institutions, with the aim to improve public investment projects in the future.

The Concept Research Program has a mandate to do trailing research on the major public investment projects that are subjected to external quality assurance under a scheme that was introduced by the Ministry of Finance year 2000. This study reports on the lessons from a pilot evaluation where four external evaluation teams were commissioned to evaluate four projects. These projects represent particularly suited cases for learning since they have all been exposed to systematic ex-ante evaluation up front.

### Methodology for comprehensive project evaluation

An ex-post evaluation should take a broad view of the project which would include both the operational perspective (was it implemented efficiently), the tactical perspective (were the anticipated benefits produced), and the strategic perspective (was it useful to society). Experience is that people's concern is mostly restricted to the operational aspects of a project and less on the tactical and strategic.

In the four pilot evaluations we decided to use the same methodology for all. Firstly, a goal oriented evaluation technique which is widely known and is recommended both by the OECD, the UN and the EU. It is commonly referred to as the OECD model or log-frame evaluation. The model stipulates an assessment of five overall evaluation criteria: that is the project's *efficiency*, *effectiveness*, *impact*, *relevance* and *sustainability*. In addition the teams were asked to perform an economic analysis, which involves an assessment of all *economic benefits and costs* accruing to the project.

---

## Findings from four pilot evaluations

The external evaluation teams were also assigned with a researcher from the Concept Research Program. Four entirely different projects were selected for the study with only one thing in common: they were all about five years into their operational phase:

1. An integrated customs and traffic control facility
2. A section of intercity rail infrastructure
3. A section of a highway
4. The production of five high-speed missile and torpedo vessels

The evaluations were carried out separately and involved a thorough review of existing documentation, field inspection and observation and a series of interviews with the key informants. Findings and conclusions are produced in separate reports. Some brief conclusions are summarized below.

### *The customs and traffic control facility*

The project involved a novel type of integrated control facility at the border to Sweden, designed to serve the purposes of both customs, road authority and police at the country's number one road connection to the European continent. The conclusion was that the project was implemented efficiently and well below budget. It has helped improve border control and reduced traffic congestion, and was considered both relevant and sustainable. It was also considered viable in economic terms.

### *Intercity rail infrastructure*

The project involved 9.5 kilometers double rail tracks, mostly in tunnel, and two railway stations in the vicinity of the capital. The project was implemented efficiently but its economic viability was limited. The anticipated road to rail effect has not materialized as anticipated. Passenger counts remain at the same level. This is partly due to an existing bottleneck in the infrastructure closer to the city which will require further investments if the potential of the IC service should be fully achieved.

### *Section of a highway*

The project involved a six kilometer section of the road between Oslo, Norway and Stockholm, Sweden: it is a small part of a larger scheme to build a highway between the two countries. The project was implemented efficiently although with

cost overrun and is likely to become economically viable once the remaining parts of the highway have been constructed, sometime in the future. It has improved traffic safety and reduced environmental problems locally, but seems to have caused a certain shift in passenger traffic from rail to road, which would be an adverse effect of the investment.

### ***High-speed missile/torpedo vessels***

The project involves the production of five high-speed vessels for the navy. They will be deployed to perform control missions in coastal areas. The project is considerably delayed and the cost is higher than budgeted. Although the vessels seems to successfully achieve anticipated performance and capability there is disagreement within the defense community whether they are relevant within the current defense strategy and given present days restrained military budgets.

## **Lessons and advice**

The scope and quality of the four evaluations varied. One of the main lessons from the exercise is that the combination of economic analysis and a objectives-oriented evaluation is a viable approach to a comprehensive evaluation of large public investment projects. The approach is highly flexible and therefore applicable to all types of investment projects.

Evaluations demonstrated that there are common weaknesses in the reference material and baseline information, such as the formally agreed objectives, estimates of cost and benefits, and also data on performance in the operational face. 26 specific recommendations have been highlighted as the result of this study, based on the experience with the four pilot evaluations.

One is that a minimum of standardization of information will be useful in future evaluations. It will make it possible to draw lessons on specific topics across several projects and within sectors. It is essential though that the model is generic and flexible and that indicators under each evaluation criterion are selected in order to capture essential features of the investment project under study. Further standardization, for example in terms of common evaluation questions, indicators or scales is discouraged. Evaluations, and in particular ex-post evaluations are initiated essentially to review a project in its wider and strategic context. Identifying all impacts that are attributable to the investment project is essential, for all involved and effected parties, both the intended and the unintended ones, the short- and long term ones, etc.



Evaluation teams should be independent and have considerable evaluation expertise, both regarding goal-oriented evaluation and economic analysis. Teams should be interdisciplinary.

Generally speaking, evaluation is not research, and the lack of data will probably always be a problem, given the time and cost constraints facing the evaluators. Evaluations should be carried out intensively during a couple of months, and the final report should be limited with respect to page numbers. One should focus on what is “good enough”, while at the same time set an acceptable quality standard. The focus should be on validity of information rather than precision. Triangulation of information of different types, from different sources, different informants with differing priorities, etc. is one of the most powerful tools evaluators can use in their efforts to ensure that their conclusions are substantiated with information that is both relevant and true.

# 1. Innledning og formål med studien

Concept-programmet utvikler kunnskap som bidrar til bedre konseptvalg, ressursutnyttning og effekt av store statlige investeringsprosjekter. Programmet har i snart 10 år systematisk samlet data om slike prosjekter. Fokus er i hovedsak på tidligfasen og kvaliteten på beslutningsunderlaget. Effektiv læring og tilbakeføring av erfaringer gjør det nødvendig å følge prosjektene over tid, fra tidligfasen via gjennomføring og videre ut i driftsfasen. Det blir helt sentralt å evaluere i hvilken grad prosjektene er effektivt gjennomført og ikke minst hvorvidt en har realisert de tiltenkte effekter for brukere og samfunn. Slike overordnede evalueringer av store statlige investeringsprosjekter gjennomføres i liten grad i dag. Dette er bakgrunnen for at vi i denne studien har prøvd ut en metodikk for overordnet ex post evaluering av store statlige investeringsprosjekter.

## 1.1. Kvalitetssikring, læring og læringsløyper

Det investeres hvert år milliardbeløp i store investeringsprosjekter over statlige budsjetter. Eksemplene spenner fra veier og jernbaneanlegg, til signalbygg i kultursektor, nytt forsvarsmateriell og store IKT-prosjekter. Etter å ha sett seg lei på store kostnadsoverskridelser, forsinkelser og rene feilinvesteringer, etablerte Finansdepartementet fra år 2000 en ordning med ekstern kvalitetssikring av beslutningsunderlaget for de aller største investeringsprosjektene (KS-ordningen)<sup>1</sup>. En startet med å innføre kontrollpunktet KS2 og utvidet deretter med KS1 fra 2005:<sup>2</sup>

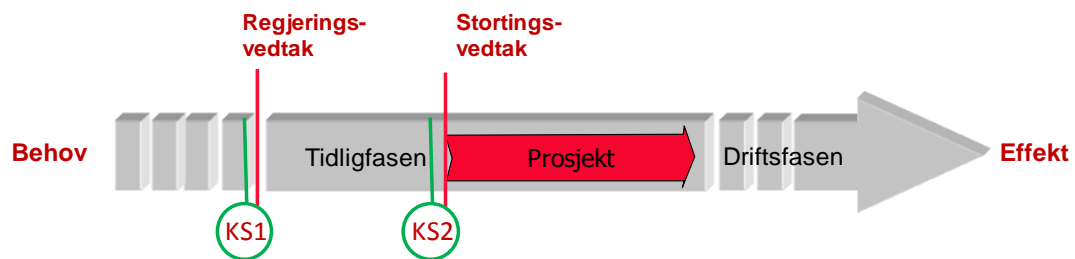
- KS1 - Kvalitetssikring av konseptvalg før vedtak i regjeringen om å starte forprosjekt. Konseptet skal være i tråd med samfunnets behov, mål og krav og det skal gjøres en samfunnsøkonomisk lønnsomhetsvurdering.
- KS2 - Kvalitetssikring av styringsunderlag samt kostnadsoverslag før fremleggelse for endelig investeringsbeslutning i Stortinget.

Til sammen skal de to kontrollpunktene sikre at den faglige kvaliteten på beslutningsunderlaget er slik at myndighetene får bedre overordnet styring med store statlige investeringer, herunder at konseptvalget undergis reell politisk styring.

---

<sup>1</sup> Se Berg m.fl. (1999) hvor behovet for bedre styring av statlige prosjekter dokumenteres

<sup>2</sup> Mer informasjon om KS-ordningen finnes på [www.concept.ntnu.no/ks-ordningen](http://www.concept.ntnu.no/ks-ordningen).



**Figur 1.1** Finansdepartementets kvalitetssikringsordning i to steg; KS1 og KS2

For å få kunne vurdere om kvalitetssikringen (og andre tiltak) virkelig fører til at staten gjør bedre investeringer, er det nødvendig å følge prosjektene over tid, helt inn i driftsfasen. Det gir mulighet for å oppnå to typer av læring, gjerne omtalt som *den lille og den store læringsløyfen*.<sup>3</sup>

Lille og store læringsløyfe er norske uttrykk for begrepene ”single-loop learning” og ”double-loop learning”, som ble lansert av psykologen Chris Argyris og filosofen Donald Schön på 1970-tallet<sup>4</sup>. De så på læring som en prosess for å oppdage og rette opp feil. De slo fast at single loop læring er tilstrekkelig i situasjoner der ting går bra og en kun har behov for å reflektere over i hvilken grad man lykkes med samme type handlinger fra gang til gang. Når ting går prinsipielt galt er det derimot nødvendig med double loop-læring der man stiller grunnleggende spørsmål om gyldigheten av det man gjør og de forutsetningene dette er basert på. Kanskje bør en utforske helt nye og mer kreative løsninger. Double loop læring vil typisk utfordre de mentale kartene som folk har vendt seg til å bruke.

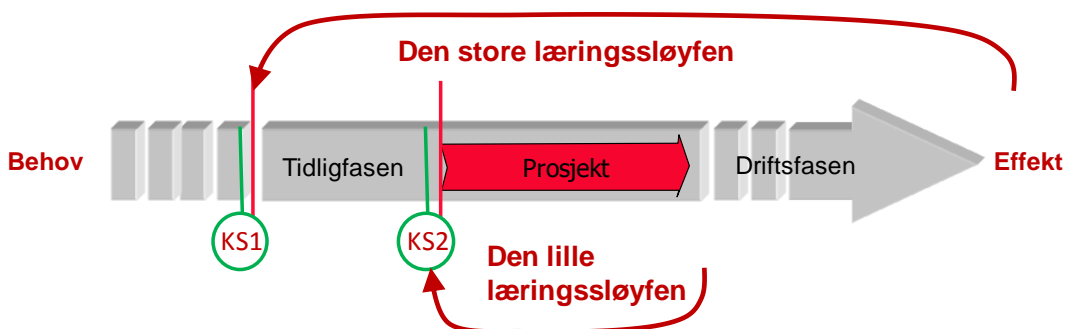
Når det gjelder investeringsprosjekter, betyr dette at:

- Den lille læringsløyfen trekker erfaringer fra avslutningen av prosjektet som gir kunnskap om hvordan man kan forbedre prosjektgjennomføringen. Det handler om å videreutvikle etablert praksis. Denne type læring kan bidra til å bedre tids- og kostnadsstyringen, risikostyringen, kontraktstrategien etc.
- Den store læringsløyfen trekker erfaringer fra etter at resultatet av prosjektet har vært i driftsfasen en tid, og gir overordnet kunnskap om hvorvidt

<sup>3</sup> Andre betegnelser er enkelt- og dobbeltkretslæring, operativ/taktisk og strategisk læring, se Samset og Christensen, 2012

<sup>4</sup> Se Argyris & Schön (1978)

prosjektet var samfunnsnyttig. Det handler om å vurdere godheten av selve konseptvalget. Fikk man realisert de bruker- og samfunnsgevinster som var forutsatt? Eller burde man valgt noe helt annet enn dette prosjektkonseptet?



**Figur 1.2** Den store og den lille læringsløyfen

Begge typer læring er viktige for å forbedre statlige investeringsprosjekter. Likevel er det spesielt behovet for å styrke den *store* læringsløyfen som motiverer denne studien. Den lille læringsløyfen er det allerede betydelig fokus på. Prosjektets resultatmål er vanligvis konkrete og målbare, og det finnes et vell av verktøy som skal hjelpe prosjektleder til å sikre at disse nås. Det er derimot mindre tradisjon for å fokusere på den store læringsløyfen, og foreløpig er det lite systematisk kunnskap om hvor godt statlige investeringsbevilgninger omsettes til samfunnsnytte i Norge.

Denne studien skal tilrettelegge for systematisk evaluering av KS-prosjektene, som vil bidra med kunnskap om dette. Det overordnede målet som er utgangspunktet for denne studien er altså bedre læringseffekter av store investeringsprosjekter i staten. Læringen bør styrkes både internt i etater og departementer, men også på tvers i staten.

## 1.2. Ex post evaluering som læringsarena

Det statlige økonomiregelverket krever at det gjennomføres evalueringer som et supplement til den løpende mål- og resultatstyringen av statlig virksomhet<sup>5</sup>. Dette kravet omfatter både investeringsprosjekter og annen virksomhet. Frekvens og omfang av evalueringene må bestemmes ut fra prosjektenes egenart, risiko og

<sup>5</sup> Reglement for økonomistyring § 16. Se også Finansdepartementet (2005b).

---

vesentlighet. En evaluering kan derfor være alt fra enkle stikkprøver til omfattende undersøkelser som likner forskningsprosjekter.

Hensikten med en evaluering er i prinsippet tredelt, den skal bidra til *kontroll*, *styring* eller *læring*. Vektleggingen kan være forskjellig i ulike evalueringer. Som allerede nevnt vil evalueringene i vårt tilfelle først og fremst fokusere på *læringsaspektet*. Å undersøke hva som skjer med investeringsprosjektene etter at de er gjennomført, gir kunnskap om hva som er fornuftige valg i fremtiden. I tillegg vil det at det her gjelder prosjekter som beløper seg til hundretalls millioner av kroner, være et selvstendig argument for at det skjer en viss form for *kontroll*, hovedsakelig ved at en fremskaffer sentral og pålitelig informasjon. Når det gjelder *styringsaspektet*, handler det om å fremskaffe resultater som gjør at prosjektet kan ledes videre på rett vei. Det siste er lite relevant ved ex post evaluering, ettersom prosjektet er avsluttet.

Erfaringer generelt tilsier at en bør bli flinkere til å bruke evalueringer til læringsformål, dette gjelder ikke minst i staten (Samsøt og Christensen, 2012). Når evalueringer gjennomføres skjer det ofte ad hoc når noe allerede har gått galt. Det er et paradoks at en ofte bruker mye tid på detaljerte utredninger og planer i forkant, men ikke legger inn ressurser for å foreta systematisk oppfølging av hva som skjer i etterkant. En forklaring kan være at evalueringer kan være ressurskrevende, samtidig som man ikke ser den direkte nytten – prosjektet er jo uansett avsluttet. Mer generelt er en i staten antakelig mer opptatt av å etablere neste års budsjett enn å kikke i bakspeilet for å finne ut hva en fikk ut av pengene som ble fordelt året før. Da går en glipp av viktig læring.

Imidlertid har en i mange år sett en tiltakende satsing på evaluering i det offentlige, og interessen for ex post evaluering har økt. Som et godt eksempel kan nevnes at Samferdselsdepartementet i 2006 tok initiativ til en ordning for etterprøving av de viktigste prissatte virkningene av et utvalg vei- og jernbaneinvesteringer. Dette for å gi et bedre faglig grunnlag for planlegging av fremtidige samferdselsprosjekter<sup>6</sup>.

Tilsvarende vil det være svært nyttig å gjennomføre systematiske evalueringer av KS-prosjektene i etterkant, som jo er Norges største statlige investeringsprosjekter. Men samtidig vil det være hensiktsmessig å begrense ambisjonsnivået. Dersom evalueringene krever for mye tid og ressurser vil de trolig ikke bli gjennomført. Det bør derfor legges en begrensning både på omfanget av det som gjøres i hvert

---

<sup>6</sup> Se seneste omtale i Stortingsproposisjon nr. 1 (2011-12), Samferdselsdepartementets budsjettproposisjon s. 189. Det er også laget en veileder for etterprøving av store veiprojekter, Statens Vegvesen (2006).

enkelt tilfelle og detaljeringsnivået. Men først og fremst bør evalueringen ha et hovedfokus på det som gjelder den store læringsløyfen og ikke gå inn i utførlig granskning av enkelte utvalgte prosesser eller prosjektleveranser. Det er de overordnede spørsmålene som skal avklares, det vil si nytten av investeringstiltakene og om man har gjort fornuftige valg. Dersom den overordnede evalueringen viser at det er behov for å følge opp med dybdestudier på avgrensede områder burde en gjøre dette i etterkant.

### **1.3. Fire pilotevalueringer og forslag til et enkelt opplegg for evaluering**

Concept-programmet har lenge vært opptatt av effektvurderinger av store investeringsprosjekter. Vi har studert prinsipper for hvordan dette kan gjøres, og erfaringer fra evalueringsvirksomhet i privat og offentlig sektor i Norge og andre land.<sup>7</sup> Programmet har drevet følgeforskning på prosjekter under KS-ordningen siden år 2002. I hele denne perioden har det ikke fantes informasjon for å kunne si noe om effekter og måloppnåelse i bred forstand for disse prosjektene. Årsaken er at det tar lang tid å planlegge og gjennomføre store investeringsprosjekter. De første prosjektene som gjennomgikk KS2 i år 2000, har i ettertid vært igjennom en planleggings- og byggefase som typisk tar 5-10 år. Dette er markert med den blå linjen i figur 1.3. Som figuren viser, var status per utgangen av 2011 at bare 36 prosjekter var ferdigstilt og kommet i driftsfasen. I tillegg bør en vente 3-5 år etter ferdigstillelse med å evaluere, slik at en er kommet inn i normal drift og det er mulig å studere effektene for brukere og samfunn.

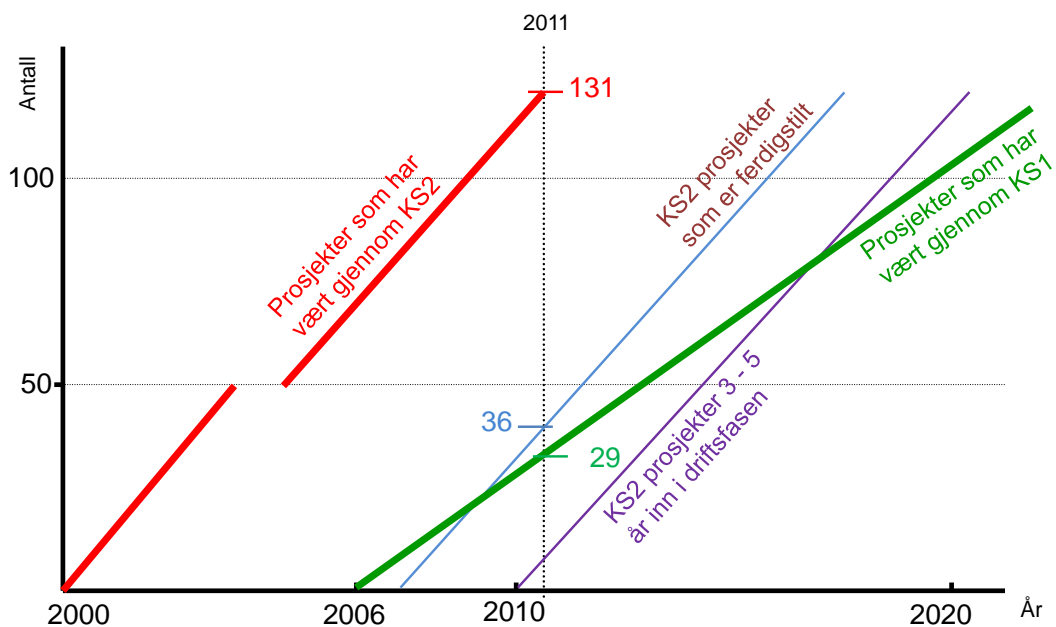
Det er altså først nå det er mulig å gjøre en overordnet, bred evaluering av de første prosjektene. Dette er antydnet i figuren med en fiolett linje. Imidlertid gjelder dette bare prosjekter som har gjennomgått KS2, det vil si prosjekter der en ikke har hatt en kvalitetssikring av underlaget for konseptvalget. Kontrollpunkt KS1 ble ikke innført før i 2005, i figur 1.3 markert med den grønne linjen. Disse prosjektene skal først gjennom en KS2-kontroll, deretter en planleggings- og gjennomføringsfase og tilslutt en driftsfasen på 3-5 år før de kan evalueres. Det sier seg selv at det ennå vil ta tid før en kan ta fatt på evalueringen av disse.

Dette betyr at de KS-prosjektene vi nå har mulighet til å evaluere kan betegnes som en kontrollgruppe av prosjekter der fokus i utgangspunktet var på risiko og kostnadsstyring, men ikke på selve konseptvalget. Både disse prosjektene og KS1 prosjektene som etter hvert kommer i driftsfasen vil bli evaluert etter samme mal, der både vurderingen av konseptvalget og effekten på sikt vil stå sentralt. Dette vil

---

<sup>7</sup> Se for eksempel Concept-rapportene nr. 7 og nr. 19.

gi en mulighet for å vurdere om KS ordningen slik den er i dag kan sies å ha en positiv effekt både på konseptvalget og gjennomføringen av prosjektene.



**Figur 1.3.** Antall prosjekter som har vært gjennom KS1 og KS 2 over tid og prognose for årene som kommer. Prinsippkisse. Status ved utgangen av 2011 var at henholdsvis 131 og 29 prosjekter hadde vært gjennom KS2 og KS1

I denne studien har vi tatt for oss et utvalg på fire KS-prosjekter:

1. Svinesund kontrollområde
2. Nytt dobbeltspor Asker-Sandvika
3. Veiparsell E 18 Momarken-Sekkelsten
4. Bygging av Skjold-klasse MTBer til Sjøforsvaret

Ved utvelgelsen har en vært ute etter å se på svært ulike prosjekter. Utvelgelsen er gjort i samråd med Finansdepartementet, og er forelagt de berørte departementene. Det viktigste kriteriet for utvalgelse har vært at prosjektene er ferdigstilt og kommet i driftsfasen. Et annet hensyn er ønsket om å få en spredning på sektor og prosjekttypen. Samferdselsprosjekter utgjør om lag halvparten av alle prosjekter i KS-ordningen, forsvarsprosjekter omtrent en fjerdedel, og den siste fjerdedelen omfatter blant annet ulike større byggeprosjekter. Utvalget gjenspeiler dette forholdet. Ingen av tiltakene antas å skille seg nevneverdig fra andre store prosjekter på sitt sektorområde, men en har heller ikke noe grunnlag for å si at de er representative for sektoren.

De fire evalueringene omtales her som ”pilotevalueringer”. Dette fordi den primære hensikten har vært å teste ut et overordnet evalueringsopplegg. Viktige forskningsspørsmål har vært: Er metodikken egnet for overordnet evaluering av store investeringsprosjekter? Passer den i så fall for alle typer prosjekter? Hvilke utfordringer bør en være oppmerksom på? Målet er å sikre gode evalueringer som gir læring til senere investeringsprosjekter i de forskjellige sektorer. I tillegg til å få kunnskap om prosjektenes vellykkethet, vil man søke læring om alle tema som er sentrale for tidligfasevurderinger, f.eks. om mål og målformulering, utnyttelse av mulighetsrommet, dokumentasjonens omfang og detaljeringsgrad, behov for gode prognoser, etc.

I denne studien har tre eksterne fagmiljøer vært engasjert for å evaluere hvert sitt prosjekt etter denne felles, overordnede evalueringsmodellen. For å sikre læringsutbyttet har Concept hatt ansvar for én evaluering i tillegg og hatt med egne forskere i hvert av teamene. Prosjektene er valgt ut i samråd med Finansdepartementet. De har alle vært om lag fem år inne i driftsfasen, men er for øvrig helt ulike. Fagmiljøene er:

- SINTEF Teknologi og samfunn. Prosjektleder Håkon Finne (Svinesund kontrollområde)
- VTI (Statens väg- och transportforskningsinstitut). Prosjektleder Johan Nyström (Nytt dobbeltspor Asker-Sandvika)
- Scanteam AS. Prosjektleder Erik Whist. (Bygging av Skjold-klasse MTB-er til Sjøforsvaret)
- Concept-programmet internt. Prosjektleder Torbjørn Aass. (E18 veiparsell Momarken-Sekkelsten)

Studien ble gjennomført i perioden oktober 2011-oktober 2012. Det ble avholdt en felles samling ved oppstarten av arbeidet, der man drøftet det felles evalueringsopplegget og metodikken som skulle benyttes. Hvert team utarbeidet deretter en mer detaljert plan for sin evaluering, ut fra særtrekk ved prosjekt, tilgjengelige data, og teamets faglige og erfaringsmessige bakgrunn. Det ble også avholdt en felles samling mot slutten av arbeidet, hvor de fire teamene presenterte foreløpige resultater og viktige erfaringer for hverandre. Til slutt gjennomførte Concept en metaevaluering og trakk ut felles erfaringer og anbefalinger, basert på de fire pilotevalueringene.

Resten av denne rapporten dokumenterer studien og våre erfaringer og anbefalinger. I kapittel 2 presenterer vi det metodiske opplegget som ble benyttet. Kapittel 3 gir en sammenstilling av resultatene av de fire pilotevalueringene. I kapittel 4 drøfter vi erfaringer og lærdommer en har fått med evalueringsmetodikken, og vi peker på utfordringer og andre forhold som det er særlig viktig å ha fokus på i senere evalueringer. Kapittel 5 avslutter med vår



anbefaling til et forenklet opplegg for overordnet evaluering av investeringsprosjekters vellykkethet.

## 2. Metodikk for prosjektevaluering

I evalueringsvirksomhet skiller en gjerne mellom målorientert evaluering og prosessorientert evaluering. I det første tilfellet er evalueringen ofte hypotetisk deduktiv, dvs. man tar man utgangspunkt i antatte konsekvenser (for eksempel uttrykt som mål) som så testes mot observerbar virkelighet når evalueringen foretas. Dette gjør det forholdsvis enkelt å velge data og tolkningsmuligheter, men det kan også føre til at perspektivet avgrenses og at en overser viktige sider ved det som skal evalueres. Et annet problem er at mål ofte er uklart eller utilstrekkelig formulert og lite egnet som utgangspunkt for analysen som skal gjøres. Dette er drøftet nedenfor.

Prossessorientert evaluering handler om å evaluere tiltak som ikke er klart avgrenset med hensyn til omfang og målorientering. Slike evalueringer vil i større grad bruke en induktiv tilnærming, dvs. man starter med mer uavklarte og spørrende forestillinger. Ved observasjon og undersøkelser reises nye temaer som krever ny kunnskap og som etter hvert fører til ny innsikt. Induktiv tilnærming benyttes også når en på forhånd har lite kjennskap til fenomenet som skal evalueres. Den er spesielt egnet til å oppnå såkalt dybdeforståelse, men er også derfor forholdsvis ressurskrevende.

I det foreliggende tilfellet handler det om å evaluere prosjekter. Ettersom dette er fenomener som er begrenset både i omfang og tid og godt definert på forhånd, har vi antatt at målorientert evaluering vil være mest hensiktsmessig. Samtidig bør evalueringen også ha elementer av induktiv tilnærming, spesielt er det viktig for å identifisere de utilsiktede virkningene av prosjektene.

I dette kapitlet drøfter vi det metodiske opplegget som pilotevalueringene er basert på. Resten av kapitlet presenterer dette opplegget<sup>8</sup>. I avsnitt 2.1. drøfter vi hva som menes med et "vellykket" investeringsprosjekt. I avsnitt 2.2. argumenterer vi for at en kombinasjon av målorientert evaluering og samfunnsøkonomisk analyse (ex post) vil være et godt utgangspunkt for evaluering av store, statlige prosjekter. Dette gir til sammen seks overordnede kriterier for vellykkethet. Siste del av kapitlet (avsnitt 2.3) diskuterer andre viktige momenter når en skal lage gode indikatorer og tilhørende datainnsamlings- og analysemetoder for disse seks kriteriene. Avsnitt 2.4. oppsummerer om metodisk tilnærming.

---

<sup>8</sup> Dette kapitlet er i stor grad basert på lærebok i prosjektevaluering, Samset (2003), supplert med retningslinjer for samfunnsøkonomisk analyse jevnfør Finansdepartementet (2005a), samt generell veiledning til evaluering av statlige virksomhet jevnfør Finansdepartementet (2005b) og DFØ (2008).

## 2.1. Definisjon av "vellykket" prosjekt

Ex post evaluering av prosjekter handler til syvende og sist om å vurdere i hvilken grad en har lyktes på sikt. Hva som legges i begrepet vellykket vil avhenge blant annet av hvem man spør, hvilket ambisjonsnivå som var lagt til grunn og hvilket tidsperspektiv man har hatt. Det kan argumenteres for at vellykkethet kan – og bør – vurderes både i et operasjonelt, taktisk og strategisk perspektiv. Når det gjelder store, statlige investeringer bør en ha et bredt og samfunnsmessig perspektiv på dette.

### Prosjektet når sine mål og genererer nytte

Prosjektet er en måte å organisere et arbeid på, for å nå bestemte mål<sup>9</sup>. Gitt at målene er relevante og viktige for samfunnet, kan vi si at et vellykket prosjekt er kjennetegnet ved dets *måloppnåelse*. Dette betyr at en greier å realisere noen ønskede *effekter*, som i sin tur bidrar til å dekke de *behovene* som var utgangspunkt for prosjektet. Begrepene behov, mål og effekt henger altså nært sammen, som drøftet i Concept rapport nr. 9 (Næss, 2004b).

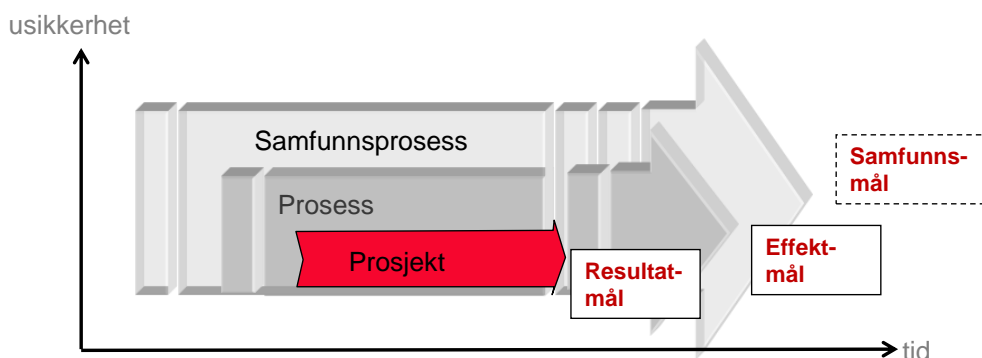
### *Ulike perspektiv på måloppnåelse*

I prosjektfaglig litteratur skilles det mellom tre nivåer av måloppnåelse: *resultatmål*, *effekt mål* og *samfunns mål*:

- Resultatmål er knyttet til selve prosjektleveransen. Dersom for eksempel prosjektet er å bygge en vei, vil resultatmålet være den ferdige veien definert ved omfang (lengde, bredde), tids- og kostnadsrammer og avtalt kvalitet (veistandard, type asfalt etc.).
- Med effekt mål mener vi første ordens effekt av prosjektet. Viktige effekt mål kan være økt kjørekomfort og tidsbesparelser for de som bruker veien. Realiseringen av effekt målene avhenger i mange tilfeller ikke bare av prosjektet alene, men er også avhengig av andre tiltak og prosesser, følgeprosjekter etc.
- Samfunns mål er enda bredere, for eksempel den økonomiske effekten av veiinvesteringen innenfor en gitt region. Samfunns mål realiseres på lang sikt, og er resultat av andre prosesser i tillegg til de som følger direkte av prosjektet.

---

<sup>9</sup> A project is a temporary endeavor undertaken to create a unique product or service (Project Management Institute, PMI).



**Figur 2.1** Tre nivåer av prosjektmål, basert på Samset (2003)

Måloppnåelse på resultatnivå kan vurderes straks prosjektet er avsluttet og leveransen foreligger. Måloppnåelse på effektnivå må normalt vurderes et stykke ut i driftsfasen. Måloppnåelse på samfunnsnivå kan sjelden måles direkte, men indikasjoner på måloppnåelse kan belyses på ulike vis. Realisering av samfunnsmålet skyldes mange forhold, hvorav prosjektet bare er ett.

Ulike aktører har forskjellige interesser i prosjektet og dermed ulike perspektiv på hva de mener er måloppnåelse. Prosjektleder blir målt på om resultatmålet nås, og vil selvsagt definere egen suksess kun ut fra dette. Viktige brukergrupper samt den finansierende part og samfunnet for øvrig er derimot mer interessert i måloppnåelse på effekt- og samfunnsnivå.

### ***Innvirkningen på andre mål samfunnet har***

Prosjektets samfunns mål vil være en viktig rettesnor når en skal utforme et offentlig tiltak. Samfunnet kan imidlertid ha flere overordnede mål, som direkte eller indirekte berøres av tiltaket. Ikke sjelden er ulike samfunns mål endog i strid med hverandre.

Et veiprosjekt som bidrar til tidsbesparelser for folk og næringsliv i regionen, er positivt sett i forhold til målet om regional økonomisk vekst. Men det samme veiprosjektet kan over tid virke negativt på samfunnets mål knyttet til miljø, både klimapolitiske og lokale mål. Dersom dette ignoreres kan tiltaket risikere å miste sin relevans over tid, i alle fall blant enkelte grupper. I samferdselssektoren har en tatt tak i dette ved å etablere et sett relevante samfunns mål i Nasjonal

transportplan, hvor departement og etat må finne gode avveininger mellom dem i det enkelte prosjekt.

Et annet eksempel på dilemma en kan stå overfor ved utforming av statlige tiltak, er hvordan forene målet om effektiv bruk og dermed nytte av en tjeneste, med målet om å støtte grupper med lav betalingsevne (fordeling). Dette gjelder også i forbindelse med investeringer i infrastruktur, hvor finansieringen av utbygging og drift ofte er et tema. Innslag av brukerbetaling må vurderes i forhold til begge samfunnsmålene.

Ved evaluering av et enkeltprosjekt, i motsetning til et helt sektorområde eller reguleringsområde, kan det likevel være rimelig å sette ett samfunns mål i høysetet, dersom dette var direkte utløsende og som investeringstiltaket klart mest direkte var innrettet mot å nå. Samtidig er det alltid viktig å ta hensyn til mål også i sektorer som blir berørt av prosjektet i form av *sidevirkninger* (se Concept-rapport nr. 5, Næss, 2004a). Ikke minst må en ex post evaluering vurdere hvilke *utviklede* virkninger prosjektet kan ha hatt, både positive og negative.

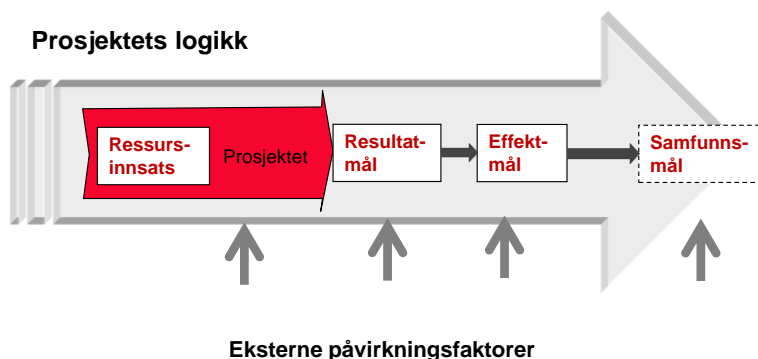
For å sikre faglig bredde i vurderingene, og for å identifisere ulike og kanskje motstridende mål og prioriteringer hos berørte parter bør en anlegge forskjellige faglige perspektiver i analysen. Internasjonalt brukes betegnelsen ”cross sectoral view”, og legger til grunn at følgende sektorielle perspektiver bør inngå i analysen av ethvert prosjekt, OECD (2002):

- Politisk forankring og hjemmel i lovverk, policy etc.
- Økonomiske og finansielle aspekter
- Samfunnsmessige aspekter, herunder fordelingsmessig, grupperes rettigheter, plikter og tilgang til viktige goder
- Miljømessige konsekvenser
- Institusjonelle aspekter som relasjoner innenfor og mellom organisasjoner
- Teknologiske forhold.

## En realistisk prosjektlogikk er en forutsetning

Et hvert prosjekt bygger på en oppfatning om en underliggende kausal sammenheng mellom de viktigste hendelsene eller målene. De ressursene som inngår utløser aktiviteter som resulterer i at resultatmål realiseres, dernest effektmål og tilslutt samfunns mål. Dette kalles prosjektets logikk eller strategi. Hvert ledd i

kjeden påvirkes av eksterne faktorer som kan true realiseringen av strategien eller deler av denne. Dette er illustrert i figur 2.2.



**Figur 2.2** Prosjektets logikk, eller strategi

I en målorientert evaluering vil utgangspunktet være at evaluator gjør seg kjent med prosjektets logikk. Evaluator må vurdere om årsak-virkning-sammenhengen mellom målene er logisk og om disse er realistiske. Formelt sett vil et prosjekt være vellykket dersom det leverer effekter som stemmer overens med målene, er i samsvar med de behovene som var utløsende for prosjektet, og er gjennomført så effektivt som råd er. Reelt sett oppstår det et problem dersom det formelt avtalte ikke er realistisk eller relevant, noe som ofte er tilfellet. Dersom de formelle målene ikke er realiserbare vil prosjektet formelt sett være mislykket selv om det reelt sett kanskje er svært vellykket vurdert i forhold til hva en med rimelighet kunne forvente.

Store investeringsprosjekter er komplekse og har gjerne mange mål som i større eller mindre grad er gjensidig avhengige. Praksis viser at målhierarkier ofte har vesentlige feil. Det kan være at de er uklart beskrevet, at det er en sammenblanding og forveksling av mål på ulike ambisjonsnivå og at enkelte mål er for ambisiøse eller urealistiske. I et målhierarki opererer en med to typer mål. Det er de som er koblet i årsak-virkningsforhold og de som er parallelle på samme nivå. Plasseringen i kjeden angir ambisjonsnivå. Ambisjonsnivået kan testes ved å vurdere sannsynlighet for realisering på hvert nivå. Parallelle målsettinger er mål en antar må realiseres samlet for å nå målet på neste ambisjonsnivå. Et meget vanlig problem er at mål presenteres i sammensatte setninger som inneholder flere mål som både er koplete og parallelle. Det vil derfor være en vanlig utfordring for evaluator å bryte opp slike komplekse utsagn og avklare hva som er det formelle grunnlaget for prosjektet. Eksempler på slike analyser er vist i Concept-rapport 21 (Samset m.fl, 2009), kapittel 5. En ytterligere kompliserende faktor er at prosjektets logikk noen ganger endres underveis i gjennomføringen, enten formelt eller i form av en stilltiende enighet mellom aktørene. Da bør også dette dokumenteres.

En gjennomgang, vurdering og eventuell revisjon av det formelt avtalte grunnlaget for prosjektet er i mange tilfeller avgjørende for å kunne utforme et hensiktsmessig design for evalueringen og velge fornuftige indikatorer. Dette vil være en oppgave for evaluator, som må tydeliggjøre denne for de berørte parter og oppnå enighet om hva som skal være det formelle grunnlaget for evalueringen før denne settes i gang.

## Operasjonell, taktisk og strategisk vellykkethet

Samlet kan vi snakke om investeringsprosjekters vellykkethet på tre nivåer (Samset, 2003) som henspeiler på de tre nivåene av måloppnåelse nevnt tidligere:

- *Operasjonell* vellykkethet. Leveransen foreligger som avtalt - og er produsert på mest mulig kostnadseffektiv måte.
- *Taktisk* vellykkethet. Tiltaket gir opphav til størst mulig nytte/måloppnåelse for brukerne, sidevirkninger er optimalisert – og tiltaket er formåls effektivt.
- *Strategisk* vellykkethet. Tiltaket bidrar samlet sett til en samfunnsmessig ønsket utvikling over tid - og tiltaket (egentlig pakken av tiltak) som har frembrakt denne utviklingen er samfunnsøkonomisk effektiv.

Det er lett å finne eksempler på prosjekter som har vært meget vellykkede i et operasjonelt perspektiv, men mislykkede i et taktisk og strategisk perspektiv. Det gjelder prosjekter der gjennomføringen har vært upåklagelig og leveransen er fremskaffet med lavere ressursbruk enn forventet – men hvor det ferdige produktet ga lite nytte for brukere og samfunn. Et prosjekt kan også være taktisk vellykket, men samtidig strategisk mislykket, eksempelvis dersom brukergruppen ikke egentlig var prioritert av storsamfunnet og/eller hvor deres behov kun ble tidsfredsstilt på kort sikt.

Motsatt kan en finne mange eksempler på at et prosjekt som var operasjonelt mislykket (for eksempel på grunn av store kostnadsoverskridelser) likevel viser seg å være strategisk vellykket, fordi i et langsiktig perspektiv er investeringskostnaden kun en liten bit av bildet – de langsiktige gevinststrømmene er viktigere. Vi innser derfor at strategisk vellykkethet er det ”viktigste” aspektet i en evaluering og det vi er mest interessert i å få vurdert.

Det vil være viktig å unngå at det brukes for mye ressurser på evalueringen av det operasjonelle perspektivet. Dette er et vanlig problem ettersom det på dette området finnes mest og lettest tilgjengelig informasjon. Langt mer krevende er å lage et evalueringsdesign som fanger opp de viktigste effektene på taktisk og strategisk nivå.

En vil ofte møte bestemte oppfatninger hos berørte om i hvilken grad prosjektet er vellykket, på hvilken måte og hva som er de sentrale problemene. Det vil være en utfordring for evaluatør å ikke bli for påvirket av slike forutinntatte synspunkter når evalueringsdesignet utformes, men å kunne åpne perspektivet, for eksempel ved å studere erfaringene fra evaluering av lignende prosjekter.

## 2.2. Valg av overordnet evalueringsmodell

Denne pilotstudien skal komme frem til en metodikk for en bred og overordnet, og mest mulig enkel vurdering av investeringsprosjekters vellykkethet. Det finnes mange ulike evalueringsmodeller. Se for eksempel Concept-rapportene nr. 7 og 19 for en oversikt over modeller, med særlig fokus på metoder for å måle og følge opp effekter (måloppnåelse). Vi har tidligere nevnt målorientert evaluering og prosessevaluering som to hovedtyper med henholdsvis deduktiv og induktiv tilnærming. I tillegg finnes modeller som særlig fokuserer på bestemte type resultater eller tema, for eksempel økonomiske modeller (fokus på effektivitet på ulike nivåer), systemmodeller (fokus på betydningen av organisering, ledelse, kompetanse, teknologi etc.) og aktørmodeller (fokus på bestemte aktører og interessenter). Finansdepartementet skiller i sin veileder i evaluering mellom hhv tiltaksanalyse, effektanalyse, årsaksanalyse, prosessanalyse, måloppfyllelsesanalyse og effektivitetsanalyse (se Finansdepartementet, 2005b).

Alle analyser som er gjennomført i tidligfasen vil dessuten kunne danne utgangspunktet for en resultatoppfølging i ettertid. Det kan være utredninger av avgrensede tema som miljøkonsekvensanalyse, livssyklus kostnadsanalyse, finansieringsanalyse eller utredning for eksempel vedrørende lokal sysselsetting og bosetting, inntektsulikheter, likestilling eller lignende. Det kan også være brede konsekvensutredninger som samfunnsøkonomisk analyse eller analyse av bærekraft (Sustainability Impact Assessment).

I denne studien har vi valgt å ta utgangspunkt i en evalueringsmodell som anbefales av OECD som den primære tilnærmingen, supplert med samfunnsøkonomisk analyse som sekundær tilnærming. OECD-modellen er en målorientert modell, som er generisk og derfor anvendbar på alle typer tiltak, og samtidig fleksibel og overordnet nok til å omfatte nærmere analyse av bestemte tema eller aktører dersom det skulle være behov for dette. Samfunnsøkonomisk analyse er en økonomisk modell, men også den er så bred at den er ment å dekke alle typer virkninger på samfunnet, ikke bare finansielle og økonomiske aspekter. Vi mener at disse to modellene til sammen gir en god og dekkende vurdering av tiltakets vellykkethet både i et operasjonelt, taktisk og strategisk perspektiv. Dette begrunnes nærmere under, samtidig som vi gir en kort introduksjon til modellene.



I tillegg til modellenes evne til å måle vellykkethet, er det et selvstendig argument at KS-ordningen (KS1) langt på vei er fundert på disse to.

## Målorientert evalueringsmodell

Det som her omtales som målorientert evalueringsmodell ble etablert av Leon J. Rosenberg (Rosenberg m. fl. 1970) og er senere spredt i betydelig omfang internasjonalt, og etablert som anbefalt metode blant annet i FN, EU og OECD (OECD, 2002). I norsk sammenheng er det særlig bistandsforvaltningen som har tatt metoden i bruk i forbindelse med evalueringer av prosjekter i utviklingsland.

OECD's definisjon av fenomenet evaluering er følgende:

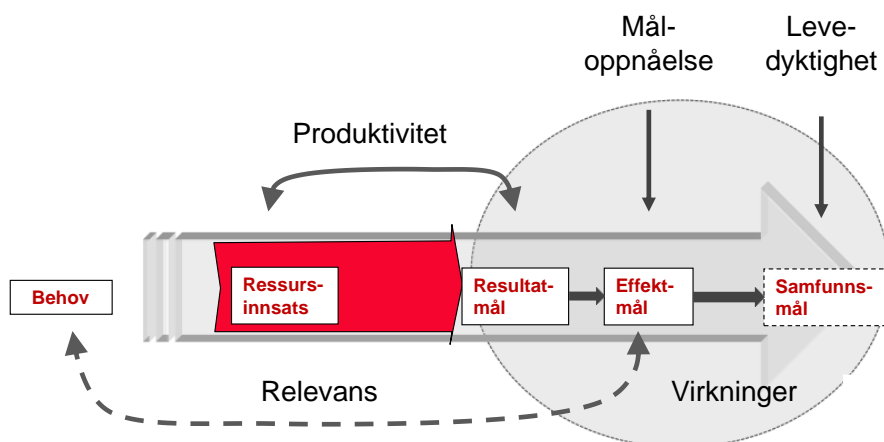
Evaluering er en undersøkelse, så systematisk og objektiv som mulig, av et igangværende eller avsluttet prosjekt, program eller en prosess – av utformingen, gjennomføringen og resultatene. Formålet er å fastslå måloppnåelsen og hvorvidt målene er relevante, om gjennomføringen har vært effektiv, de samlede virkningene av tiltaket, og levedyktigheten på lengre sikt. Evalueringen skal fremskaffe informasjon som er etterrettelig og nyttig, og bidra til at læring kan tilføres beslutningsprosessene” (OECD, 1999).

Dette sammenfatter det som er modellens fem overordnede evalueringskriterier; *produktivitet*, *måloppnåelse*, *relevans*, *virkinger* og *levedyktighet*. Disse er vist i figur 2.3. og beskrevet nærmere nedenfor. Modellen har et bredt perspektiv på prosjektet. Vi kan si at *produktivitet* måler operasjonell vellykkethet, *måloppnåelse* måler taktisk vellykkethet, mens *relevans*, *virkinger* og *levedyktighet* måler strategisk vellykkethet.

Metoden og bruken av denne er grundig beskrevet i en egen håndbok (Samset, 2003). De fem evalueringskriteriene kan sammenfatningsvis beskrives som nedenfor:

### Produktivitet

Dette handler om gjennomføringen av selve investeringsprosjektet, og spørsmålet om hvor effektivt prosjektorganisasjonen har omsatt ressurser til leveranser. En må spørre om resultatmålene for henholdsvis kostnad, tid og kvalitet ble nådd som avtalt. Det bør i prinsippet være det en lettest kan svare på, da det normalt vil utarbeides sluttrapport som viser prosjektrekskap, tidslinjer og innhold i leveransene. I tillegg må en vurdere om arbeidet kunne vært gjort billigere, raskere eller med bedre kvalitet. Det siste vil kreve en referansesjekk mot andre liknende prosjekter, samt en gjennomgang av hvilke usikkerhetsfaktorer som har slått til og hvordan disse er håndtert.



**Figur 2.3** Målorientert evalueringsmodell med fem evalueringskriterier

### Måloppnåelse

Når vi snakker om måloppnåelse gjelder dette realiseringen av effektmålet, dvs. behovstilfredsstillelse for de aktuelle brukergrupper. En spør om de avtalte effektmålene ble nådd, og i hvilken grad prosjektet vurderes å ha bidratt til dette. Dersom det avtalte målet er lite egnet til evaluering, for eksempel fordi det er trivielt eller urealistisk, bør evaluator avklare hvilket ambisjonsnivå man legger til grunn. Hvis mulig bør en benchmarke mot andre tilsvarende eller liknende tiltak for å avklare hva som er et rimelig ambisjonsnivå.

### Virknings

Prosjektet kan også ha utilsiktede virkninger Disse må dokumenteres og en må da tenke bredere og mer kreativt og vurdere de samlede konsekvensene av tiltaket, positive og negative, tilsiktede og utilsiktede, kort- og langsiktige, for målgruppen og for andre berørte parter. Dette er ofte en krevende del av evalueringen. En skal bare ta med de virkningene som med rimelighet kan tilbakeføres som resultat av tiltaket.

### Relevans

Et prosjekt er relevant dersom det er i samsvar med viktige prioriteringer i samfunnet og for viktige brukergrupper, altså dersom det er *behov* for det prosjektet leverer. Oppnåelse av formelt avtalte mål er ikke synonymt med relevans – men at disse målene er i samsvar med brukernes og samfunnets behov eller prioriteringer. Dette krever altså igjen en kritisk vurdering av de formelt avtalte målene. Hvorvidt målene fremdeles er relevante kan også variere over tid, dersom behovene eller etterspørselen etter de ytelser eller goder som prosjektet gir opphav til endrer seg. En utfordring med dette kriteriet er at brukernes og samfunnets behov kan være uavklart eller motsetningsfylt. Evaluators må derfor bruke sunn fornuft og vurdere

de formelle målene i forhold til kunnskap om politiske prioriteringer, marked og trender, lover og bestemmelser.

### **Levedyktighet**

Det langsiktige, strategiske perspektivet omtales gjerne som levedyktighet (eng. sustainability). En spør da om de positive effektene av tiltaket kan vedvare over hele levetiden, og om nytten for samfunnet er positiv på sikt. Dette har både en økonomisk, miljømessig og sosial/fordelingsmessig dimensjon og handler blant annet om stabiliteten i behov og prioriteringer, fremtidig ressurstilgang, fremtidig finansieringsevne og -vilje, fleksibilitet til å tilpasse seg nye rammebetingelser etc. (se Concept-rapport nr. 29, Lædre m.fl., 2012).

Ifølge Samset og Christensen (2012) er det de strategiske kriteriene, og da særlig relevans og levedyktighet, man er notorisk dårlig på å vurdere i forkant. Samtidig er det disse forholdene som er viktigst å få avklart på et tidlig tidspunkt før endelig valg av prosjekt. Ved evaluering i etterkant er det følgelig ekstra viktig å få avklart i hvilken grad dette ble gjort i utgangspunktet, og om investeringen i ettertid fremdeles kan sies å være relevant og levedyktig.

En grunnleggende utfordring med denne modellen er at evaluator nødvendigvis har stor frihet til å definere hvordan de fem kriteriene skal operasjonaliseres, det vil si hvilke problemstillinger som skal belyses, hvem sine perspektiv som skal tas inn og hvordan informasjonen og konklusjonene til sist skal vektas. Evaluator blir på denne måten en «overdommer» i noe som er viktige samfunnsspørsmål og det vil kunne stilles spørsmål ved legitimiteten i vurderingene. Det gjelder kanskje særlig for de strategiske spørsmålene. Dette tilsier at habilitet og erfaring med slik evaluering er grunnleggende viktig, i tillegg til at en må ha et høyt fokus på informasjonens validitet, at alle parter blir hørt, at det metodiske opplegget er konsistent og robust etc.

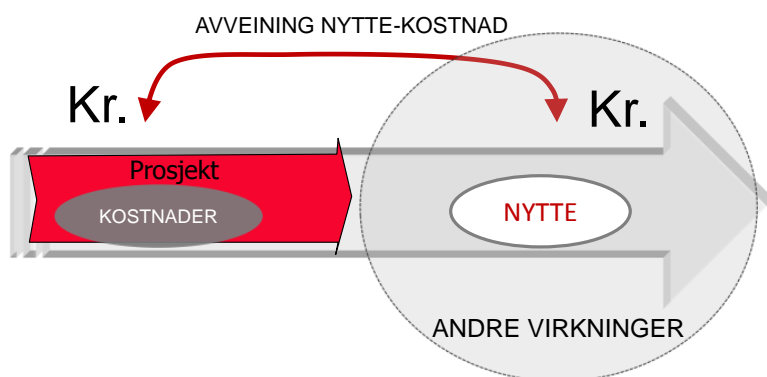
### **Samfunnsøkonomisk lønnsomhet**

Samfunnet har ikke ubegrenset med ressurser, og intet mål er så viktig at det må nås *for enhver pris*. Midler som brukes på ett offentlig tiltak går på bekostning av andre goder innbyggerne kunne ha fått. Vi kan derfor ikke bruke måloppnåelse eller nytte alene som mål på prosjektets vellykkethet, men må forsikre oss om at gevinstene er fremskaffet effektivt og uten sløsing.

Effektivitet handler, enkelt sagt, om sammenhengen mellom den ressursbruken som settes inn og det som kommer ut. Spørsmålet kan stilles begge veier; hva koster det å frembringe en gitt leveranse – og hvor stor leveranse får vi ut av et gitt budsjett? Effektivitet er således en relativ størrelse som i prinsippet er lett å sammenlikne mellom ”like” prosjekter.

Samfunnsøkonomisk analyse er en metode for å systematisk klarlegge og synliggjøre konsekvensene av alternative offentlige tiltak før beslutninger fattes. Metoden er mye brukt av statlige myndigheter ved prioritering av hvilke tiltak som skal få finansiering med knappe offentlige budsjettmidler. Grunnlaget for dagens norske praksis finnes i NOU 1997:27 og Finansdepartementet (2005). Direktoratet for økonomistyring har også utarbeidet en praktisk håndbok for samfunnsøkonomiske analyser (DFØ, 2010). Det er helt nylig kommet en ny utredning fra et ekspertutvalg nedsatt for å vurdere metoden, NOU 2012:16, som må ventes å få innvirkning på praksis i fremtiden.

Alle virkninger skal i prinsippet inkluderes, både på kort og lang sikt, for alle grupper etc. En er ikke opptatt av formelt avtalte mål i prosjektet, og skiller derfor ikke eksplisitt mellom måloppnåelse og andre virkninger. Virkningene kvantifiseres og verdsettes så langt det er mulig og faglig forsvarlig. Nytte måles som summen av befolkningens betalingsvillighet for de godene som fremskaffes, mens kostnader verdsettes til alternativkostnad. Virkninger på ulike tidspunkt sammenstilles ved bruk av en diskonteringsrente. Hensikten med analysen er å svare på om prosjektet er *samfunnsøkonomisk lønnsomt* eller ikke, og en kan rangere prosjekter etter hvilke som er mest lønnsomme.



**Figur 2.4** Avveining av nytte og kostnader i samfunnsøkonomisk analyse

En forsøker altså så langt det er mulig og faglig forsvarlig, å måle alle virkninger i penger. Det er hensiktsmessig fordi det gir en felles målestokk, og ved å benytte betalingsvillighet unngår man i prinsippet skjønsmessige vurderinger. I virkeligheten vil det alltid være en rekke virkninger som ikke kan verdsettes i penger. Finansdepartementet (2005a) skiller mellom tre praktisk anvendbare former for samfunnsøkonomiske analyser, avhengig av muligheten for å verdsette virkningene;

- Nytte-kostnadsanalyse. Dette er den mest fullstendige formen hvor alle sentrale virkninger er verdsatt i penger, slik at den samfunnsøkonomiske lønnsomheten (netto nåverdi) kan utregnes.
- Kostnadseffektivitets-analyse. Dersom nyttesiden kan tas for gitt kan det gjøres en analyse av kostnadene med sikte på å identifisere det mest kostnadseffektive tiltaket for å nå en gitt målsetting.
- Kostnads-virknings-analyse. Der nyttesiden verken kan verdsettes eller tas for gitt kan en ikke uten videre velge det billigste tiltaket. Kostnadssiden verdsettes og nytten gis en kvalitativ beskrivelse. Dette er således en mer fleksibel analyseform, som kan fange et bredere sett av virkninger, disse kan vurderes langs en +/- skala og oppsummeres i en tabell. Kostnads-virkningsanalysen gir ikke noe entydig svar på om prosjektet er lønnsomt eller ikke, dette må vurderes skjønnsmessig til slutt.

Samfunnsøkonomiske analyser er utbredt som del av beslutningsunderlaget for store investeringsprosjekter i staten<sup>10</sup>. Dette er i seg selv et viktig argument for å benytte dette rammeverket også til ex post evaluering. Dersom det foreligger en samfunnsøkonomisk analyse ex ante så vil det i prinsippet være en relativt enkel sak å kontrollere om resultatet ex post ble slik man forventet.

Metoden er imidlertid også mye kritisert, det gjelder både det nytteteoretiske grunnlaget og den praksis man typisk ser ved gjennomføringen av analyser<sup>11</sup>. Den viktigste kritikken kan oppsummeres slik:

- Et grunnleggende problem er bruken av *betalingsvillighet som mål på nytte*. En slik tilnærming har problemer med å fange opp de av samfunnets verdier og prioriteringer som *ikke* reflekteres i individenes betalingsvillighet. Betalingsvillighet har nær sammenheng med betalings*evne* og betyr at ressurssterke grupper tillegges systematisk større vekt enn ressursvake. Fordelingsvirkninger for ulike grupper skal i prinsippet også beskrives, men i praksis gjøres dette i liten grad (Lædre m.fl., 2012).
- Et annet grunnleggende problem som mange har vært opptatt av, er at bruken av en *diskonteringsrente* skaper utfordringer i forhold til å vektlegge virkninger for senere generasjoner. Også dette er fordelingsvirkninger som bør beskrives som del av analysen, men det gjøres svært sjelden.

---

<sup>10</sup> Det er krav om dette etter Utredningsinstruksen (Fornyings-, administrasjons- og kirkedepartementet, 2005) samt krav i KS1 (Finansdepartementet, 2010).

<sup>11</sup> Se blant annet Heinzerling og Ackerman (2002), Næss (2004a) og Nyborg (2012).

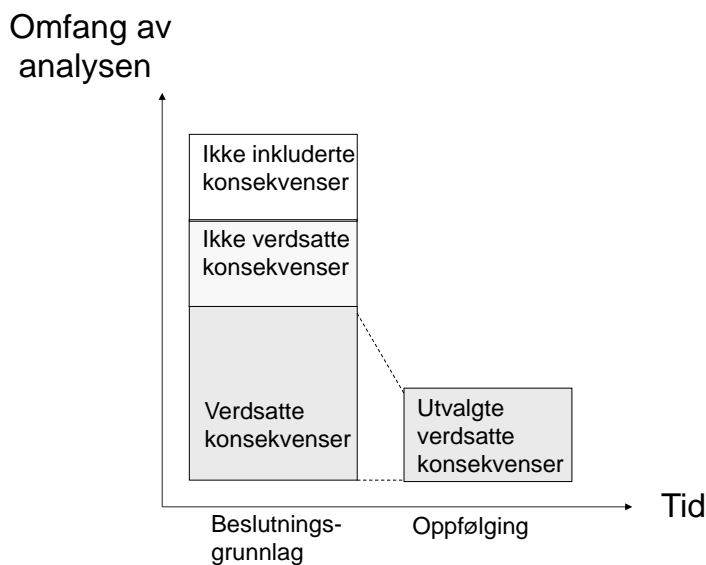
- I praksis er det en vanlig utfordring at *mange virkninger ikke kan verdsettes*. Dersom en kan argumentere for at disse virkningene er små og indirekte, gjør en ikke nødvendigvis noen stor feil ved å se bort fra dem. Men i andre tilfeller handler det om selve måloppnåelsen. For eksempel er det svært vanskelig å transformere nytten av et forsvarsprosjekt til kroner.
- Til slutt er det et betydelig problem at man ofte gjør bruk av unødige tunge modeller, som presenteres på en *lite transparent* form, slik at resultatene blir vanskelige å forstå for både beslutningstakere og folk flest.<sup>12</sup> Dette gjelder ikke minst i nyttekostnadsanalyser av samferdselsprosjekter.

En konklusjon kan være at samfunnsøkonomisk lønnsomhet er en bedre oppsummering av tiltakets vellykkethet desto flere virkninger som kan verdsettes, desto bedre samsvar det er mellom betalingsvillighet og samfunnets nytte, og mer generelt desto mindre de etiske og fordelingsmessige problemstillingene er i tid og rom. I tillegg må resultatene alltid presenteres på en forståelig og oversiktlig måte.

Men også i tilfeller hvor samfunnsøkonomisk lønnsomhet *ikke* alene er en god oppsummering av tiltakets vellykkethet, kan det være interessant å fremskaffe denne informasjonen. Nyborg og Spangen (1996) fant at medlemmene av Stortingets samferdselskomité ikke la særlig vekt på veiprosjektenes nytte-kostnadsbrøk når de skulle prioritere mellom prosjektene, fordi distriktspolitiske hensyn (dvs. fordelingshensyn) ble ansett som vel så viktige. De fleste mente imidlertid at det var interessant å ha tilgang til informasjon om nytte-kostnadsforholdet, slik at de kunne vurdere dette opp mot andre indikatorer på prosjektenes vellykkethet.

---

<sup>12</sup> Grønn, 2003, s 272: ”Når alt kommer til alt, dreier ikke dette seg om mer enn en systematisk anvendelse av sunn fornuft. Det er en beklagelig kjensgjerning at det ofte synes unødvendig med relativt vidløftige teoretiske studier (...) for å holde fast ved alminnelig sunn fornuft”.

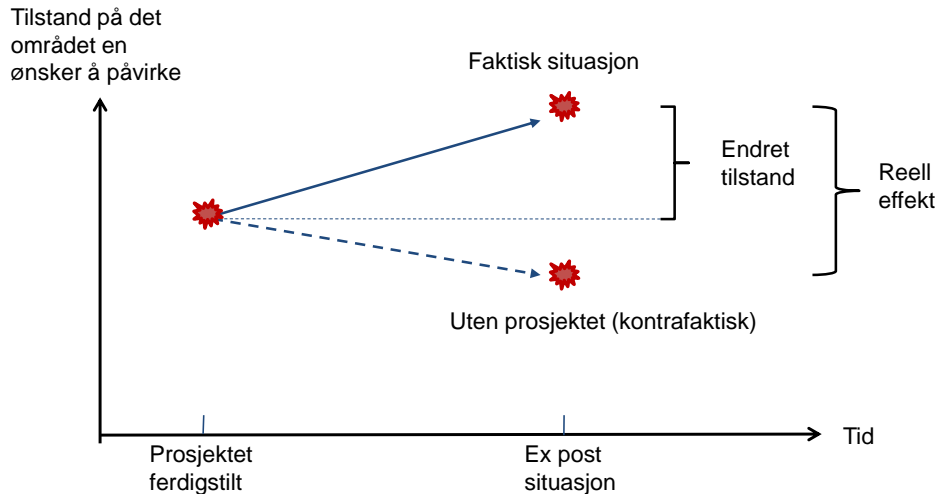


**Figur 2.5** Samfunnsøkonomisk analyse, hva som typisk omfattes før og etter (Olsson og Veiseth, 2011)

Figur 2.5 illustrerer hva som ofte er praksis hva gjelder samfunnsøkonomiske analyser ex ante og ex post. Analysen vil typisk bestå både av verdsatte og ikke-verdsatte virkninger. Samtidig er det nesten alltid en del virkninger som har tendens til å falle ut ("ikke-inkluderte konsekvenser"). I den grad det overhodet gjennomføres samfunnsøkonomiske analyser ex post, er tendensen i enda større grad at det som er vanskelig å måle faller ut. Det kan til og med gjelde virkninger som ex ante ble kvantifisert og verdsatt, for eksempel basert på modellsimuleringer. I ettertid klarer man ikke å samle inn god informasjon om prosjektenes effekter med en realistisk arbeidsinnsats. Dette er viktig å ta med seg før en legger ambisjonsnivået for høyt.

Sammenlikningsgrunnlaget for en samfunnsøkonomisk analyse ex post, er hva som ville skjedd i fravær av tiltaket, dvs. et kontrafaktisk alternativ må være definert. I ex ante analysen har en gjerne definert et såkalt nullalternativ (ingen investering, kun nødvendig vedlikehold) som utgangspunkt for å vurdere lønnsomhet av ulike utbyggingsalternativer, se figur 2.6. Dette nullalternativet kan i prinsippet også legges til grunn som referanse for å vurdere den reelle effekten av prosjektet i ex post-analysen. Det kan imidlertid være vanskelig å vurdere hva som mest sannsynlig ville skjedd uten tiltaket. En må være oppmerksom på at effekten av utbyggingen fremstår som bedre desto dårligere en definerer det kontrafaktiske alternativet, og således er det åpenbart en fare for strategisk påvirkning av resultatene. I praksis ser en ofte at det er tilstandsendringen før-etter som legges til grunn for å beregne samfunnsøkonomiske effekter. Også ved bruk av målorientert

evaluering er det tilstandsendringen før-etter som normalt er utgangspunktet for vurdering av effekten (måloppnåelse).



**Figur 2.6** Før-etter-måling av effekt, versus måling i forhold til nullalternativet. Basert på DFØ (2007)

I ex post-analysen vil det i utgangspunktet være naturlig å bruke siste år som referansetidspunkt (som alle andre verdier diskonteres tilbake eller frem til) fremfor det året som ble brukt i ex ante-analysen. Men dersom det er et viktig poeng å sammenlikne med de verdier som ble estimert ex ante kan en også bruke det gamle tidspunktet, resultatet bør bli det samme.

## Konkurrerende eller komplementære modeller?

Målorientert evaluering og samfunnsøkonomisk analyse kan betraktes som dels konkurrerende og dels komplementære modeller.

Modellen for målorientert evaluering er generisk og fleksibel, og gir ikke konkrete føringer om valg av indikatorer, datatyper eller beregninger. Det er i prinsippet ikke noe i veien for å inkludere samfunnsøkonomisk lønnsomhet som indikator i modellen (for eksempel under det overordnede kriteriet ”måloppnåelse” eller evt. ”virkninger”). På den annen side er det heller ingen automatikk i at en i målorientert evaluering spør om prosjektets samfunnsøkonomiske lønnsomhet. Den samfunnsøkonomiske analysen sammenstiller informasjonen på en helt bestemt måte, for å forsøke å svare på om den samlede betalingsvilligheten for et prosjekt overstiger kostnadene. Dette er et interessant mål på vellykkethet som en dessuten krever at beregnes ex ante.



Også samfunnsøkonomisk analyse betraktes av enkelte som et overordnet kriterium for vellykkethet. Dersom et prosjekt er samfunnsøkonomisk lønnsomt kan en argumentere med at det også vil være kostnadseffektivt, ha positive nettovirkninger og være relevant og levedyktig, altså dekke alle kriterier i modellen for målorientert evaluering. I praksis er analysen ofte mer snever i tid og rom, og spesielt dekker den ofte ikke det strategiske nivået på en tilfredsstillende måte. Et annet moment er at den samfunnsøkonomiske analysen mangler det eksplisitte målstyringsfokus som ofte er viktig ved store statlige investeringsprosjekter.

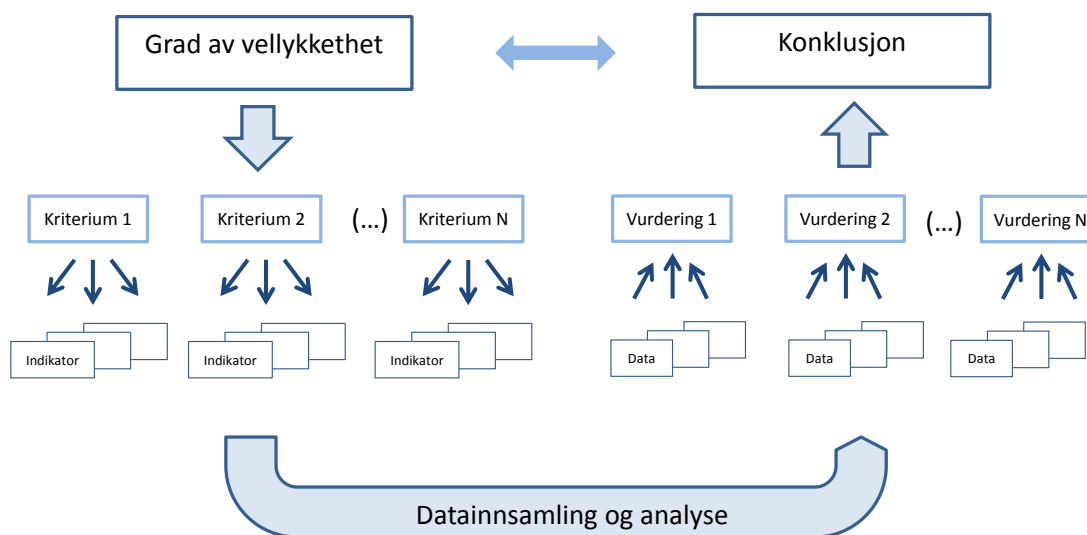
Alt i alt mener vi at hver av modellene dekker aspekter som ikke dekkes av den andre. Det synes derfor fornuftig å inkludere både de fem evalueringskriteriene i målorientert evaluering og i tillegg samfunnsøkonomisk lønnsomhet som et sjette evalueringskriterium. Det meste av databehovet til en samfunnsøkonomisk analyse, i form av fysiske effekter og virkninger, vil uansett være felles med det ettersøker i målorientert evaluering, og således er den ekstra ressursbruken for evaluator liten.<sup>13</sup>

## 2.3. Datainnsamling og analyse

Anvendelsen av en slik utvidet modell for å evaluere prosjekters vellykkethet forutsetter en prosess med systematisk disaggregering og aggregering av informasjon på overordnet og underordnet nivå. Prinsippet er at en med utgangspunkt i overordnede evalueringskriterier velger ut et sett med mer konkrete forhold som skal undersøkes, såkalte indikatorer, som til sammen kan gi et svar på i hvilken grad en har lyktes på hvert av områdene. Med andre ord, man bryter ned de overordnede spørsmålene i mer konkrete spørsmål knyttet til det som skal evalueres, samler informasjon som kan gi svar på dette og får på den måten et grunnlag for å gi svar på de overordnede spørsmålene. Dette er illustrert i figur 2.7.

---

<sup>13</sup> Kostnader finnes under ”produktivitet”, sentrale nyttevirksomheter er beskrevet under ”måloppnåelse”, og ”virkninger” dekker ulike sidevirkninger som bør kunne splittes i positive (nytte) og negative (kostnad). Samfunnsøkonomisk analyse spør ikke om målene er ”relevante” – i stedet fokuserer man på direkte verdsetting av de effektene og virkningene prosjektet genererer. ”Levedyktighet” fanges i den grad en greier å inkludere alle nytte- og kostnadsstrømmer på så lang sikt som prosjektet kan ventes å generere effekter.



**Figur 2.7** Planlegging og gjennomføring av en prosjektevaluering

Prosedyren blir altså som følger:

1. Fastsett det/de overordnede kriteriene for vellykkethet som prosjektets resultater skal måles mot. Her tar en utgangspunkt i modellene for målorientert evaluering og samfunnsøkonomisk lønnsomhet som drøftet over.
2. Disaggreger de overordnede evalueringskriteriene til mer konkrete evaluerings spørsmål og indikatorer, som er særlig relevante for evalueringsobjektet
3. Innhent og analyser data for å besvare evaluerings spørsmålene
4. Aggreger resultatene og trekk en overordnet konklusjon for hvert evalueringskriterium og for prosjektet samlet

Valg av konkrete indikatorer, datakilder etc. vil selvsagt være forskjellig fra prosjekt til prosjekt. I denne delen av kapitlet diskuteres hvilke forhold som er viktige når en skal identifisere konkrete indikatorer, og samle inn og analysere data for å vurdere hvert av de overordnede evalueringskriteriene omtalt i forrige avsnitt, det vil si *produktivitet, måloppnåelse, virkninger, relevans, levedyktighet og samfunnsøkonomisk lønnsomhet*.

Indikatorerne kan være alt fra tallmål til en beskrivelse av ulike informanternes meninger og holdninger. Til sammen skal et sett av indikatorer per overordnet

kriterium gi evaluator et tilstrekkelig grunnlag for å konkludere om graden av vellykkethet for det aktuelle kriteriet.

Et viktig bakteppe er at mange evalueringer skjer under tidspress og med begrenset budsjett. Dersom en er for ambisiøs, eller betrakter evaluering som synonymt med forskning, vil resultatet kanskje ikke stå i forhold til innsatsen. I evaluering må en fokusere på hva som er ”godt nok” innenfor realistiske rammer, men likevel sette akseptable kvalitetskrav. Dette er drøftet i Concept-rapport nr. 17 (Sunnevåg, 2007). Ikke minst kommer spørsmålet om hvilken informasjon som finnes tilgjengelig, hvilken informasjon det er regningsvarende å fremskaffe, og hvilket presisjonsnivå en skal legge opp til. Ofte vil en måtte kompromisere på alle disse områdene, men uten at dette i vesentlig grad går ut over kvaliteten eller troverdigheten av det en kommer frem til. Erfaring tilsier at en kan redusere på detaljeringsgrad og presisjon, som normalt er sterkt kostnadsdrivende, uten at en får tilsvarende uttelling når det gjelder nytten. Men det som er avgjørende for kvaliteten er at en sikrer informasjonens *validitet* og *reliabilitet* - og det er ikke nødvendigvis fordyrende.

## Informasjonens kvalitet. Validitet og reliabilitet

Evalueringer har mye til felles med forskning og utredningsarbeid. I praksis er en imidlertid ofte nødt til å senke både ambisjons- og presisjonsnivået fordi pålitelig informasjon er vanskelig tilgjengelig eller at en ikke har tid eller ressurser til å skaffe slik informasjon. Uansett skal det stilles krav til å opprettholde en akseptabel faglig standard, hvilket betyr at et visst krav til ”vitenskapelighet” også må tillegges betydning ved evaluering:

- Dette gjelder for det første kravet til etterprøvbarehet (*reliabilitet*) som innebærer at informasjonen skal være så pålitelig at en uavhengig undersøkelse av samme fenomen ville kunne gi tilnærmet identisk resultat. Det stilles derfor krav til at evalueringen framskaffer god nok informasjon, men også at det er redegjort for de metodene som brukes slik at det kan være mulig for utenforstående å etterprøve resultatet.
- For det andre må det stilles krav til informasjonens gyldighet (*validitet*). Det innebærer at informasjonen må være relevant for nettopp den problemstillingen som skal belyses. Mens etterprøvbarehet kan kontrolleres ved å foreta omtest, er det slik at gyldigheten ikke lar seg fastslå direkte empirisk. Her må det brukes skjønn. Validitet sikres først og fremst ved at man velger variable som gir et mest mulig direkte mål på det som skal undersøkes. I tillegg at man bruker flere variable eller informasjonskilder for å beskrive samme fenomen, slik at en både får en mer fullstendig

beskrivelse, men også en trianguleringseffekt som bedrer gyldigheten av det en har funnet.

- For det tredje må det stilles krav til at de antakelsene som gjøres og konklusjonene som trekkes i evalueringen har støtte i faktaopplysninger eller konkret erfaring fra virkeligheten. Den som evaluerer plikter å oppgi kilder i de tilfellene der andre har dokumentert de sammenhengene som beskrives.
- For det fjerde må det stilles krav til bredde og fordomsfrihet. Det betyr at beskrivelser og konklusjoner bør ha gyldighet utover en enkelt situasjon eller institusjon. En må derfor legge vekt på forskjellige sider ved et problemkompleks og være åpen for å teste ut ulike oppfatninger, hypoteser eller forklaringer.

Som nevnt ovenfor må en i mange tilfeller redusere ambisjonsnivået som følge av de forholdene en arbeider under. I enkelte tilfeller kan for eksempel kravene til gyldighet og etterprøvarhet komme i konflikt. Kombinasjonen av faglig skjønn, støtte i andre undersøkelser og egne data kan til sammen ofte gi et langt bedre bilde enn konvensjonelle oppfatninger. Men samtidig er det viktig å angi hvor langt en kan trekke konklusjonene. Dette gjøres best gjennom ryddige utsagn om hvilken del av virkeligheten en har faktaopplysninger om, og hvilke forhold en uttaler seg om.

## **Kvantitativ eller kvalitativ analyse. Ja takk, begge deler**

Et sentralt spørsmål i evalueringer er hvor avanserte analyser man skal legge opp til. Tabellariske oversikter og histogrammer, prosenttall og gjennomsnittstall representerer et basisnivå som også gjør det lett å kommunisere med ikke-spesialister. Informasjon om fordelinger, standardavvik og samsvar gir ofte nyttig og nødvendig informasjon. I tillegg kan man tenke seg et utall av avanserte analysemetoder som kan tas i bruk.

Imidlertid er det både mulig og ønskelig å i størst mulig grad basere evaluering på enkle og grunnleggende analyser. Tross alt, både utvalgsstørrelser og kvaliteten på data er ofte slik at de ikke egner seg for avansert analyse. Dessuten vil en måtte benytte både kvantitativ og kvalitativ analyse.

*Kvantitativ* analyse er tallbasert. Det kan være objektiv faktainformasjon eller mer subjektive vurderingsdata. Kvantitativ informasjon vil typisk være mengder og frekvenser, resultater av modellberegninger, parametriske statistikk etc., som brukes til å belyse sammenhenger mellom variabler og til å generalisere funn. For å skaffe slik informasjon er evaluator opptatt av å følge klart definerte prosedyrer, for eksempel tilstrebe at et utvalg er statistisk representativt. Også data som i

utgangspunktet er kvalitativ, for eksempel resultater fra menings- og holdningsundersøkelser, kan ofte kvantifiseres og uttrykkes på en skala.

Den kvantitative analysen har i prinsippet tre hovedformål:

- Å beskrive fenomener i et eksakt format ved hjelp av statistikk
- Å teste samsvar mellom utvalgte variable
- Å generalisere funn fra et utvalg til en større populasjon.

En fordel med kvantitativ informasjon er at en kan generere og analysere data om mange undersøkelsesenheter. En fordel med statistisk analyse er at en kan sammenfatte forskjellige funn på en klar, presis og pålitelig måte. Men det er mange typer informasjon som ikke kan analyseres kvantitativt. For eksempel informasjon fra intervjuer eller spørreundersøkelser med åpne spørsmål, hvor en mister mye av informasjonen om en forsøker å kategorisere denne. Et sentralt problem er at resultatet vil være avhengig av at dataene har en viss kvalitet. Hvis ikke svekkes gyldigheten av konklusjonene uten at dette nødvendigvis synliggjøres. Det er lett å gå i den fellen at en får funnene til å virke mer objektive og av høyere kvalitet enn de i virkeligheten er. Kvaliteten på analysene avhenger alltid av kvaliteten på de dataene som inngår inn, og en ser ofte eksempler på at analysene blir mer sofistikerte enn dataene skulle tilsi.

Oversikt over viktige datainnsamlings- og analysemetoder:<sup>14</sup>

#### Kvantitative

- Registerbaserte undersøkelser (for eksempel SSB)
- Egen spørreskjemaundersøkelse/formell survey med statistisk representativt utvalg
- Andre egne målinger/registreringer
- Etterprøving av en nytte-kostnadsanalyse

#### Kvalitative

- Dokumentanalyse (prosjektdokumenter, tidligere evalueringer etc.)
- Direkte observasjon (for eksempel befarings, teste bruken av en tjeneste)
- Case-studie
- Personlig intervju med nøkkelinformanter (ustrukturert eller semistrukturert), enkeltvis eller fokusgruppeintervju

<sup>14</sup> Se for eksempel Finansdepartementets veileder i gjennomføring av evaluering (2005b) for en nærmere omtale av den enkelte metode

*Kvalitativ analyse* bygger på tekstlig informasjon i form av beskrivelser, observasjoner, direkte sitater fra åpne spørreundersøkelser, intervjuer etc. Kvalitativ analyse omfatter innholdsanalyser, analyse av casestudier, induktiv analyse og logisk og sekvensiell analyse. Resultatet er beskrivelser (mønstre, tematikk, tendenser, trender etc.) og fortolkninger og forklaringer av dette. Kvalitativ analyse bidrar til å gi bredde og dybde i det som evalueres og er helt nødvendig som supplement til kvantitative funn. Imidlertid vil data basert på ustrukturerte intervjuer, egne observasjoner og tolkninger, i mindre grad være etterprøvbare, og i en viss grad påvirket av subjektiv vurdering og skjev fremstilling. Kvaliteten på analysen hviler i stor grad på datautvelgelse, perspektiv og metoder, samt på evaluators troverdighet mer generelt. Det blir derfor et ufravikelig krav å gjøre rede for gyldigheten (validitet) og etterprøvbarheten (reliabilitet) av det en legger til grunn og det en kommer frem til.

En vil måtte bruke både kvantitativ og kvalitativ analyse ved evaluering basert på de fem evalueringskriteriene. Det kan være nyttig å starte med en åpen og i stor grad kvalitativ tilnærming, hvor en gjennom intervjuer, observasjon og lesing av bakgrunnsdokumenter, forsøker å skaffe oversikt over relevante tema, virkninger, tilgjengelige data etc. Deretter benyttes i større grad kvantitative metoder for å besvare mer konkrete hypoteser. En slik kombinasjon innebærer dessuten en form for metodetriangulering som er viktig for evalueringens troverdighet og kvalitet, ikke minst i tilfeller hvor konklusjonene kan være kontroversielle.

### **Hva er "gode data"? Kvalitet og nytte av evalueringen**

Evalueringer kjennetegnes blant annet ved at de gjennomføres i løpet av kort tid, under vanskelige forhold og at de tar opp vesentlige spørsmål til kritisk vurdering. Evaluering kjennetegnes også ved at en i større grad enn forskning bruker enkle, mindre ressurskrevende metoder i tillegg til formelle metoder. Konklusjoner og anbefalinger vil ofte bli oppfattet som kontroversielle, og det er derfor vanlig at også selve evalueringen blir gjenstand for kritisk vurdering. I slike tilfeller reises det ofte spørsmål som gjelder valg av metoder for datainnsamling og analyse.

*Formelle metoder* kjennetegnes ved at prosedyrene i utgangspunktet er klart definert. Resultatet vil normalt være kvalitativ og kvantitativ informasjon med høy presisjon, det vil si med høy validitet og reliabilitet. Problemet med formelle metoder er først og fremst at de er tid- og ressurskrevende. Eksempler kan være formell spørreundersøkelse med statistisk representativt utvalg av respondenter, eller systematisk dybdeobservasjon over lengre tid. De brukes i noen grad i forbindelse med forundersøkelser til evaluering, men sjelden som en del av selve evalueringen.

Motsatsen til formelle metoder er de *enkle, helt uformelle metodene* uten klart etablerte prosedyrer å holde seg til. Slike metoder benyttes i stor grad ved evaluering, ikke minst har det vist seg at samtaler med nøkkelpersoner er den mest benyttede metoden for datainnsamling. Fordelene med uformelle metoder er at de gjerne krever lite tid og ressurser og ofte resulterer i informasjon som mer eller mindre er tilrettelagt for bruk. Problemet er først og fremst usikkerheten med hensyn til kvaliteten. Personlige oppfatninger kan påvirke både validitet og reliabilitet. Ofte er det slik at den som gir informasjon systematisk unngår observasjoner som ikke understøtter en bestemt konklusjon.

Det er derfor behov for mer *tilpassede metoder* som ligger mellom de formelle og de helt uformelle metodene. Dette gjelder uformelle surveys med små utvalg som ikke er statistisk representative, men som bruker mindre kvoteutvalg for eksempel på grunnlag av holdninger, rolle, status eller lignende. Det kan også gjelde gruppeintervju der tilstedeværelsen og medvirkningen av flere respondenter gir triangulerende effekter og viktig tilleggsmåling. Og det kan være fokusgruppeprosesser der en gjennomfører styrt dialog i en gruppe nøkkelpersoner omkring bestemte temaer.

Slike forenklete metoder vil ofte være bedre egnet som grunnlag for forståelse og analyse enn for presis måling. Men utgangspunktet for bruken vil ofte være kvantitativ statistikk og faktainformasjon.

Bruk av tilpassede eller helt uformelle metoder krever at en tar hensyn til hvilke kvalitetsmessige konsekvenser dette vil kunne få, særlig med hensyn til validitet og reliabilitet, og at en i rapporten redegjør for metodevalg og de begrensningene som ligger i data og analyse.

#### Ressurskrevende - Høy validitet/reliabilitet

—————	Formell survey
—————	Dybde-observasjon
—————	Case-studier
—————	Uformell survey
—————	Fokusgruppe
—————	Evaluators egen observasjon
—————	Direkte måling
—————	Intervju med nøkkelinformanter
—————	Litteratursøk, sekundærdata

#### Lite ressurskrevende - Lavere validitet/reliabilitet

**Figur 2.8** Datainnsamlingsmetoder - ressursbruk og validitet, basert på Samset (2003)

Konklusjonen må være at en veier kostnad mot kvalitet når en velger datainnsamlingsmetoder. Det betyr normalt at enkle og rimelige metoder alltid bør benyttes, mens mer ressurskrevende metoder brukes mer selektivt.

Triangulering eller kryssvalidering betyr å studere et fenomen fra ulike vinkler eller med ulike tilnærminger, og på denne måten oppnå en mer helhetlig forståelse. Begrepet kommer fra landmåling og navigasjon. Innen evaluering handler triangulering om å benytte ulike teorier og metoder, data, forskere etc. Hensikten er både å få et mer komplett datagrunnlag og å kunne trekke mer robuste konklusjoner. Triangulering er kritisk viktig for å sikre evalueringsresultatenes legitimitet og troverdighet. Dette er spesielt viktig dersom evalueringen er kontroversiell, dersom det er motstridende interesser, og dersom evaluator har hatt begrenset tilgang til informasjon av høy kvalitet.



---

## 3. Evaluering av fire prosjekter. Sammenfatning

I dette kapitlet presenteres de fire pilotevalueringene som ble gjennomført som del av denne studien. Det gjelder fire investeringsprosjekter som er avsluttet for ca. 5 år siden, med unntak av MTB Skjold hvor det første serieproduserte fartøyet ble levert for 6 år siden mens det siste ennå ikke er levert. Evalueringen av MTB Skjold er således ikke en reell *ex post* evaluering, men en lykkes likevel i å gi noen foreløpige vurderinger av prosjektets vellykkethet.

Prosjektene er helt forskjellige og representerer ulike sektorer. Ikke minst derfor var det interessant å teste muligheten for å benytte samme evalueringsmodell på alle, dvs. de fem overordnede evalueringskriteriene fra målorientert evaluering samt kriteriet samfunnsøkonomisk lønnsomhet.

Ingen av prosjektene har gjennomgått KS1, og derfor har det vært utfordrende å finne et referansegrunnlag for de overordnede vurderingene av konseptvalget. Evaluator har i flere tilfeller selv måttet stipulere hva en anså som realistiske målformuleringer og prosjektlogikk fra bakgrunnsdokumentene, ettersom målformuleringer på høyere nivå manglet eller var urealistiske. Dette vil en forhåpentligvis slippe i fremtiden, når de fleste store investeringsprosjektene har gjennomgått ekstern kvalitetssikring av konseptvalget før beslutningen ble fattet.

Nedenfor gis kun en oppsummering av de viktigste funnene i evalueringene. Hvert team ble bedt om å gi en score for det enkelte overordnede evalueringskriterium på en skala 1 – 6 som grunnlag for en overordnet konklusjon av prosjektets vellykkethet.

### 3.1. Svinesund kontrollområde

#### Kort introduksjon til prosjektet

Prosjektet omfatter byggingen av et nytt felles kontrollområde for Tollvesenet, Statens vegvesen og Politiet, ved landets mest trafikkerte grenseovergang mot Sverige. Byggeprosjektet ble gjennomført av Statsbygg i perioden 2002-05. Leveransen omfattet bygningsarealer på nær 8 000 kvm, hvorav 75 % til Tollvesenet. Det ble også satt av lokaler til spedisjonsfirmaer på stedet.



Prosjektet kom opp som et følgeprosjekt av den nye Svinesundforbindelsen, som bestod av ny bro ved Svinesund og oppgradering av E6 til firefelts motorveg langs ny trasé.

Det nye kontrollområdet avløste tidligere separate kontrollstasjoner for Tollvesenet og Vegvesenet langs den gamle E6-traséen (politiet hadde ingen fast stasjonering). Et felles grensekontrollpunkt for personer, gods og kjøretøy i denne skala skal være unikt i internasjonal sammenheng.

Både vegen, broen og kontrollområdet ble tatt i bruk i juni 2005, i forbindelse med 100-årsjubileet for oppløsningen av Norges union med Sverige. Den norske og den svenske kongen var til stede ved åpningen.

### **Bakgrunn og tidligfase**

Behovet for ny Svinesundforbindelse kom som en følge av sterk trafikkøkning over grensen gjennom mange år. Planene om sammenhengende firefelts motorveg mellom Göteborg og Oslo verserte siden 1980-tallet, og grensepasseringen på Svinesund ble sett på som en av flaskehalsene. Det lå i kortene at dersom vegens

trasé ble flyttet, måtte kontrollstasjonene flytte etter, men det var lenge uklart om og når dette ville skje.

Ideen om samlokalisering av kontrolletatene ble nevnt første gang i 1998, i forbindelse med utredning av den nye Svinesundbroen og veiparsellen Nordby (Sverige) – Svingenskogen. Når en først skulle flytte, var det også klart at den nye stasjonen måtte bli større enn summen av de gamle, da i alle fall Tollvesenet ville måtte øke sin grenseklaringsaktivitet betraktelig på grunn av trafikkøkningen. En større stasjon ville også gi mulighet for økt kontrollvirksomhet.

Statsbygg fikk i juni 2002 oppdraget fra Finansdepartementet, Samferdselsdepartementet og Politimesteren i Østfold å utarbeide forprosjekt om felles kontrollområde langs den nye veien. Deretter gikk det hele meget raskt. Ingen andre konsepter synes å ha blitt utredet.

### **Evalueringen**

Avgrensningen av evalueringsobjektet blir diskutert i evalueringen. Prosjektet nytt kontrollområde kom som en nødvendig følge av ny Svinesundforbindelse. En kunne da tenke seg at man måtte se på investeringen i kontrollområdet som en integrert del av hele samferdselsinvesteringen. Dersom vegprosjektet ikke hadde endret trasé, er det ikke sikkert at det hadde vært nødvendig å investere i ny(e) kontrollstasjon(er), i alle fall ikke på det angjeldende tidspunkt.

På den annen side ga dette en mulighet for å velge et annet konsept for kontrollområdet, basert på samlokalisering av tre etater, og en generell mulighet for modernisering og vekst. For å identifisere gevinstene av dette er det mer naturlig å betrakte kontrollområdet som et avgrenset evalueringsobjekt. Evaluator valgte det siste. Det er også naturlig siden Kontrollområdet og Svinesundforbindelsen ble gjennomført som to forskjellige prosjekter. Evaluator har likevel passet på å se kontrollområdet i sammenheng med veg- og broprosjektet for øvrig.

Evaluator har lagt vekt på i størst mulig grad å benytte allerede innsamlet datamateriale, men har også supplert med egne intervjuer. En har lagt vekt på triangulering, og det er samlet både dokumenter, statistiske data, egen befarings samt intervjuer fra en rekke kilder.

### **Prosjektlogikk og målstruktur**

I motsetning til mange andre større statlige investeringer, har dette prosjektet vært lite omdiskutert. Målformuleringer finnes ikke i de stortingsproposisjoner der prosjektet omtales, annet enn at det er et nødvendig tiltak og at det må stå ferdig til bruk samtidig med at Svinesundforbindelsen åpner.

I prosjektets styringsdokument er det formulert mål på alle tre nivåer. Resultatmålet er, i tillegg til tid og kostnad, *Et karakterfullt anlegg*. Effektmålet er *Synergisk samhandling mellom tollmyndighet, politimyndighet og samferdselsmyndighet*. Samfunnsmålet er *En autoritetsskapende og gjestfri grensekontroll på Svinesund*.



Evaluatør tegner opp og drøfter prosjektets logikk, men konkluderer med at den ikke er logisk og realistisk. Resultatmålet er misvisende, effektmålet for lite konkret, deler av samfunnsmålet ("gjestfri") fremstår som motsetningsfullt i forhold til resten ("autoritetsskapende") og logikken mellom nivåene mangler. Evaluatør velger derfor å legge til grunn andre og mer egnede mål fra relevante underlagsdokumenter samt fra målene for de tre etatene. Resultatmålet bør fokusere på at anlegget skal være *funksjonsriktig*. Effektmålet bør handle om 1) *Mer effektiv grensekontroll*, som gjelder etatenes operasjonelle og taktiske effektivitet, og 2) *Økt framkommelighet* for (de lovlydige) trafikantene. Samfunnsmålene bør knyttes til redusert kriminalitet, trafikkulykker etc. – eller *samfunnsøkonomisk nytte* mer generelt.

Denne prosjektlogikken ga grunnlag for å identifisere konkrete indikatorer på måloppnåelse for oppfølging i evalueringen.

### **Produktivitet**

Kvalitetsmålet gjelder som nevnt anleggets funksjonalitet. Dette er bare delvis innfridd, mener evaluatør. Bygningsmassen holder en enkel og kanskje for enkel standard på noen områder, noe som også er bekreftet gjennom intervju med de ansatte. Stasjonen er likevel utformet i samsvar med spesifisering og for å legge til

rette for samhandling i tråd med intensjonen, og nødvendig teknologisk utstyr er på plass.

I dette prosjektet har tidsmålet vært det viktigste, og det var stor prestisje forbundet med å nå målet om ferdigstilling 7. juni 2005. Tidsmålet ble innfridd.

Leveransen skjedde også innenfor den definerte styringsrammen på 275 mill. kroner. Det har vært diskutert om dette skyldes god kostnadsstyring eller om budsjettet rett og slett var for romslig. Det faktum at rammene var kvalitetssikret (KS2) – riktignok først etter at prosjektet var godt i gang – taler for det første. I dette prosjektet testet man dessuten ut en ny prosjektstyringsmodell som skulle styrke den overordnede eierstyringen, noe som også kan ha bidratt til en vellykket gjennomføring.

### ***Måloppnåelse***

Evalueringen har lagt til grunn ulike indikatorer både for Tollvesenets og Vegvesenets kontrollaktivitet, herunder å deklare legal vareimport, avsløre smugling og å avdekke brudd på Vegtrafikkloven. En har også sett på effektiviteten hos spedisjonsfirmaene på kontrollområdet som påvirker den samlede effektiviteten for grensekontrollen.

Det er flere indikasjoner på at resultatene av kontrollaktiviteten er bedret. Blant annet har Tollvesenet hatt en flerdobling i beslag av bestemte typer illegale varer. Statens Vegvesen har muligens hatt en forbedret treffsikkerhet i sine kontroller, men endringer i registreringsrutiner gjør konklusjonen usikker. Evaluator peker på at en viktig årsak, i alle fall hva gjelder Tollvesenet, er det økte aktivitetsnivået (ca. 50 flere ansatte). I tillegg har bruken av IT-systemer, administrative prosedyrer og ny teknologi hatt positiv effekt.

Noe av effektivitetsøkningen skyldes nok også utveksling av informasjon, for eksempel fra Tollvesenet til Vegvesenet om hvilke biler som bør inn til kontroll før de forlater kontrollområdet. Evaluator mener likevel at potensialet for samhandling bare i begrenset grad er tatt ut. En rimelig konklusjon etter syv års erfaring synes å være at det ikke er tilstrekkelig med fysisk samlokalisering, det er også nødvendig at de tre etatene tar felles initiativ til mer samhandling. Hvorvidt en faktisk ønsker at dette skal skje synes noe uklart i dag.

Det andre effektmålet er knyttet til fremkommeligheten. Denne er bedret for alle grupper trafikanter takket være at den nye vegen har fire felt i stedet for to. Det er likevel periodevis store opphopinger og lange forsinkelser for kjøretøy på rødsone, altså de som skal gjennom fortolling. Det medfører tidstap for lastebilnæringen, og omfanget på dette problemet har økt til et nivå som ikke

aksepteres verken i næringen eller på politisk hold. Etatene arbeider med mulige løsninger, fortrinnsvis innenfor eksisterende rammer for kontrollområdet.

### ***Virkninger***

Evaluatør har også sett på andre mulige virkninger som kanskje ikke var tilsiktet. For det første kan omfanget av *smuglingslekkasje* ha økt. Sentralisering av grensekontrollen i et stort kontrollområde kan ha en avskrekkende virkning på smuglere og føre til at noen velger å passere på andre grenseoverganger lenger nord. Tollerne har imidlertid strategier for å avdekke dette, og en har ikke klare indikasjoner på at det skjer i større omfang.

For det andre rapporteres det om at de ukentlige opphopningene av lastebiler på området, kan utgjøre et visst *sikkerhetsproblem* for tjenestemenn som arbeider ute. Dette bør kunne løses med flere parkeringsplasser og eventuelt andre tiltak i samarbeid mellom etatene og lastebilnæringen. Bilister som snur og kjører tilbake til Sverige mot kjøreretningen når de ser at bilene foran dem sluses inn til kontroll, er også observert i et lite, men økende omfang.

En mulig positiv utilsiktet virkning er at Tollvesenet i stor grad bruker den nye kontrollstasjonen som en arena for opplæring. Dersom Tollvesenet har nytte av denne opplæringsarenaen vil det være positivt. Det må imidlertid holdes opp mot den mulig negative effekten det kan ha på kontrolleffektiviteten.

### ***Relevans***

Hovedaktiviteten ved kontrollområdet er deklarerering av legal vareimport. Utgangspunktet er derfor at så lenge Norge skal ha et eget tollregime, så er det relevant å ha en kontrollstasjon på grensen langs landets største innfartsåre fra Sør-Europa. I den grad grensekontrollen ved den gamle Svinesundbroen ble vurdert som relevant, er det all grunn til å tro at det samme gjelder den nye stasjonen. Forutsetningen for at selve flyttetiltaket er relevant, vil derfor være at hovedtyngden av trafikken faktisk flyttes til den nye broen, og dessuten at effektiviteten faktisk forbedres i tråd med intensjonen. Det trekker åpenbart ned at om lag en tredjedel av trafikken fremdeles går over den gamle broen og bare sporadisk utsettes for kontroll.

Det synes ikke som at noen andre konsepter er blitt utredet i forkant, og en kan derfor ikke utelukke at det finnes andre konsepter som er enda mer relevante. Evaluatør peker blant annet på et konsept med felles tollstasjon for Norge og Sverige, enten på svensk eller norsk side, som ble diskutert på et tidlig stadium men aldri utredet nærmere.



### **Levedyktighet**

Kontrollområdets levedyktighet vil åpenbart stå og falle med Tolletatens tilstedeværelse og funksjon. Tolletaten er den dominerende part med mer enn 90 prosent av de fast stasjonerte tjenestemenn. Det eksisterende fortollingsregimet mellom EU og EØS-land kan teoretisk sett opphøre ved et pennestrøk gjennom internasjonale forhandlinger, selv om det ikke er noen indikasjoner i dag på at dette vil skje. En radikal nedbemanning og en mer symbolsk tilstedeværelse på grensen ville bli resultatet. En mer profesjonell transportnæring i landene sør og øst på kontinentet og smidigere spedisjonsbehandling vil også kunne redusere behovet for kapasitet på tollstasjonen betraktelig.

Hva ville i så fall være alternativt anvendelse av kontrollområdet? Vegetaten ville kunne utvide sin kontrollvirksomhet, spesielt av tunge kjøretøyer, forutsatt at dette ble prioritert og at en fikk midler til dette. Det er et utvilsomt behov for slike kontroller i grenseområdet, særlig på vinterstid, på grunn av glatte veier, slitte dekk, manglende kjetting, etc. Det har imidlertid vært fremholdt som en ulempe ved dagens kontrollområde at det for Vegvesenets del helst burde ligge utenfor synsrekkevidde til vegen.

Når det gjelder levedyktigheten til bygningsmassen er avskrivningsperioden fra Statsbyggs side satt til å være 60 år. Dette kan være et rimelig estimat. Den forrige stasjonen ble bygget i 1986 og slik sett hadde den kun en levetid/aktualitet på 20 år, men dette har sin årsak i behovet for ny vegtrasé og bro.

### **Samfunnsøkonomisk lønnsomhet**

Evaluator har gjennomført en egen samfunnsøkonomisk analyse av tiltaket. Det presiseres imidlertid at datagrunnlaget er usikkert og mangelfullt, også når det gjelder nullalternativet. Nyttensiden er derfor diskutert rent kvalitativt, og er langt på vei sammenfallende med vurderingen av måloppnåelse og virkninger over. Konklusjonen er at det er vanskelig å si om den økte nytteverdien er mindre eller større enn endringen i kostnadene, og dermed om tiltaket er samfunnsøkonomisk lønnsomt.

### **Oppsummert om prosjektets vellykkethet**

Tabellen under gir en oversikt over de viktigste (mest valide) indikatorene som ble benyttet for å belyse hvert av evalueringskriteriene.

**Tabell 3.1** Svinesund kontrollområde. Indikatorer

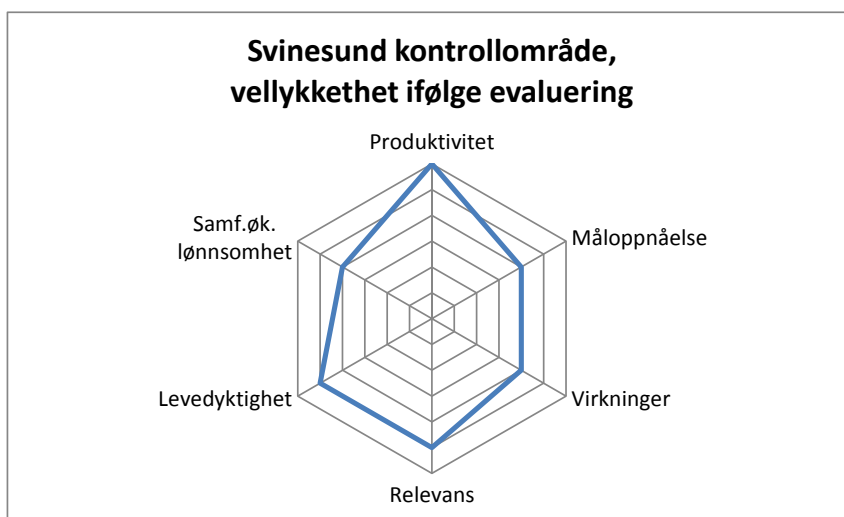
<b>Viktigste indikatorer</b>	
<b>Produktivitet</b>	<ul style="list-style-type: none"> <li>– Anleggets funksjonalitet, vurdert ved intervju og egen observasjon</li> <li>– Tidsmål, vurdert ved åpningstidspunkt mot målet om å være klar for drift innen 07.06.05</li> <li>– Kostnadsmål, vurdert ved prosjektregnskap mot styringsramme og kostnadsramme i KS2</li> <li>– Kostnad per kvm, målt mot gjennomsnitt for andre bygg</li> <li>– Egen analyse av gjennomføringen, inkl. organisering</li> </ul>
<b>Måloppnåelse</b>	<ul style="list-style-type: none"> <li>– Omfang av samhandling, basert på intervju</li> <li>– Indikatorer for omfang og effektivitet i etatenes kontrollaktivitet (statistikk fra etat)</li> <li>– Fremkommelighet gjennom kontrollstasjonen (rød sone), basert på tellinger og intervju</li> </ul>
<b>Virksomheter</b>	<ul style="list-style-type: none"> <li>– Beslag, som indikator på smuglingslekkasje</li> <li>– Bruken av stasjonen som opplæringsarena, intervju</li> </ul>
<b>Relevans</b>	<ul style="list-style-type: none"> <li>– Trafikkvolum trender, herunder gammel versus ny bro</li> </ul>
<b>Levedyktighet</b>	<ul style="list-style-type: none"> <li>– Trender med hensyn til kontrollvirksomhet, spedisjonsbehandling mv.</li> <li>– Egen vurdering av alternativ bruk av stasjonen (realopsjon)</li> <li>– Vurdering av levetid på bygningsmassen</li> </ul>
<b>Samf.øk. lønnsomhet</b>	<ul style="list-style-type: none"> <li>– I hovedsak basert på ovennevnte indikatorer</li> </ul>

Evaluators samlede vurdering er at byggingen av den nye kontrollstasjonen på Svinesund synes å være samfunnsnyttig og i så måte et vellykket prosjekt. Gjennomføringen av prosjektet fremstår som effektiv. Effektmålene er i overveiende grad oppnådd, selv om potensialet trolig er større, og den reduserte fremkommeligheten for tungtrafikken de siste årene må adresseres. Det er få negative sidevirkninger, og prosjektet anses som relevant og levedyktig gitt det



tollpolitiske regimet som gjelder. Den samfunnsøkonomiske lønnsomheten er ikke mulig å konkludere om.

Diagrammet nedenfor gir en grov oppsummering av evaluators vurdering for de seks evalueringskriteriene (helt skjønsmessig hva gjelder samfunnsøkonomisk lønnsomhet, da evaluator ikke oppga noen verdi her).



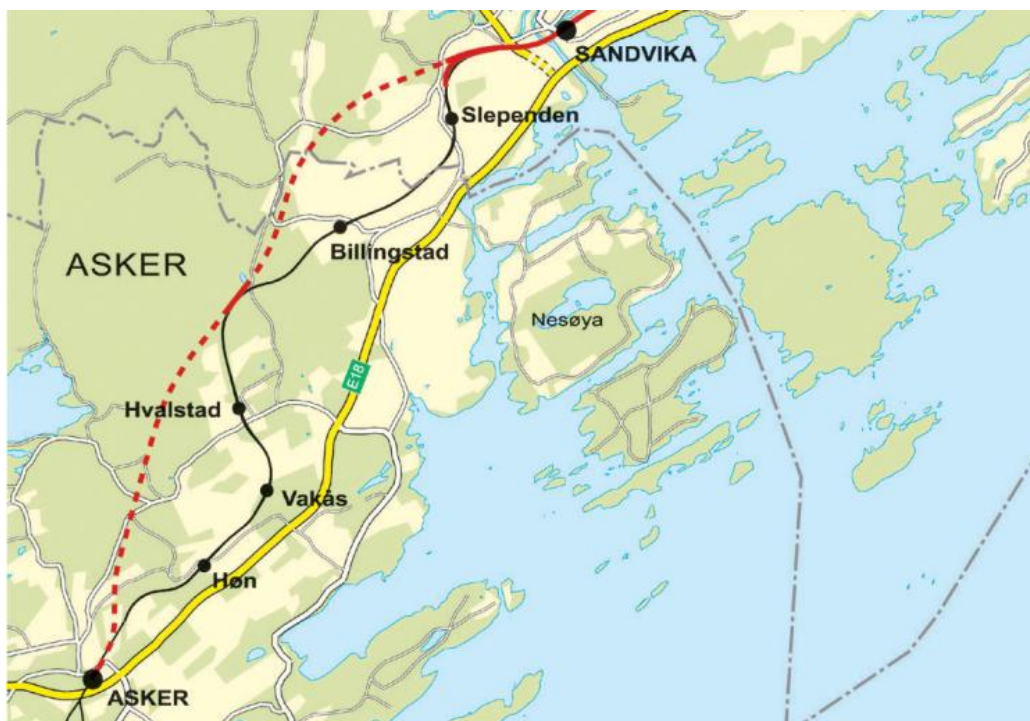
**Figur 3.1.** Svinesund kontrollområde. Evalueringsteamets samlede vurdering av prosjektets vellykkethet

Denne evalueringen er basert på et omfattende materiale av dokumenter, rapporter, styringsreferater, foreliggende statistikk, observasjon på stedet, samtaler med alle sentrale parter, og flere runder med oppfølging og verifisering av informasjonen i etterkant. Teamet bestod av to personer med omfattende evalueringskompetanse, og en samfunnsøkonom. Oppdragsgivers vurdering er at en i dette tilfellet har benyttet den målorienterte metodikken til fulle, og gitt en bred, overordnet og solid fundert vurdering av prosjektet. Den samfunnsøkonomiske analysen ble gjennomført som en kostnads-virkningsanalyse da det ikke var mulig å kvantifisere nytten i dette tilfellet. Det antydes at en kanskje ikke fikk så mye tilleggsnytte ut av denne analysen, ut over at den gir et mer korrekt bilde av den samfunnsøkonomiske kostnaden, både investeringskostnad og kostnadsutvikling ut igjennom driftsfasen.

## 3.2. Nytt dobbeltspor Asker-Sandvika

### Kort introduksjon til prosjektet

Prosjektet omfatter utbygging av 9,5 km dobbeltspor for jernbane på strekningen Asker-Sandvika samt arbeid på begge stasjoner og nytt signalanlegg. Mesteparten av strekningen går i tunnel. Prosjektet var del av den delvis bompengefinansierte Oslopakke 2, som var en satsning på kollektivtiltak i Osloregionen. Arbeidet startet i 2001 og det nye sporet åpnet for trafikk i august 2005. Figuren nedenfor viser eksisterende (sort) og nytt (rødt) dobbeltspor.



Asker-Sandvika var første byggetrinn i et større utbyggingsprosjekt for nye dobbeltspor helt fra Asker til Skøyen (17,4 km), jevnfør figur under. Utbyggingen av Lysaker stasjon ble ferdigstilt i 2009 og strekningen Sandvika-Lysaker åpnet høsten 2011. Det siste byggetrinnet, Lysaker-Skøyen, er nå lag på is og vil ses i sammenheng med Skøyen-Oslo, som ikke var med i den opprinnelige planen.

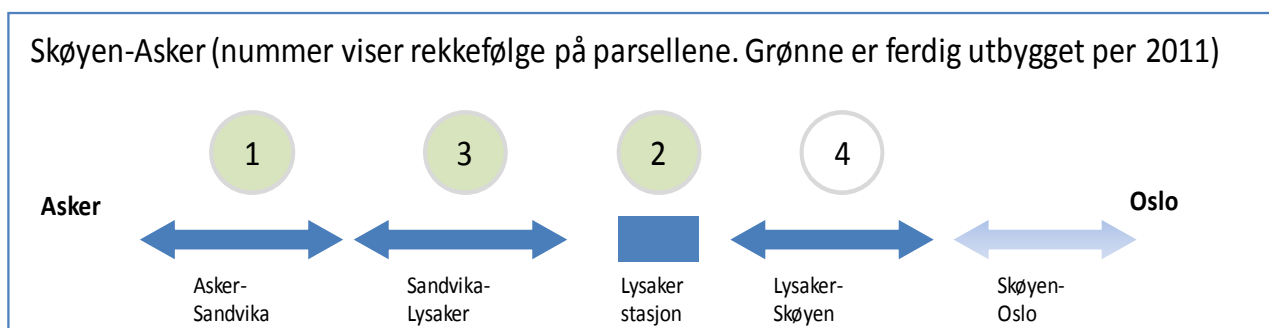
### Bakgrunn og tidligfase

Toglinjen mellom Oslo og Asker er den mest trafikkerte i Norge, og det har siden tidlig på 1990-tallet vært et økende kapasitetsproblem. Med kun ett dobbeltspor ble hurtiggående tog sinket av lokaltogene som stanset på alle stasjoner, og det var derfor ønske om eget dobbeltspor for hurtigtogene. I tillegg visste man at det

lokale transportbehovet ville øke ytterligere med avviklingen av Fornebu som hovedflyplass fra 1998 og utbygging av dette området til boliger og næringsformål.

Det ble gjennomført en rekke samfunnsøkonomiske analyser og andre utredninger om nytt dobbeltspor, fra tidlig 1990-tall og frem til prosjektet ble vedtatt av Stortinget i 2001. Analysene spriker med hensyn til resultater og anbefalinger, men også med hensyn til avgrensning av problemstilling, forutsetninger, underliggende prognoser etc. I én av analysene så man på ulike delstrekninger separat, deriblant Asker-Sandvika. Mangelen på transparens i analysene gjør det vanskelig å sammenlikne resultatene.

De fleste utredningene gjaldt hele strekningen Asker-Skøyen under ett. Ingen av studiene så på tiltak mot den virkelig store flaskehalsen, nemlig tunnelen under og gjennom Oslo. Studiene har dels vurdert ulike traseer, herunder måter å betjene Fornebu på, og dels ulike utbyggingsrekkefølger. Valget om å starte ytterst (Asker-Sandvika) var særlig omstridt, og når man til slutt gikk for dette var begrunnelsen primært at en der hadde kommet lengst i planarbeidet.



## Evalueringen

Det nye dobbeltsporet har vært i drift i syv år, noe som skulle tilsi at tidspunktet for evaluering er egnet. Siden Asker-Sandvika imidlertid er et delprosjekt i en større utbygging, kan en ikke forvente at de fulle gevinstene tas ut før hele strekningen til Skøyen er utbygd med nytt dobbeltspor. Det er uklart når dette skal skje. Evaluator har forsøkt å dokumentere de effekter av Asker-Sandvika som hittil er realisert, og sammenholdt dem med ex ante analysen som anslo at 55 % av nytten for prosjektet samlet ville bli tatt ut når denne delstrekningen var på plass.

Evalueringsteamet har valgt en samfunnsøkonomisk analyseramme, med fokus på oppfølging av kostnader og nytte. Dette anses å være ekvivalent med en vurdering av de to første evalueringskriteriene (produktivitet og måloppnåelse), idet måloppnåelse for dette prosjektet ligger nært opp mot ”samfunnsøkonomisk nytte”. Som et tillegg til den samfunnsøkonomiske vurderingen har man også

drøftet strategisk vellykkethet innenfor rammen av de tre siste evalueringskriteriene (virkninger, relevans og levedyktighet).

Evaluator har videre vært opptatt av å *dokumentere* de konklusjoner som trekkes om samfunnsøkonomisk lønnsomhet, med kvantitative, skriftlige og mest mulig objektive data. En ønsket å begrense bruken av intervju til å utfylle og forklare det dataene viser. En betydelig utfordring har derfor vært mangelen på gode kvantitative data, spesielt om utviklingen i antall reisende. Når det gjelder strategisk vellykkethet har en derfor likevel i stor grad valgt en kvalitativ tilnærming.

### **Prosjektlogikk og målstruktur**

Selve utbyggingsprosjektet Asker-Sandvika var klart definert med resultatmål for tid, kostnader og kvalitet.

Evaluator har ikke funnet eksplisitte effektmål, og har derfor valgt å selv trekke ut fire relevante mål fra underlagsdokumentene for det større prosjektet Asker-Skøyen: *Kortere reisetid*, både for fjerntog og lokal pendlertrafikk, står da sentralt. Videre *økt kapasitet* på strekningen, slik at togtilbudet kan utvides og omfanget av forsinkelser reduseres. Asker-Sandvika var del av Oslopakke 2, hvor *økning i kollektivandelen* var et viktig mål, og det samme var *bedre lokalmiljø*. Hva som var ambisjonsnivået for byggetrinn Asker-Sandvika er noe uklart, men i en av de sentrale nytte-kostnadsanalysene fremgår som nevnt at 55 % av samlet nytte forventes realisert straks Asker-Sandvika står ferdig.

Som uttrykk for samfunns mål valgte evaluator å vise til de overordnede transportpolitiske målene (daværende NTP). Dette er generelle mål, som ivaretar både generell velferd, trafiksikkerhet, distriktspolitikk og miljø.

Den logiske sammenhengen mellom disse målene synes å være tilfredsstillende, dog avhenger realiseringen av effektmålene sterkt av at resten av strekningen bygges ut slik at ikke nye flaskehalser motvirker gevinstene.

### **Produktivitet**

Evaluator har ikke undersøkt resultatmålenes godhet og realisme, men legger til grunn at de er i samsvar med en effektiv gjennomføring (blant annet fordi prosjektet har gjennomgått KS2).

En uforutsett kostnad (400 mill. kroner) var at signalanlegget måtte byttes et godt stykke ut i prosjektet. Det skyldtes en overordnet føring om å benytte en bestemt type anlegg som senere skulle vise seg lite hensiktsmessig. For øvrig synes styringen av prosjektet å ha vært tilfredsstillende på alle nivåer.



Sluttkostnaden ble på 3,7 mrd. kroner, noe som var innenfor kostnadsrammen. Prosjektledelsen var sterkt opptatt av kostnadskontroll og fulgte blant annet deler av KS2-rapportens forslag til ”kuttliste” da man stod i fare for overskridelser. Dette er isolert sett positivt, men samtidig påpeker evaluatør at kutt ofte gir redusert nytte i et livssyklusperspektiv. Som eksempel nevnes at maksimalhastigheten i tunnelen ble satt ned – riktig nok kan dagens tog allikevel ikke akselerere opp til denne hastigheten, men man taper samtidig opsjonen på å sette inn raskere tog i fremtiden.

Tidsmessig leverte prosjektet som planlagt.



### ***Måloppnåelse***

Evaluatør har som nevnt tolket måloppnåelse som samfunnsøkonomisk nytte. Det betyr at det primært er de to første effektmålene, kortere reisetid og økt kapasitet, en forsøker å følge opp under dette kriteriet.

*Kortere reisetid* kan kontrolleres ved å studere rutetabeller før og etter utbyggingen. En finner at enkelte av togtypene har fått en liten reduksjon i reisetid over denne perioden, men langt mindre enn det som var anslått i forkant.

*Økt kapasitet* ventes å manifestere seg i form av 1) økt antall tog som bruker sporet og eventuelt 2) bedre punktlighet. Når det gjelder 1) så kan også dette kontrolleres

ved å studere togtabeller. Evaluatør finner at antall tog har økt noe, dog kun for regiontogene. Når det gjelder 2) har en fått noe data fra Jernbaneverket, men ikke greid å dokumentere bedre punktlighet. Det påpekes at det er metodisk meget krevende å dokumentere effekten av en konkret investering på omfanget av forsinkelser andre steder i nettet.

Nytten av en kortere reisetid og økt kapasitet avhenger videre av hvor mange passasjerer som benytter togtilbudet, samt hvor store godsmengder som fraktes med tog. Til tross for iherdig innsats har evaluatør ikke lyktes med å få tilgang til trafikkdata, annet enn et svært grovt estimat som antyder 22 % passasjervekst på strekningen. Interessant nok ligger dette *under* den forventede veksten i nullalternativet (37 %). Det tyder på at veksten ikke skyldes nytt dobbeltspor men eksterne faktorer som vekst i inntekt, folketall, sysselsetting etc. Det fremgår ikke klart av ex ante-analysene hvilke prognoser som der ble brukt for disse variablene. En annen potensiell forklaring på veksten kunne vært økende kapasitetsproblemer på vei (E18), men telldata fra Statens Vegvesen viser at trafikken her har vært stabil.

Samlet sett synes gevinstene derfor å være små, noe som tyder på at den samfunnsøkonomiske lønnsomheten var overvurdert. I intervjuene oppgis de påfølgende byggetrinn som forklaring. Fulle gevinster kan ikke realiseres før hele strekningen er bygd ut. I tillegg nevnes at en fullstendig reoptimalisering av ruteplaner ikke kan skje før i 2014.

Når det gjelder de siste to effektmålene, *økning i kollektivandelen* og *bedre lokalmiljø* kan en argumentere for at måloppnåelse for disse betinger økt togtrafikk på bekostning av veitrafikk. Kollektivandelen vil, alt annet likt, øke når antall reisende på toget øker, noe som altså ikke har skjedd i den utstrekning som prognosene tilsa. Dermed er det vanskelig å tenke seg at målet om økt kollektivandel nås.

### ***Virkninger***

Under dette punktet diskuteres relevante virkninger som *ikke* fanges i tradisjonelle samfunnsøkonomiske analyser.

Et viktig spørsmål er mulighetene for bredere ringvirkninger (omtalt som "wider economic benefits"), i betydningen økt effektivitet i regionen som følge av økt pendleromland. Dette er virkninger som transportøkonomer har vært veldig opptatt av de senere år. Evaluatør mener imidlertid at når *direkte gevinster*, i form av redusert reisetid og bedre togtilbud, kun i begrenset grad er realisert, kan en heller ikke vente store *indirekte gevinster*.

Videre diskuteres samfunnsålet knyttet til livskraftig næringsliv over hele landet, et fordelingspolitisk mål som heller ikke fanges nytte-kostnadsanalyser. I dette

ligger imidlertid et ønske om å tilgodese økonomisk svake regioner, noe Osloregionen ikke er.

### **Relevans**

I en større sammenheng er det samfunnsnyttene av å bygge nye dobbeltspor i hele vestkorridoren som må vurderes. Og gitt at hele strekningen skal bygges ut, er spørsmålet om det var fornuftig å starte med ytterstrekningen. Evaluatoren mener at ingen av disse spørsmålene besvares godt i analysene som ble gjort *ex ante*. Manglende transparens gjør at det er vanskelig å forstå hva analysene innebærer. Videre påpekes det at når utbyggingen av indre deler forsinkes, og ruteplaner ikke legges om i tide, trekkes dette uansett lønnsomheten ned *ex post*.

Transport er sjelden et mål i seg selv, men følger av andre behov. Utviklingen i demografi, inntekter og andre relevante variabler kan derfor gi en indikasjon på transportbehovet. Dette er å betrakte som et relativt ”stabilt” behov da det endrer seg sakte over tid. På den annen side avhenger folks reiseomfang og valg av transportform, også av myndighetenes virkemiddelbruk. Økonomiske virkemidler for å løse kapasitetsproblemer benyttes i liten grad i dag, selv om en kan argumentere for at dette ville være et langt billigere konsept enn utbygging av infrastrukturen. Dersom myndighetene begynner å prise infrastruktur mer optimalt, kan etterspørselen etter både tog og andre transportformer endres dramatisk.

### **Levedyktighet**

Evaluatoren drøfter ulike aspekter ved prosjektets levedyktighet. For det første kan en si at kollektivtransport er viktig i et miljøperspektiv, både i forhold til klima og også lokalmiljø som var viktig i Oslopakke 2. Dersom prosjektet over tid vil føre til en betydelig overgang fra bil til tog, vil det bidra positivt. Evaluatoren påpeker imidlertid at også jernbaneutbygginger har negative klimaeffekter, selv om dette aldri ble utredet i forkant.

For det andre kan kollektivtransportprosjekter sies å være positive i et sosialt perspektiv, da ikke alle har tilgang til egen bil.

For det tredje bør levedyktighet ses i et økonomisk perspektiv. Det er uklart hva man la til grunn om utviklingen i drifts- og vedlikeholdskostnader i dette tilfellet. Drifts- og vedlikeholdskostnader for jernbane har ligget langt over konsumprisindeksen det siste tiåret. Dette trekker lønnsomheten av prosjektet ned, og sannsynligheten for at andre investeringsprosjekter må vike trekkes opp.

### **Samfunnsøkonomisk lønnsomhet**

Det har altså ikke vært mulig å etterprøve den samfunnsøkonomiske analysen fullt ut, selv om evaluatoren har lagt stor vekt på å få til dette. Som drøftet over, kan det

ikke utelukkes at investeringen per i dag (syv år etter) har hatt meget små effekter og at gevinstene var overvurdert i forkant. Dette begrunnes bl.a. med at togtilbudet og rutetabellene kun er marginalt forbedret. Samtidig indikerer dataene at det har vært passasjervekst i perioden, men da trolig av andre årsaker enn bedre togtilbud. Det er derfor en betydelig risiko for at den samfunnsøkonomiske lønnsomheten var overvurdert, i alle fall hva gjelder de kvantifiserbare virkningene.

### Oppsummert om prosjektets vellykkethet

Tabellen under gir en oversikt over de viktigste (mest valide) indikatorene som ble benyttet for å belyse hvert av evalueringskriteriene.

**Tabell 3.2** Nytt dobbeltspor Asker-Sandvika. Indikatorer

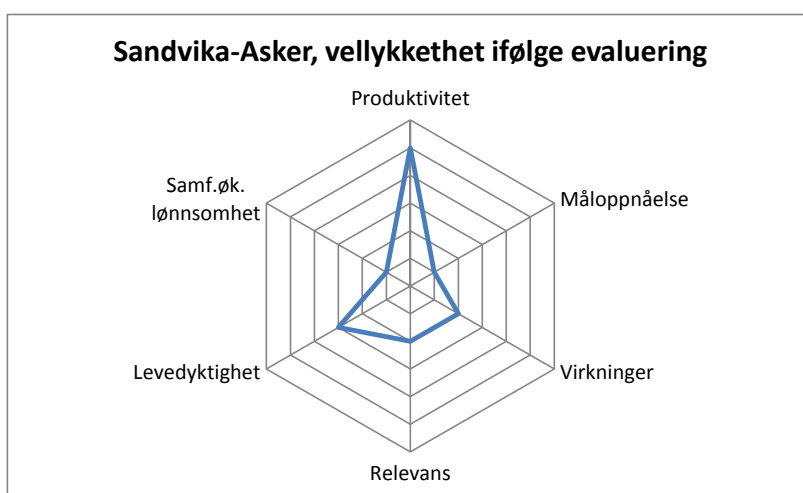
Viktigste indikatorer	
<b>Produktivitet</b>	<ul style="list-style-type: none"> <li>– Tidsmål, vurdert ved åpningstidspunkt mot plan</li> <li>– Kostnadsmål, vurdert ved prosjektregnskap opp mot styrings- og kostnadsramme i KS2</li> <li>– Egen analyse av gjennomføringen og kuttbrukens antatte effekt på kvalitet</li> </ul>
<b>Måloppnåelse</b>	<ul style="list-style-type: none"> <li>– Reisetid, basert på rutetabeller</li> <li>– Frekvens, basert på rutetabeller</li> <li>– Punktlighet, basert på data fra etat</li> <li>– Trafikkmengder, kun grovt estimat fra NSB</li> </ul>
<b>Virknings</b>	<ul style="list-style-type: none"> <li>– Indirekte virkninger, egen drøfting</li> </ul>
<b>Relevans</b>	<ul style="list-style-type: none"> <li>– Langsiktig behov og optimal rekkefølge på utbygginger, egen drøfting og studier av tidligere nytte-kostnadsanalyser</li> </ul>
<b>Levedyktighet</b>	<ul style="list-style-type: none"> <li>– Drifts- og vedlikeholdskostnader for jernbane, statistikk over tid</li> <li>– Egen drøfting</li> </ul>
<b>Samf.øk. lønnsomhet</b>	<ul style="list-style-type: none"> <li>– Basert på ovennevnte indikatorer</li> </ul>

Generelt er det svært begrensende å vurdere et prosjekt som kun er ett byggetrinn i en større utbygging. Det evaluators mener å kunne slå fast er at kostnadskontrollen i Asker-Sandvika-prosjektet var god, dog medførte de omfattende kuttene en



reduksjon i kvaliteten. Måloppnåelsen/nyttjen har derimot vært meget begrenset så langt. I et strategisk perspektiv kan dette likevel være en fornuftig investering, dog mener evaluatoren at dette burde vært bedre utredet, herunder opp mot alternative tiltak som for eksempel prising av infrastruktur. En burde også gjort grep for å sikre raskere gevinstrealisering.

Diagrammet nedenfor gir en grov oppsummering av evaluators vurdering for de seks evalueringskriteriene.



**Figur 3.2** Nytt dobbeltspor Asker-Sandvika. Evalueringsteamets samlede vurdering av prosjektets vellykkethet

Denne evalueringen er bygget opp rundt en samfunnsøkonomisk analyseramme, og er i stor grad basert på foreliggende skriftlig materiale, kostnadstall og statistikk, men i liten grad egen observasjon og samtaler med sentrale informanter. Teamet bestod av tre personer med omfattende transportøkonomisk kompetanse, men med begrenset evalueringskompetanse og liten erfaring med den målorienterte modellen. Oppdragsgivers vurdering er at en i dette tilfellet ikke har benyttet den målorienterte evalueringsmetodikken etter intensjonen og at perspektivet dermed ble noe smalt. Spesielt har en ikke lyktes i å gjøre en bred og overordnet vurdering av prosjektets strategiske vellykkethet. Dette er da også krevende når prosjektet kun er en parsell i en større utbygging. Hva gjelder datainnsamlings- og analysemetoder kunne en med fordel ha kombinert den kvantitative tilnærmingen med mer kvalitative vurderinger. Evalueringen har imidlertid satt fokus på tilgjengelighet av viktige nøkkeltall i jernbanesektor, noe som er viktig å ta med seg i senere evalueringer.

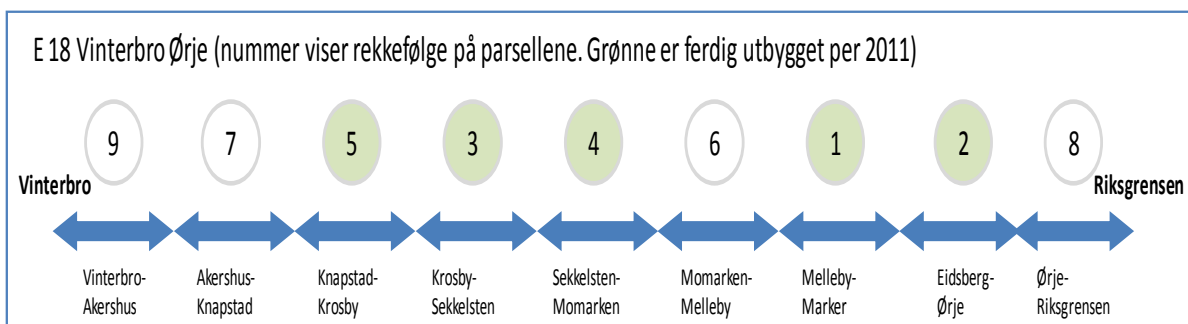
### 3.3. Veiparsell Momarken-Sekkelsten

#### Kort introduksjon til prosjektet

Prosjektet omfatter en 6,2 kilometer lang veiparsell på E18 mellom Mysen og Askim i Østfold. Veien er bygget som en smal fire felts motorvei. I tillegg er det som del av prosjektet bygget to toplanskryss, en miljøkulvert og to overgangsbruer, samt en rekke tiltak på sideveger.

Endelig investeringsbeslutning for utbyggingen ble tatt i 2004 med en styringsramme på 450 mill. kr. og en kostnadsramme på 490 mill. kr. Anleggsperioden pågikk fra høsten 2005 til høsten 2007, med offisiell åpning av vegen 31. august 2007. Investeringskostnaden kom på 643 mill. kr. (2010) og ble finansiert med bompenger og statlige bevilgninger.

Parsellen er en del av et større prosjekt E18 Ørje – Vinterbro, som skal realisere en ny E18 gjennom Østfold fylke. Totalt består prosjektet av ni parseller på til sammen ca. 70 kilometer.



Per i dag er fem av de ni parsellene ferdig bygget, en parsell er under bygging, mens tre parseller er under planlegging. Vest for Momarken – Sekkelsten ligger parsellen Krosby – Sekkelsten som stod ferdig høsten 2005. I øst ligger parsellen Momarken – Melleby som forventes ferdigstilt i løpet av 2014.

#### Bakgrunn og tidligfase

E18 fra Oslo til svenskegrensen ved Ørje utgjør sammen med Fylkesveg 2 og Østfoldbanen Norges nest viktigste transportkorridor til utlandet over land. Dette gjelder både for gods- og persontrafikk. I 2004 gikk 25 % av all godstransport på veg til utlandet over E18.

Trafikkavviklingen på gamle E18 i Østfold har lenge vært belastet med kø og ulykker. For å bedre fremkommeligheten og sikkerheten på vegen ble det på 1990-tallet startet et arbeid med å etablere en ny veg med høyere standard. Rundt årtusenskiftet ble planene for utbygging av E18 samlet under betegnelsen

«Østfoldpakka», som også inkluderte oppgradering av E6 i Østfold, vegforbindelse til havner, utbedringer på jernbane og transportsystemet i byene i fylket. Blant lokale politikere og næringsliv var det et sterkt ønske om å få gjennomført utbyggingen raskt. Det ble derfor besluttet å finansiere deler av kostnadene gjennom bruk av bompengefinansiering. Gjennom St.prp. nr. 26 (1999-2000) «Østfoldpakka» som ble vedtatt november 1999 ga Stortinget aksept for opprettelsen av bompengeselskap for dette formålet, og i St.prp. nr. 1 (2004-2005) ble det gitt klarsignal for oppstart av anleggsarbeidene.



*E18 gjennom Monaryggen. Foto: Statens Vegvesen*

## **Evalueringen**

Evaluator mener at Momarken – Sekkelsten må betraktes som et lite komplisert prosjekt å evaluere. Målstrukturen var forholdsvis enkel, og indikatorer for måloppnåelse kunne lett identifiseres. Evaluator benyttet i hovedsak en kvantitativ tilnærming, hvor data og prognoser for trafikkmengder og ulykker var sentrale parametere. I tillegg til å fungere direkte som indikatorer for prosjektets måloppnåelse, benyttet man også slike data som indikasjon på indirekte virkninger som bør undersøkes nærmere. Trafikkdata må imidlertid settes i en kontekst for å kunne tolkes rett og evaluator har derfor også supplert med enkelte kvalitative vurderinger.

En har i evalueringen også vært opptatt av at prosjektet må sees i et systemperspektiv. Vegparsellen er vurdert både som del av et lokalt, regionalt og et nasjonalt transportsystem.

## Prosjektlogikk og målstruktur

Prosjektets begrunnelse og mål må sees i sammenheng med utbyggingen av hele E18 gjennom Østfold. Det var et sterkt behov for å bedre trafikkavviklingen og redusere ulykkesfrekvensen på denne veien. Fremkommeligheten var lav også ved lite trafikk på grunn av mange avkjørsler og nedsatt fartsgrense gjennom tettsteder. Det var også stor miljøbelastning for tettstedene.

I underlaget for E18-utbyggingen var det ikke formulert konkrete samfunns mål, men evaluator oppfatter målene om reduksjon av ulykkes-, reise- og miljøkostnader som samfunns mål i seg selv (jf. Nasjonal Transportplan). Implisitt er det også lagt til grunn et mål om å bedre rammebetingelsene for utenlandshandelen, samt bosetting og næringsaktivitet i regionen.

Det var spesifisert tre konkrete effektmål:

- Redusere antallet ulykker og alvorlighetsgraden på de ulykkene som oppstår i E18-korridoren
- Redusere reisetiden for trafikanter i E18-korridoren
- Gi beboere i tettstedsområdene langs dagens E18 et bedre nærmiljø ved å redusere lokal støy- og luftforurensing

Sekundært skulle prosjektet gi beboerne og trafikanter i området økt trivsel og trygghetsfølelse.

### **Produktivitet**

Evaluator mener at prosjektet kommer forholdsvis godt ut på kriteriet produktivitet, gitt at en vurderer dette i et snevert perspektiv som evnen til å levere de spesifiserte leveransene tids- og kostnadseffektivt. Veien tilfredsstiller tekniske og funksjonelle kvalitetskrav, og prosjektet ble gjennomført raskt og veien åpnet før planlagt tidspunkt. Evaluator har beregnet at kostnadsrammen ble overskredet med om lag 5 % og styringsrammen med om lag 14 % basert på prisjusterte tall (2010). En hevder at dette dels kan forklares med mangelfull planlegging og mangel på interne ressurser, men at de fleste avvikene kan forklares som normale utslag av usikkerhet. En stiller spørsmål ved om disse usikkerhetene i tilstrekkelig grad var reflektert i kostnadsestimatet som lå til grunn for rammen.

Betraktes utbyggingen som et delprosjekt i det store prosjektet med å bygge ny E18 gjennom Østfold, er bildet mer problematisk. Den overordnede styringen av planprosessene i etaten synes å ha vært mangelfull, hvilket gjorde den tidlige planperioden i prosjektet unødvendig lang og kostbar. Oppstykkningen i mange små parseller har dessuten nødvendiggjort en del investeringer i midlertidige tiltak på

atkomstveger i påvente av finansiering og utbygging av tilstøtende parseller. Dette fremstår som en lite kostnadseffektiv måte å bygge veg på, men er delvis en følge av rammebetingelsene for veibygging i Norge.

### **Måloppnåelse**

Prosjektet E18 Momarken – Sekkelsten hadde tre effektmål som beskrevet over. Alle tre målene vurderes som oppnådd. En har fått mer forutsigbar og noe kortere reisetid og stor nedgang i antallet ulykker, spesielt ulykker med alvorlig utfall. I tillegg har trafikken langs den gamle veien blitt redusert med over 80 %. En antar at det kan forventes en ytterligere gevinst på alle målområder når den østlig tilstøtende parsellen, Melleby – Momarken, åpner i 2014.

### **Virkninger**

I evalueringen av investeringstiltakets virkninger har denne evalueringen skilt mellom tre geografiske nivåer; lokale, regionale og nasjonale virkninger.

På lokalt plan har vegen gitt et bedre nærmiljø med hensyn på luftforurensing og støy. Vegene er blitt anlagt i en ny trasé, noe som har medført at en del landbruks- og næringsareal er blitt båndlagt. Siden vegen er bompengefinansiert har også de betalbare reisekostnadene for trafikantene i området blitt noe høyere.

For Indre Østfold som region har den nye vegen medført kortere reisetider for trafikanter og gods på veg, og bruk av bil er nå blitt relativt mer attraktivt for blant annet pendlerreiser til og fra Osloområdet. NSB som driver persontransport på Østfoldbanen oppgir at ny E18 har gitt en merkbar nedgang i passasjertallene for jernbane. Busstrafikken (som fortsatt benytter den gamle veien) har ikke merket noen endring i antallet passasjerer – en har ikke hatt nedgang, men heller ikke oppgang til tross for at bussen nå har bedre fremkommelighet og punktlighet.

En har også stilt spørsmål ved om utbyggingen har gitt økonomiske og sysselsettingsmessige ringvirkninger i området. Det er imidlertid ikke noe som tyder på dette, og det er uansett vanskelig å påvise noen bredere økonomiske virkninger av vegutbyggingen uten detaljerte studier.

Nasjonale virkninger er vurdert innenfor rammen av en samfunnsøkonomisk analyse, se under.

### **Relevans**

Evaluator legger til grunn at behovet bak investeringen var knyttet til de tre målområdene nevnt over, og konkluderer med at prosjektet er relevant både nå og i fremtiden. Det er dokumentert at veien har hatt en trafikkvekst som er langt høyere enn de estimerer som forelå før bygging. Trafikkprognosene for E18 også i årene som kommer viser en betydelig vekst. En antar derfor at problemene som lå

til grunn for tiltaket ville vært enda større i dag, og ville blitt stadig forsterket fremover, i en situasjon uten utbygging.

En har imidlertid ikke gått nærmere inn på det mer overordnede spørsmålet om veiutbygging er det mest relevante av mulige tiltak for å løse disse og andre samfunnsutfordringer.

### **Levedyktighet**

Den nye vegen har god kapasitet og antas å kunne ta unna fremtidig trafikk selv med den veksten som forventes de nærmeste tiårene. En antar videre at ulykkesfrekvensen heller ikke vil øke som en følge av mer trafikk, blant annet siden trafikken legges utenom tettstedene. Dette tilsier at de samfunnsøkonomiske nyttestrømmene vil genereres som forventet i tiden fremover.

Erfaringene etter fem års drift er at drifts- og vedlikeholdskostnadene er innenfor det som kan forventes og håndteres innenfor normale budsjetttrammer. Grunnet høyere trafikk og dermed høyere bompenginntekter enn forventet, antar en at denne delen av prosjektet sannsynligvis kan nedbetales på kortere tid.

I et miljø-/klimaperspektiv kan en derimot si at levedyktigheten er mindre god, da dette er et tiltak som skaper enda mer biltrafikk på bekostning av andre, mer miljøvennlige transportformer.

### **Samfunnsøkonomisk lønnsomhet**

Det er gjennomført en nytte-kostnadsanalyse av investeringen, som omfatter de prissatte virkningene av tiltaket. Analysen er basert på simulering over hele analyseperioden og indikerer at prosjektet kan vise seg å bli svært lønnsomt. Netto nytte estimeres til mellom 800 – 1400 mill. kr (avrundet), avhengig av hvilken trafikkvekst som legges til grunn.

Dette er betydelig høyere enn de resultatene som forelå i beslutningsgrunnlaget for prosjektet. Analysene viste da en negativ til marginalt positiv netto nytte. Dels kan dette tilskrives forskjeller i inngangsparameterne og forutsetningene i analysen (f.eks. nedsettelse av kalkulasjonsrenten) men det skyldes også at trafikkveksten ble mye høyere enn ventet. Evaluatør konkluderer med at avviket er så stort at beslutningsunderlaget for prosjektet med hensyn på samfunnsøkonomisk lønnsomhet må ha vært mangelfullt.

### **Oppsummert om prosjektets vellykkethet**

Tabellen under gir en oversikt over de viktigste (mest valide) indikatorene som ble benyttet for å belyse hvert av evalueringskriteriene.

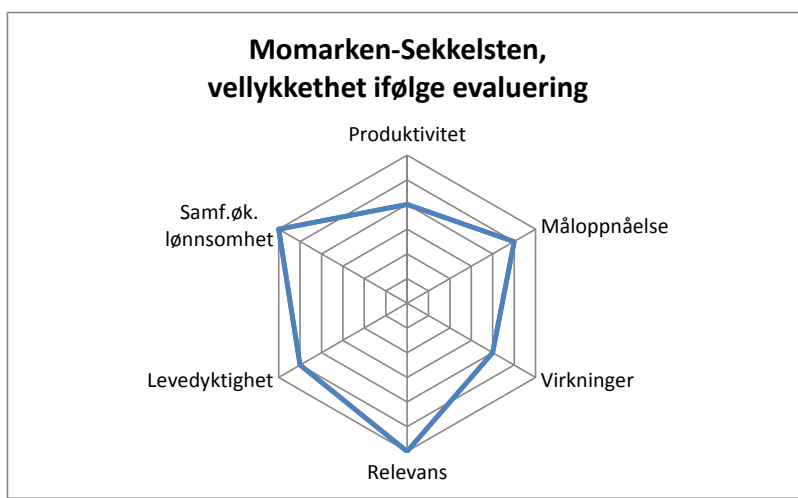
**Tabell 3.3** Veiparsell Momarken-Sekkelsten. Indikatorer

<b>Viktigste indikatorer</b>	
<b>Produktivitet</b>	<ul style="list-style-type: none"> <li>– Kvalitet og funksjonalitet vurdert av aktører som drifter vegen, samt tekniske standarder</li> <li>– Tidsbruk, vurdert ved analyse av prosjektprosessen og referansesjekk</li> <li>– Kostnadseffektivitet, vurdert ved analyse av prosjektprosessen og referansesjekk</li> </ul>
<b>Måloppnåelse</b>	– Måloppnåelse vurdert ved trafikkdata og ulykkesstatistikk
<b>Virkninger</b>	<ul style="list-style-type: none"> <li>– Virkninger på lokalt nivå, vurdert ved egne observasjoner, trafikkdata, ulykkesdata og intervjuer.</li> <li>– Virkninger på regionalt nivå vurdert gjennom innhenting av uttalelser og trafikkdata</li> <li>– Virkninger på regionalt nivå vurdert ved data fra SSB vedrørende sysselsetting og bosetting.</li> <li>– Virkninger på nasjonalt nivå er vurdert gjennom nytte-kostnadsanalyse</li> </ul>
<b>Relevans</b>	– Behov uttrykt i nasjonal transportplan og øvrig beslutningsunderlag sammenliknet med faktisk utvikling (trafikk- og ulykkesdata).
<b>Levedyktighet</b>	<ul style="list-style-type: none"> <li>– Vurdering av veiens kapasitet mot prognoser for trafikkmengder og ulykker og for perioden fram til 2040.</li> <li>– Analyse av lønnsomheten i bompengeneinnkrevingen, både inntekter og utgifter.</li> <li>– Egen vurdering av evne til å vedlikeholde veganlegget i framtiden</li> <li>– Flexibiliteten vurdert med hensyn på muligheter for endring av utforming av vegen og tilpasning til andre fremtidige tiltak.</li> </ul>
<b>Samf.øk. lønnsomhet</b>	– Nytt-kostnadsanalyse basert på modellen EFFEKT versjon 6.43.

Evaluatør mener samlet sett at dette har vært et rimelig vellykket prosjekt. Selve utbyggingen skjedde effektivt, til tross for at prosjektet kom ut med en kostnadsoverskridelse. Dersom en ser på E18-utbyggingen samlet kan det stilles

spørsmål ved effektiviteten i denne type parsellvise utbygging. Måloppnåelsen antas i dag meget god grunnet høy trafikk, og nyttekostnadsanalysen indikerer av samme grunn god fremtidig lønnsomhet. Evaluator har også gitt god score på de strategiske kriteriene, selv om det i et overordnet perspektiv kan stilles spørsmål ved om økt veiutbygging er det samfunnet vil ønske å prioritere fremover dersom det skjer på bekostning av kollektivtrafikk/gods på bane.

Diagrammet nedenfor gir en grov oppsummering av evaluators vurdering for de seks evalueringskriteriene.



**Figur 3.3** Veiparsell Momarken-Sekkelsten. Evalueringsteamets samlede vurdering av prosjektets vellykkethet

Denne evalueringen er, i likhet med evalueringen av Sandvika-Asker, bygget opp rundt en samfunnsøkonomisk analyseramme og i stor grad basert på foreliggende skriftlig materiale, kostnadstall og statistikk. Informantene er hovedsakelig avgrenset til veietaten og prosjektnivået. Ingen i evalueringsteamet hadde erfaring med målorientert evaluering, men begge hadde betydelig erfaring med transportøkonomiske vurderinger. Oppdragsgivers vurdering er at en i dette tilfellet ikke har benyttet den målorienterte evalueringsmetodikken etter intensjonen. En betydelig stor del av arbeidet har gått med til å vurdere den operasjonelle gjennomføringen av prosjektet, og til å forklare den observerte kostnadsoverskridelsen. Videre er det lagt stor vekt på den samfunnsøkonomiske analysen, men denne er hypotetisk og basert på modellberegning av en fremtidig helhet. I dette tilfellet har evalueringsteamet ikke fremskaffet tilstrekkelig grunnlag til å gi en troverdig vurdering av strategiske forhold ved prosjektet slik det fremstår i figur 3.3.



### 3.4. Bygging av Skjold-klasse MTB-er til Sjøforsvaret

#### Kort introduksjon til prosjektet

Prosjektet (2003-) omfatter bygging av fem MTB-fartøyer (missiltorpedobåter) i Skjoldklassen samt oppgradering av en prototyp som var utviklet i et forprosjekt som var slutført i 1999. Fartøyene skal utføre overflateoperasjoner som ledd i Sjøforsvarets utøvelse av sjøkontroll og den nye serien etterfølger tidligere MTB-serie. Til leveransen hører også våpensystemer, opplærings- og treningsystemer, dokumentasjon og reservedeler. Nye sjømissiler (åtte per fartøy) er utviklet i et annet prosjekt, MTB-ene og fregattene var grunnlaget for dette prosjektet.

Båtene omtales som luftputekatamaran, og prosjektet har hatt betydelige elementer av teknologiutvikling. Dette var hovedårsaken til at en først utviklet en prototyp. De nye fartøyene blir noe helt annet enn tradisjonelle MTB-er, med større slagkraft og mobilitet, og evne både til å operere selvstendig og sammen med andre fartøy.

Per 2012 er fortsatt to av fartøyene ikke ferdigstilt og satt i drift. Overlevering av det siste fartøyet er forventet i slutten av 2012 og tidlig i 2013.



#### Bakgrunn og tidligfase

Allerede i 1986 startet man utredninger av alternative konsepter da de gamle MTB-ene begynte å nå sin tekniske og operative levealder. Det er uklart om det ble foretatt reelle konsekvensvurderinger før man landet på det valgte konseptet; få,

store MTB-er med stor mobilitet og slagkraft. MTB-er har tradisjonelt vært mindre fartøy med begrenset utholdenhet og våpenlast.

Det ble tidlig klart at Skjold MTB-ene skulle bygges i Norge, og det er liten tvil om at de næringspolitiske hensynene veide tungt for mange av aktørene. I 1995 hadde Sjøforsvaret ferdigutviklet krav og spesifikasjoner. På dette grunnlag ble det – uten anbud – inngått kontrakt med Kværner Mandal (senere Umoe) om detaljert design og utvikling av en prototyp. Kontrakten inneholdt også en opsjon for senere bygging av seriefartøyer. Prototypen ble levert i 1999 og viste tilfredsstillende resultater for det valgte teknologiske konseptet.

Opsjonen på å kjøpe seriefartøyer utløp i 2000 uten at det ble tatt noen beslutning om anskaffelse. Perioden 2000-03 ble preget av konflikt om prosjektet. Forsvarssjefen mente at en investering i MTB-våpen uten ekstra tilførsel av midler til Forsvaret, ikke kunne anbefales. Han innså at ikke bare investeringskostnaden men også de årlige driftskostnadene ville bli høye, og at dette kunne fortrenge andre og høyere prioriterte anskaffelser og drift. Regjeringen sluttet seg til dette. Stortinget vedtok likevel til slutt prosjektet, under henvisning til at det ville gi en styrket kapasitet til suverenitetshåndheving, territorialsikring og til støtte for sivile operasjoner, som er spesielt viktig i nordområdene.

Det ble gjennomført KS2 som konkluderte med at prosjektet var godt nok forberedt. Endelige bevilgning ble gjort 21. oktober 2003. Deretter ble det inngått kontrakt med Skjold Prime Consortium (SPC), som bestod av Umoe Mandal, med ansvar for skrog og maskineri, Kongsberg Defence & Aerospace (KDA), med ansvar for våpensystemer og DCN International med ansvar for kommando-, kontroll- og informasjonssystemer.

Det kan nevnes at konfliktene om prosjektet fortsatte også etter stortingsvedtaket. Til sammen tre forsvarssjefer har advart mot MTB-prosjektet. Senest i 2007, da 80 % av fartøyene var så godt som ferdigbygde, fremmet daværende forsvarssjef et forslag om å stoppe prosjektet og han omtalte det hele som en feilinvestering.

## **Evalueringen**

De to siste fartøyene er ennå ikke endelig overlevert og kommet i drift. Således er det ikke mulig å gjøre en fullstendig ex post evaluering på nåværende tidspunkt. Spesielt er erfaringen med drift av fartøyene foreløpig begrenset.

Måloppnåelse er uansett vanskelig å måle i fredstid, og krever at en gjør vurderinger rundt situasjonen i hypotetiske krigsscenarier. Her er en i stor grad henvist til subjektive vurderinger. En kompliserende faktor er at de data som tross alt eksisterer, i et forsvarsmateriellprosjekt er konfidensiell og kan dermed ikke omtales i evalueringsrapporten.

Evaluatør har altså vært henvist til dokumenter som beskriver *intensjonene* med prosjektet, samt intervjuer med aktører som har oppfatninger om hvordan prosjektet vil komme ut. Et bredt utvalg av informanter er intervjuet, herunder fra Forsvaret, Forsvarsdepartementet og leverandørene.

### **Prosjektlogikk og målstruktur**

Leveransen og innholdet i prosjektet er klart definert, med resultatmål for tid, kostnader og kvalitet.

Det er ikke formulert konkrete effekt- og samfunns mål. Evaluatør har løst dette ved å trekke ut målformuleringer for effektmål fra den stortingsproposisjon som lå til grunn for bevilgningen. Det er snakk om forsvarsmessige mål knyttet til MTB-enes funksjon i forhold til å utføre overflateoperasjoner, overvåke og kontrollere store sjøområder, levere langtrekkende ildkraft mot sjø- og landmål, samt en del andre oppgaver som å støtte Kystvakten, beskytte petroleumsinstallasjonene og delta i søk og redningsaksjoner.

Videre identifiserer evaluatør fem overordnede samfunns mål for sektoren som er de nyeste sikkerhetspolitiske hovedmålsettinger; *Forebygge krig og trusler*, *Bidra til fred, stabilitet og FNs rettsorden*, *Norsk suverenitet*, *Forsvare Norge og NATO* og *Forsvare mot statlige og ikke-statlige angrep*. Evaluatør vurderer at den logiske sammenhengen fra effektmål til samfunns mål er ”relativt god”. Dog påvirkes samfunns målet *Bidra til fred, stabilitet og FNs rettsorden* i liten grad av MTB-ene. Det påpekes også at effektmål knyttet til andre oppgaver som søk og redningsaksjoner, er av mindre relevans for samfunns målene og bør trolig betraktes som sidevirkning heller enn som mål.

Som nevnt var også næringspolitiske hensyn viktige, selv om dette ikke var formelle mål for prosjektet.

### **Produktivitet**

De seks fartøyene vil, når de er klare, ha en ytelse, funksjonalitet og kvalitet som tilfredsstillende resultatmålene. I mangel av referanseprosjekter vites ikke om tids- og kostnadsrammene var fornuftige i utgangspunktet, selv om det faktisk at de hadde vært gjenstand for KS2 burde tale for dette. Evaluatør har gjennomført en egen studie av gjennomføringsfasen med sikte på å avdekke hendelser eller faktorer som kan ha påvirket produktiviteten.

Ut over normal og forventet prisstigning har det vært to utvidelser av budsjettammene i gjennomføringsperioden. Begge var fornuftige i lys av behov som oppstod. Sluttkostnad ligger per 2012 an til å bli på 5 mrd. kroner, som er godt under både kostnads- og styringsramme når disse prisjusteres til 2012-kroner.

I så måte synes prosjektet å ha vært kostnadseffektivt gjennomført. (Nye sjømissiler til ca. 2 mrd. kroner kommer i tillegg, som eget prosjekt).

Det var likevel en viktig hendelse *i forkant av prosjektet*, som tilsier at prisen på fartøyene kunne blitt betraktelig lavere – inntil 1 mrd. kroner. Det gjelder det faktum at opsjonen om levering av seriefartøyer ikke ble utløst innen fristen i 2000. Da det endelig ble vedtatt å starte serieproduksjon var leverandørkonsortiet i realiteten i en monopolsituasjon, noe de utnyttet ved forhandlingsbordet. I tillegg var den teknologiske risikoen økt, fordi det nå var behov for en viss teknologisk videreutvikling, samt at enkelte komponenter ikke lenger var i markedet.

Når det gjelder tidseffektiviteten så ble prosjektet 3-4 år forsinket i forhold til de kontraktsfestede leveransedatoer. Tidsplanen som ble fastsatt i 2003 synes imidlertid å ha vært for ambisiøs, i lys av det nevnte oppholdet mellom prototyp og serieproduksjon, og de teknologiske utfordringene en da måtte forvente.

### **Måloppnåelse**

Som nevnt er det er ikke formulert konkrete effekt- og samfunns mål for prosjektet, men evaluator har selv trukket effektmål fra underlagsdokumentene, og i tillegg benyttet overordnede mål for Forsvaret som samfunns mål. Intervjuene har i stor grad tatt utgangspunkt i målene og hvorvidt disse er oppnådd.

Samlet indikerer intervjuene at fartøyene, når de er ferdige, trolig i stor grad vil nå effektmålene. Flere informanter påpeker imidlertid at det er vanskelig å gjøre denne type isolert vurdering av en enkelt fartøyklasse fordi ulike fartøyklasser fungerer innenfor et integrert forsvar. Derfor blir MTB-enes relevans i forhold til Sjøforsvaret i sin helhet et nøkkelspørsmål jevnfør nedenfor.

Når det gjelder de næringspolitiske målene, finner evaluator at disse er oppnådd kun i begrenset grad. For Umoe var kontrakten på 2 milliarder kr og bidro til at bedriften overlevde. Kontrakten innebar en sysselsetting som på det høyeste var på ca. 400 personer. På grunn av de store forsinkelsene og de derav følgende dagsbøter ble imidlertid inntjeningen begrenset og etter hvert gikk verftet med betydelige tap.

For begge de to norske leverandørene var det et mål at prosjektet skulle bidra til videre utvikling av produkter, teknologi og kompetanse for videre salg. Dette har heller ikke skjedd så langt.

### **Virkninger**

Evaluator gjør også en samlet vurdering av prosjektets virkninger for ulike aktører. Virkningene for Forsvaret og leverandørene har vært både positive og negative. Med Skjoldfartøyene er Norge i ferd med å få en av Europas mest moderne

mariner. Samtidig er denne anskaffelsen neppe optimal i en større forsvarsmessig sammenheng, og kan f.eks. innebære at fregattene, som har et større bruksområde, må redusere sin aktivitet.

For Umoe var prosjektet i utgangspunktet positivt fordi det var medvirkende til at verftet bestod, men negativt at det så ble et tapsprosjekt. For de andre leverandørene oppnådde man ikke å komme inn på eksportmarkedet for produkter og tjenester som ble utviklet og levert til Skjoldprosjektet, men sammen med søsterprosjektene kan det fortsatt gi utviklingsperspektiver for bedriftene.

### **Relevans**

Spørsmålet om prosjektets relevans er omstridt og er det kriterium som reiser flest spørsmål. Et hovedspørsmål er hvilken relevans Skjoldfartøyene har i forhold til Norges forsvarsmessige behov, herunder i forhold til andre fartøysier og plattformer. Svaret på dette spørsmålet har endret seg over tid.

Da man begynte å utrede prosjektet på slutten av 80 tallet, var den sikkerhetspolitiske situasjonen preget av invasjonstrussel og et hovedforsvarsområde mot en slik trussel var lokalisert i Troms. Da den kalde krigen var slutt, var invasjonstrusselen redusert, og Forsvaret har vært igjennom en betydelig omstilling, hvor omfanget av kostbar infrastruktur og statiske forsvarsanlegg har blitt redusert til fordel for såkalte ekspedisjonære styrker, som i mindre grad er bundet til forsvar av eget land. MTB-ene er ikke egnet for å operere i åpent hav og derfor mindre relevant for dette behovet.

Dessuten er det betydelig overlapp mellom MTB-fartøyene og andre fartøysier, spesielt fregattene og ubåter, men også neste generasjons kampfly vil ha evnen til å uskadeliggjøre store overflatefartøyer med langtrekkende missiler. Dette er mye av forklaringen på den fagmilitære motstanden mot MTB-prosjektet. Et annet moment som er blitt trukket frem er at MTB-fartøyene er overdimensjonert, med sitt avanserte og kostbare utstyr, i forhold til de oppgavene de primært skal utføre (tilstedeværelse og suverenitetshevdelse i kystnære områder).

Til slutt drøfter evaluator prosjektets relevans i forhold til markedet. Dersom prosjektet utvikler teknologi som det er behov for ut over Skjold anskaffelsen, målt ved verdi av spin-off-salg, kan en si at dette bidrar positivt til prosjektets relevans. Erfaringene hittil er at dette i liten grad har skjedd. Riktignok har fartøyene ennå ikke fullt ut fått vise hva de duger til, men i lys av at de fleste land er svært proteksjonistiske hva gjelder egen forsvarsindustri, er sannsynligheten for å lykkes liten. Det sivile markedet er en mulighet, men trolig vil denne type hurtiggående katamaraner bli for kostbare. Nye Sjømissiler (NSM), som er utviklet i et parallelt prosjekt, kan imidlertid ha et fremtidig marked også i utlandet og særlig når det nå utvikles til å kunne rettes mot landmål.

### **Levedyktighet**

Skjoldfartøyet er et høyteknologisk produkt som er meget kostbart å drifte. Besetningen er på 20 personer og årlig driftsbudsjett er estimert til 200 millioner kroner. Det er i tillegg klart at vi snakker om omfattende vedlikehold. Flere informanter har pekt på at behovet for oppgradering nok vil bli enda større enn man har forutsett, både hva gjelder fartøy og våpen. I lys av Forsvarets begrensede totalbudsjett knytter det seg derfor usikkerhet til den økonomiske levedyktigheten av MTB-fartøyene.

Usikkerheten forsterkes ytterligere av at det er tendens mot høy realkostnadsvekst i de aktuelle markedene for vedlikehold og særlig oppgradering, samt at det vil være et spørsmål om hva som er tilgjengelig. Kompetansen for å vedlikeholde og oppgradere MTB-ene ligger først og fremst ved Umoe og en har således gjort seg avhengig av at Umoe overlever i markedet.

Levedyktigheten er også kritisk avhengig av fartøyenes fortsatte relevans i forhold til effekt- og samfunns mål, noe som i sin tur påvirkes av utviklingen i trusselbilde samt teknologisk utvikling ved andre plattformer. Det er stor usikkerhet om dette. Dersom nordområdene blir viktigere for Norge, kan en spekulere i om prosjektets relevans vil øke i fremtiden.

Samtidig sier informanter at Sjøforsvaret er i en presset personellsituasjon og at de allerede har en stor oppgave i å drifte fregattene på en så god måte som mulig. Noen argumenterer derfor for at Forsvaret bør konsentrere seg om dette.

Alt i alt er et høyst sannsynlig scenario at Forsvaret velger å fase ut Skjold-MTB-ene lenge før levetiden på 40 år er utløpt.

### **Samfunnsøkonomisk lønnsomhet**

Evalueringen inneholder ingen egen samfunnsøkonomisk analyse, da man har vurdert dette til å være vanskelig og kanskje også lite interessant, siden nyttesiden i svært liten grad kan kvantifiseres. Men ut fra en samlet vurdering, og spesielt usikkerheten knyttet til hvilken rolle MTB-ene vil ha i den samlede forsvarsstrukturen i årene som kommer, anslår man at lønnsomheten vil være begrenset.

### **Oppsummert om prosjektets vellykkethet**

Tabellen under gir en oversikt over de viktigste (mest valide) indikatorene som ble benyttet for å belyse hvert av evalueringskriteriene.

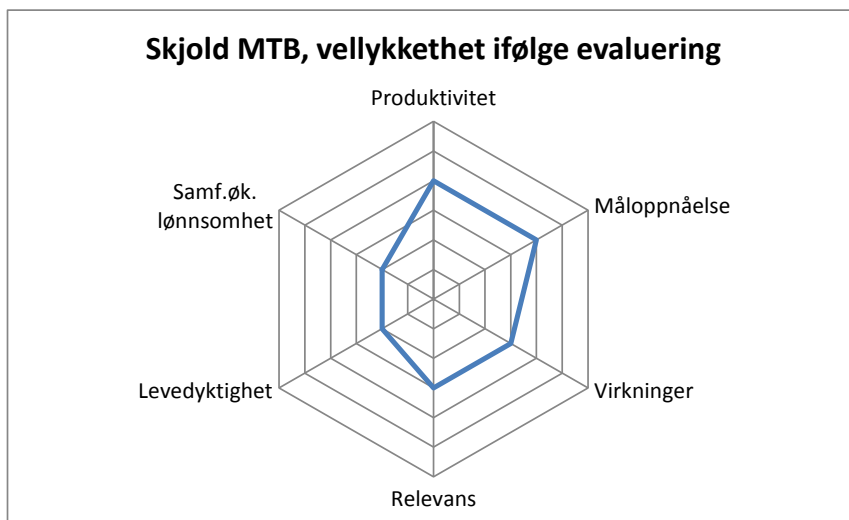
**Tabell 3.4.** Innkjøp av MTB klasse Skjold til Sjøforsvaret. Indikatorer

<b>Viktigste indikatorer</b>	
<b>Produktivitet</b>	<ul style="list-style-type: none"> <li>– Ytelse, funksjonalitet og kvalitet, basert på intervju</li> <li>– Tidsmål, status i dag vurdert opp mot plan</li> <li>– Kostnadsmål, vurdert ved sluttkostnad mot styrings- og kostnadsramme fra KS2</li> <li>– Egen analyse av gjennomføringen</li> </ul>
<b>Måloppnåelse</b>	<ul style="list-style-type: none"> <li>– Forsvarsmessige mål vurdert ved intervju</li> <li>– Næringspolitiske mål (virkninger hos leverandørene) vurdert ved intervju</li> </ul>
<b>Virkninger</b>	<ul style="list-style-type: none"> <li>– Virkninger for andre plattformer, vurdert ved intervju</li> <li>– Salg av produkter og teknologi til andre markeder, intervju</li> <li>– Lokale og regionale virkninger, rapporter og intervju</li> </ul>
<b>Relevans</b>	<ul style="list-style-type: none"> <li>– Forsvarsmessig relevans basert på analyse av dokumenter, herunder forsvarsstudier og medieklipp</li> <li>– Markedsmessig/teknologisk relevant, vurdert ved intervju</li> </ul>
<b>Levedyktighet</b>	<ul style="list-style-type: none"> <li>– Driftskostnader, basert på intervju</li> <li>– Fortsatt forsvarsmessig relevans, se over</li> </ul>
<b>Samf.øk. lønnsomhet</b>	<ul style="list-style-type: none"> <li>– Ikke gjort</li> </ul>

Prosjektet er i dag i overgangen fra gjennomføringsfasen til driftsfasen. Det betyr at evalueringen som er gjennomført får en noe ”foreløpig” karakter.

Evaluator finner at det mest interessante å påpeke fra denne evalueringen er hvordan ulike informanter og ulike analyseperspektiver, gir helt forskjellige svar på om prosjektet er vellykket eller ikke. Produktiviteten i gjennomføringen er vurdert å ha vært rimelig god, de forsvarsmessige målene vil mest sannsynlig nås, og virkningene for sysselsettingen i verftsindustrien er trolig positiv, dog usikker på lang sikt. På den annen side stilles det grunnleggende spørsmål ved både relevans, levedyktighet og samfunnsøkonomisk lønnsomhet i forhold til de oppgavene MTB-ene skal løse.

Diagrammet nedenfor gir en oppsummering av evaluators vurdering for de fem evalueringskriteriene som er besvart, samt en skjønsmessig vurdering av samfunnsøkonomisk lønnsomhet.



**Figur 3.4.** Innkjøp av MTB klasse Skjold til Sjøforsvaret. Evalueringsteamets samlede vurdering av prosjektets vellykkethet

Denne evalueringen er basert på et omfattende materiale av dokumenter, observasjon, samtaler med alle sentrale parter, og flere runder med oppfølging og verifisering av informasjonen i etterkant. Derimot har en hatt lite kvantitativ informasjon og få «objektive» kilder, en har først og fremst benyttet intervju med et bredt spekter av interessenter. Teamet bestod av tre personer og teamleder har omfattende evalueringskompetanse og erfaring med den målorienterte modellen. Oppdragsgivers vurdering er at en i dette tilfellet har benyttet metodikken for målorientert evaluering på en god måte, og evalueringen fremstår som troverdig selv om det i prinsippet innebærer en risiko å kun basere seg på informanternes informasjon og vurderinger. Man har imidlertid ikke svart ut den samfunnsøkonomiske analysen.



---

## 4. Drøfting av lærdommer fra evalueringene

I dette kapitlet oppsummeres og drøftes de viktigste erfaringene og lærdommene fra gjennomføringen av pilotevalueringene. Dette bygger deretter opp under vårt forslag til et standardisert opplegg for evaluering som ikke er for ressurskrevende. Det skal sikre at en allikevel får dekket inn de mest sentrale problemstillingene når KS-prosjektene skal evalueres i årene som kommer. Dette presenteres i kapittel 5.

### 4.1. Generelle erfaringer fra pilotevalueringene

Selv om en benyttet et felles overordnet evalueringsopplegg, valgte de fire teamene som nevnt noe ulike tilnærminger. Vektleggingen av evalueringskriteriene var noe forskjellig, og det samme var balansen mellom kvalitativ og kvantitativ tilnærming. Dette har trolig sammenheng med ulik faglig bakgrunn hos evaluatorene, som varierte fra ingeniører til samfunnsøkonomer, statsvitere og sosiologer. Noen kjente modellen for målorientert evaluering best og andre var mest komfortabel med samfunnsøkonomisk analyse. Det siste gjaldt naturlig nok de av evaluatorene som var samfunnsøkonomer, og særlig de som i hovedsak har erfaring fra transportsektoren, hvor nytte-kostnadsforholdet i prosjektene beregnes rutinemessig. I ettertid ser vi at to av teamene var for homogene. I tillegg var prosjektene ulike, og evaluator hadde ulik tilgang på informasjon.

Teamene møtte mange av de samme utfordringene. Flere fant det krevende å holde et tilstrekkelig fokus på de strategiske evalueringskriteriene, til tross for at man skulle ha høy bevissthet på nettopp dette. Flere fant det også krevende å sortere de ulike delene av analysen på ”riktig” evalueringskriterium, samt finne den samfunnsøkonomiske analysens plass innen rammen av målorientert evaluering. Dette til tross for at teamene hadde fått en innføring i felles metode ved oppstart av studien. En lærdom er derfor at det åpenbart er behov for en klar og forståelig avgrensning mellom evalueringskriteriene, og at det gis en grundig innføring og veiledning i metodikken for alle som skal gjennomføre evalueringene.

En annen felles utfordring var at referansegrunnlaget for prosjektene i mange tilfeller var for mangelfullt til å kunne brukes som grunnlag for evaluering. Målstruktur og prosjektlogikk måtte helt eller delvis reetableres av evaluator, nullalternativet var delvis uklart etc. I en av evalueringene, Asker-Sandvika, hadde en betydelige problemer med å få ut helt nødvendige data både om før- og ettersituasjonen. Dette kan gjelde for jernbaneprosjekter mer generelt. I forsvarsmateriellprosjekter har en den utfordring at ikke all informasjon er åpent

tilgjengelig – og heller ikke kan offentliggjøres i en evalueringsrapport, noe en fikk oppleve i Skjold-evalueringen.

Når det gjelder selve gjennomføringen av studien, var det vår hensikt at de fire pilotevalueringene skulle skje parallelt, slik at vi fikk en felles prosess hvor utfordringer knyttet til den enkelte fase var relevante for alle samtidig. Dette fungerte bare delvis. Det oppstod ulike utfordringer som gjorde at evalueringene for en stor del befant seg i ulike faser. Én evaluering hadde oppstartsproblemer, mens andre ble skjøvet ut i tid. Dette skyldtes dels uforutsette hendelser (for eksempel problem med sikkerhetsklarering i tilfellet Skjold-klasse MTB-er), og dels undervurdering av arbeidsinnsatsen og manglende prioritering. Resultatet ble da at teamene måtte jobbe noe mer isolert i slutfasen enn det som ville vært optimalt. Det er vanskelig å si om en ellers ville endt opp med mer like evalueringsopplegg og tilnærming.

Da evalueringsteamene selv oppsummerte sine erfaringer med prosessen, var de viktigste tilbakemeldingene følgende:

- Valget av indikatorer og datainnsamlings- og analysemetoder må være fleksibelt og tilpasses den enkelte evaluering. Det er imidlertid viktig at evalueringsteamet tidlig spesifiserer et nokså detaljert opplegg. Det er også viktig å være tidlig i gang med tentative konklusjoner for å finne ut hva man mangler.
- Konsentrert feltarbeid og observasjon bør gjennomføres på et relativt tidlig tidspunkt
- Skriveprosessen bør starte raskt slik at en får en periode med modning frem til leveranse

Som et resultat av slike utfordringer, var resultatet av evalueringene noe ujevnt og læringsutbyttet imellom dem i prosessen ble begrenset. Særlig gjaldt det i hvilken grad en besvarte spørsmålene om prosjektenes vellykkethet både operasjonelt, taktisk og strategisk.

Vår vurdering er at evalueringsteamets grunnlag for å foreta en samlet vurdering av hvert enkelt prosjekt i noen tilfeller er mangelfullt og i noen tilfeller tvilsomt. I tillegg er det vanskelig og kanskje misvisende å sammenstille ulike evaluators vurderinger av ulike prosjekter, uten at en har foretatt en kalibrering av hvordan de bruker vurderingsskalaen.

Ettersom dette er en metaevaluering har vi i tabell 4.1 tatt dette i betraktning og på grunnlag av foreliggende informasjon og styrke og svakheter ved denne foretatt en skjønsmessig og grov vurdering av hvordan vi mener de fire prosjektene scorer operasjonelt, taktisk og strategisk. Det kan åpenbart oppfattes som en overprøving av evalueringsteamenes konklusjoner, og det må presiseres at det hefter usikkerhet

ved disse vurderingene. Tabellen er tatt med først og fremst for å gi et eksempel på hvordan resultatene fra slike evalueringer på et overordnet nivå kan illustreres.

**Tabell 4.1.** Pilotstudiens samlede vurdering av prosjektenes vellykkethet basert på resultatet fra de fire evalueringene

Evaluerings-kriterium	Kontrollområde Svinesund	Dobbeltspor Sandvika - Asker	E18 Momarken - Sekkelsten	MTB klasse Skjold til Sjøforsvaret
Operasjonell vellykkethet				
Taktisk vellykkethet				
Strategisk vellykkethet				
Samlet vurdering				

Samlet sett har både gjennomføring og resultatene fra pilotevalueringene vært svært lærerikt og interessant, både for forskerne og de eksterne evaluatorene, men også for denne metaevalueringen. Resten av kapitlet tar for seg de viktigste erfaringene og problemstillingene en møtte underveis.

## 4.2. Tidspunkt for ex post evaluering

Tre av prosjektene i studien var i normal drift på evalueringstidspunktet og har i prinsippet vært det mer enn tilstrekkelig lenge (ca. 5 år) til at resultatene kan vurderes både i et operasjonelt og taktisk perspektiv. En evaluator påpeker at dersom en hadde ventet lengre ville en kanskje ikke finne informanter som har prosjektet friskt i minne og dette kunne vanskeliggjort datainnsamlingen.

Noen vil kanskje hevde at det strategiske perspektivet krever at en går enda lenger ut i prosjektets levetid: Det endelige svaret på prosjektets virkninger, relevans og levedyktighet får man strengt tatt ikke før prosjektet er avsluttet. Men det er normalt verken hensiktsmessig eller nødvendig å vente så lenge. En kan vanligvis si mye om et prosjekts strategiske vellykkethet allerede etter 3 - 5 år, se Andersen m.fl. (2007). Som en kuriositet kan nevnes at en studie av historiske prosjekter der en har foretatt en «evaluering» av ti prosjekter i et virkelig langt historisk perspektiv på opptil flere hundre år, bekrefter at ganske sikre indikasjoner på hvor det bærer hen viser seg allerede etter ganske få år, se Samset (2012). En evaluator påpeker at en i et læringsperspektiv ikke bare er opptatt av å finne ut *om* prosjektet var strategisk vellykket, men også av hvordan *forbedre de strategiske vurderingene* i tidligfase. En trenger ikke vente mange år før en får indikasjoner på om en hadde

et godt beslutningsunderlag i tidligfasen eller om en manglet informasjon, hadde feil prognoser og metoder, dårlige mål, burde vurdert andre alternativer etc.

Svinesund kontrollområde hadde stått ferdig og vært i drift i syv år da evalueringen startet, og evaluator opplevde dette som et egnet tidspunkt for evaluering. Vi slutter oss til denne vurderingen. Effektene var betinget av et parallelt veiprojekt, men dette var også ferdig innen samme tidspunkt.

For begge de to samferdselsprosjektene var det en utfordring at evalueringsobjektet kun var en parsell i en større utbygging, og hvor samlet effekt strengt tatt ikke kan evalueres før etter at alle parsellene er ferdigstilt. Dette gir isolert sett et argument for å vente enda lenger med å evaluere. I tilfellet Momarken-Sekkelsten drøfter evaluator om en kanskje burde ventet til den ene tilstøtende parsellen var ferdig i 2014, da nytten i stor grad avhenger av denne. Samtidig ville en da møte nye utfordringer. For eksempel kan det da bli enda vanskeligere å identifisere noen kausal sammenheng mellom prosjekt og observerte tilstandsendringer. I jernbaneprojektet er dessuten siste parsell nå lagt på is på ubestemt tid, og dersom den blir realisert vil den trolig bli sett i sammenheng med en annen utbygging. Det valgte evalueringstidspunktet syntes derfor likevel å være hensiktsmessig.

Det siste prosjektet, bygging av Skjold-klasse MTB-er, er et særtilfelle, da de to siste fartøyene ennå ikke er levert, selv om de fire første har vært i drift i flere år. Evaluator rapporterer i dette tilfellet at det har vært en klar begrensning for evalueringen at prosjektet ikke var avsluttet på evalueringstidspunktet. Ideelt burde en her ha ventet noe lenger, i det minste til en hadde oppnådd noe mer erfaring med driften av båtene.

### ***Konklusjon***

Erfaringen generelt er at evaluering noen år ut i driftsfasen er passende for ex post evaluering av KS-prosjekter. Hva som er et optimalt tidspunkt må uansett vurderes fra prosjekt til prosjekt, avhengig av hvilke typer av effekter det skal generere og hva annet disse effektene avhenger av.

## **4.3. Avgrensning – avhengighet mellom prosjekter**

Denne problemstillingen henger i noen grad sammen med den forrige. De færreste prosjekter kan betraktes som isolerte tiltak. Prosjektet kan være del av et større program eller en prosjektpakke, eller det kan være et følgeprosjekt av et annet tiltak, eventuelt selv lede til sentrale følgeprosjekter. Uansett vil det ofte eksistere et komplekst sett av andre virkemidler med samme eller liknende mål som det prosjektet vi skal evaluere.

Dette var en utfordring i de to samferdselsprosjektene, hvor prosjektet kun var delprosjekt i en større utbygging og hvor nytten av delprosjektet var nært knyttet til nytten av hele prosjektet. Særlig ga evaluator i evalueringen av Momarken-Sekkelsten uttrykk for at denne avgrensningen var problematisk. Statens Vegvesen hadde antydnet at parsellinndelingen var gjort ut fra hensyn som størrelsen på byggherreorganisasjon, planleggingsressurser og årlige budsjetter, og ikke ut fra hvilke mål de ulike parsellene skulle nå. Evaluatorene valgte likevel, i begge tilfeller, å evaluere delprosjektet isolert, først og fremst på grunn av det lange tidsspennet mellom parsellene og usikkerheten om tidspunkt for endelig ferdigstillelse. Fordelen er som nevnt at evalueringen blir håndterbar. Den viktigste ulempen er at det blir krevende å finne et godt sammenlikningsgrunnlag, samt at vurderingene av taktisk og strategisk vellykkethet ikke vil være reell men basert på mange usikre antakelser. Begge evaluatorene forsøkte å løse dette ved å også se delprosjektet i sammenheng med hovedprosjektet.

Avhengighet av andre prosjekter var også i høyeste grad en utfordring for Kontrollområde Svinesund, som var et direkte følgeprosjekt av en samferdselsinvestering (ny bro over Svinesund og E6 i ny trasé). Det faktum at de to prosjektene ble gjennomført samtidig kunne tale for å evaluere dem samlet. Evaluator har likevel argumentert for å se på kontrollområdet alene, da det gir mulighet til å studere effekten av det nye konseptet med samlokalisering av de tre etatene, uavhengig av effekten av ny vei og bro. (Prosjektmålene handlet om samlokaliseringen). Evaluator har også her sørget for å se det som skjer på kontrollområdet i sammenheng med veg- og broprosjektet der dette var relevant, først og fremst endringene i trafikkvolum og bruk av alternative transportveier.

Også ved evalueringen av Skjold-fartøyene melder evaluator at avgrensningen av evalueringsobjektet representerte noe av en utfordring. Man tok utgangspunkt i prosjektets nokså snevre avgrensning av prosjekt P3600 Skjold Serie, som omfattet bygging av fem seriefartøyer. Men dette var en videreføring av et eget prosjekt som var å utvikle et forseriefartøy som var ferdig i 1999. Videre var anskaffelsen av Skjold MTBene sammen med anskaffelsen av fregattene grunnlaget for utvikling og leveranse av nye sjømissiler (NSM). Helst burde alle disse tre prosjektene vært vurdert sammen. På den annen side er oppsplittingen i prosjekter en måte å gjøre gjennomføringen av evalueringen praktisk håndterbar på. Evaluator har i dette tilfellet løst det ved å ha ett øye på de tilknyttede prosjektene og drøfte resultatene i sammenheng med disse der det faller naturlig.

### **Konklusjon**

Erfaringen er at avgrensningen i det store og hele må være det enkelte prosjektet, først og fremst begrunnet med å redusere arbeidsomfanget i evalueringen. Men prosjektet må alltid studeres i sin kontekst, noe som medfører at man også må se det i sammenheng med relaterte tiltak og eventuelle følgeprosjekter. Der prosjektet er en parsell i en større utbygging gir det begrenset mening å projisere antatt effekt

for helheten ned på den parsellen som studeres. En bør trolig heller avgrense vurderingene til det som gjelder det faktiske og nåtidige, og påpeke begrensningene som ligger i å evalueres isolerte elementer av en helhet. Den fullstendige effekten av hele prosjektet, kan allikevel ikke gjøres før den dagen det er ferdig og alle deler er kommet i drift.

#### **4.4. Metodikken: Målorientert evaluering og samfunnsøkonomisk analyse**

Et viktig mål med studien var å teste valget av evalueringsmodell, det vil si målorientert evaluering (med fem evalueringskriterier) i kombinasjon med samfunnsøkonomisk lønnsomhet som det sjetteste kriteriet. Ville de to tilnærmingene utfylle hverandre på en god måte, eller ville de medføre unødig dobbeltarbeid? Erfaringene tyder på at modellene i starten ble oppfattet til dels som konkurrerende fremstillinger av det samme. Teamene valgte som nevnt å legge noe ulik vekt på de to modellene, men alle evalueringsteamene kom etter hvert til at modellene vil være nyttige sammen.

##### ***Målorientert evaluering***

Den målorienterte evalueringsmodellen er generisk, i den forstand at den er generelt anvendbar på alle typer prosesser eller investeringstiltak, i det den bare spesifiserer fem helt generelle, overordnede evalueringskriterier. Den lar seg derfor lett tilpasse investeringsprosjekter, også slike som gjennomføres etter en strategi med mål på tre nivåer. Hensikten med modellen er i samsvar med hva en ønsket i dette tilfellet, nemlig å løfte analysen fra det laveste nivå, det vil si prosjektgjennomføringen, til det overordnede hvor en vurderer nytten av prosjektet i sin kontekst og over tid.

En viktig utfordring oppstår når målene ikke er formelt avklart, eller som en så i flere av evalueringene: at målene enten var trivielle eller urealistiske. Evaluator måtte da forsøke å fortolke målene slik at de var meningsfulle og egnet som referanseverdier i en evaluering. I slike tilfeller må en sikre at legitimiteten opprettholdes. Flere av evaluatorene påpeker det selvfølgelig ved at evaluator ikke skal endre målene. I de tilfellene der justeringer eller omformuleringer er nødvendig må det skje i råd med eierdepartement. De fire pilotevalueringene løste dette på litt ulike vis, men alle greide å etablere rammer for å vurdere måloppnåelse på en slik måte at det ga et realistisk signal om hvorvidt prosjektet var vellykket eller ikke.

Alle evalueringsteamene opplevde at det var betydelige mengder data tilgjengelig for å vurdere prosjektgjennomføringen, men langt mindre faktainformasjon for å vurdere de strategiske aspektene. Det er da lett å gå i den fellen som er vanlig,

nemlig å bruke det meste av ressursene i evalueringen på det minst viktige kriteriet, mens spørsmålet om strategisk vellykkethet blir salderingsposten. Dette var tilfellet i et par av pilotevalueringene. Samtidig gir den målorienterte evalueringsmodellen sikkerhet for å ikke helt tape fokus på den strategiske dimensjonen. Flere av evaluatorene, herunder de som ikke hadde brukt denne modellen før, påpekte dette som den primære nytteverdien ved modellen – å sikre at alle perspektiver skal med.

Det kan likevel ikke underslås at det er krevende å skulle vurdere prosjekters strategiske vellykkethet, spesielt kriteriene «relevans» og «levedyktighet». Det gjelder særlig i de tilfellene hvor det ikke er formulert gode samfunns mål og/eller dersom vurderingen av relevans i stor grad påvirkes av hvilket ståsted man har. Noen av evaluatorene ga uttrykk for at de opplevde det som vanskelig å skulle etablere indikatorer for hva som er «samfunnets beste», da dette ligger utenfor det man som fagspesialist er vant til å arbeide med. Flere hadde først og fremst erfaring med å vurdere samfunnsøkonomisk lønnsomhet, der skal en benytte betalingsvillighetsprinsippet, og analytiker kan i prinsippet forholde seg «nøytralt» til dette. Skjold-evalueringen er imidlertid et godt eksempel på at en greide å formidle ulike parter syn på prosjektets relevans, selv om dette var et kontroversielt prosjekt. I denne evalueringen var vurderingen av nettopp relevans kanskje den aller mest interessante.

Dette forteller oss først og fremst at det er behov for samfunnsfaglig kompetanse, god opplæring og veiledning i bruk av den målorienterte modellen, og da spesielt de strategiske kriteriene. I tillegg er evaluators habilitet selvsagt også svært viktig.

En annen utfordring er at de fem vurderingskriteriene må forstås og operasjonaliseres likt av alle impliserte parter. Selv om kriteriene i og for seg er uavhengige av hverandre, er de uttrykk for fem ulike innsynsvinkler på det samme prosjektet. Det kan derfor oppstå sorteringsproblemer i forhold til foreliggende informasjon. En bør unngå overlapp mellom datagrunnlaget og resonnementene for de fem vurderingskriteriene. Hvordan skal en for eksempel sortere mellom ”måloppnåelse” og ”virkninger” når en viktig del av motivet for tiltaket aldri formelt ble nedfelt som vedtatte mål (jevnfør Skjold MTB-ene og næringspolitiske mål).

### ***Samfunnsøkonomisk analyse***

Slik Finansdepartementet og DFØ definerer samfunnsøkonomisk analyse er dette en metode for å klarlegge og synliggjøre alle konsekvenser av offentlige tiltak. Det er ikke et krav at konsekvensene må kvantifiseres og verdsettes for å kunne inngå i analysen. Det skilles mellom nytte-kostnadsanalyse, kostnads-effektivitetsanalyse og kostnads-virkningsanalyse.

I de to samferdselsprosjektene valgte evaluatorene å satse på en nytte-kostnadsanalyse. I tilfellet Momarken-Sekkelsten fikk man ulykkes- og trafikkdata og kunne gjøre oppdaterte nåverdiberegninger. Dette indikerte at både parsellen og hovedprosjektet lå an til å bli langt mer lønnsomt enn forutsatt. I tilfellet Sandvika-Asker hadde evaluator derimot problemer med datatilgangen og konkluderte med at det ikke var mulig å gjennomføre nytte-kostnadsberegningen ex post. Man gjorde i stedet noen drøftinger basert på det begrensede tallmaterialet som var tilgjengelig, og argumenterte – på svært usikkert grunnlag – med at lønnsomheten trolig var overvurdert i forkant.

Det er god grunn til å tro at det i begge prosjektene også er andre, ikke-verdsette virkninger av betydning som burde vært identifisert og drøftet systematisk som del av den samfunnsøkonomiske analysen (for eksempel i en +/-tabell). Dessuten vil en slik lønnsomhetsanalyse som gjøres kun av en parsell, og hvor tallene for mesteparten av levetiden i tillegg fortsatt er hypotetisk, være beheftet med stor usikkerhet.

I Svinesund-evalueringen valgte evaluator å gjennomføre en kostnads-virkningsanalyse. Denne ga en bred og systematisk oversikt over virkningene. Få av virkningene kunne imidlertid kvantifiseres og verdsettes med særlig grad av sikkerhet, og evaluator var derfor meget forsiktig med å konkludere om lønnsomheten. Man antyder at analysen gir begrenset tilleggsnytte ut over de vurderingene som alt er gjort under den målorienterte modellen. Vi stiller spørsmål ved om evaluator kunne gått lengre i å konkludere med hensyn til lønnsomheten i dette tilfellet. På samme måte som for de fem kriteriene i den målorienterte modellen bør evaluator forsøke å gjøre en samlet vurdering basert på den informasjon som foreligger. I tilfeller hvor en står overfor politiserte eller etiske vanskelige spørsmål kan dette være vanskelig, men det gjaldt ikke her.

I evalueringen av Skjold gjorde man ingen samfunnsøkonomisk analyse i det hele tatt. Dette begrunnes delvis med at en ser det som vanskelig å verdsette nyttesiden, samtidig som evaluator også antyder at analysen ikke ville gi særlig stor tilleggsverdi. Men det er heller ikke gjort noe forsøk på å sette opp f.eks. en kostnads-effektivitets- eller kostnads-virkningsanalyse. Trolig er mangel på kompetanse i evalueringsteamet, kanskje også tidspress, medvirkende årsaker til dette. Evalueringsrapporten påpeker selv at det er fullt mulig å gjøre samfunnsøkonomiske analyser også for forsvarsmateriellprosjekter<sup>15</sup>.

---

<sup>15</sup> Viser til at Forsvarsdepartementet siden 2010 har hatt en egen Veileder i samfunnsøkonomiske analyser for investeringsvirksomheten i forsvarssektoren.



Å få klarlagt den samfunnsøkonomiske lønnsomheten av offentlige investeringer ex post er normalt ønskelig fordi dette er et sentralt beslutningskriterium på et tidlig tidspunkt og fordi det ofte legges ned et betydelig arbeid i å utrede dette ex ante. Selv om et prosjekt viser seg å bli ulønnsomt, betyr ikke dette nødvendigvis at man tok en feil beslutning. Mer generelt er det to grunner til at det kan være nyttig å sammenstille informasjonen på en slik måte at en kan vurdere prosjektets samfunnsøkonomisk lønnsomhet:

- 1) vi kan svare på om prosjektet ble *lønnsomt eller ikke*
- 2) vi kan etterprøve de samfunnsøkonomiske vurderingene (kvalitative og kvantitative) som var gjort ex ante. Da kan vi svare på om lønnsomhetsvurderingen som ble lagt til grunn var *realistisk eller ikke*.

Selv om punkt 1) skulle være lite meningsfullt (slik man antyder i to av prosjektene), så kan punkt 2) like fullt være viktig og nyttig. Men nr. 2) betinger at det finnes en samfunnsøkonomisk analyse ex ante som sammenlikningsgrunnlag, og at denne er transparent nok til at evaluator forstår hva den er basert på. I den grad det der er gitt helt konkrete forutsetninger og prognoser om relevante størrelser som påvirker lønnsomheten (eks. befolknings- og inntektsutviklingen for transportprosjekter) er det åpenbart nyttig å vite om man treffer, da dette kan forbedre senere analyser.

Ingen av prosjektene i pilotevalueringene hadde vært igjennom KS1, og for de fleste forelå ingen god og fullstendig samfunnsøkonomisk analyse i forkant. En hadde dermed ikke noe godt sammenlikningsgrunnlag. I slike tilfeller blir analysen av lønnsomhet ex post nødvendigvis mer krevende, noe som ble kommentert også i dette tilfellet (kanskje blir analysen også mindre interessant siden punkt 2) over ikke er aktuelt). I senere evalueringer av prosjekter som har vært gjennom KS1 vil en trolig ikke ha den samme utfordringen med referansegrunnlaget.

Som nevnt i kapittel 2 er erfaringen generelt at det ofte er vanskelig å gjennomføre en komplett samfunnsøkonomisk analyse av prosjekter, og at slike ofte har en tendens til å bli ufullstendige. Dette har vi også erfart også i dette tilfellet. Vi tror at en medvirkende årsak kan være uklarhet om definisjonen av samfunnsøkonomisk analyse, og dermed hvilket ambisjonsnivå evaluator legger til grunn med hensyn til verdsetting. Dersom en tolker samfunnsøkonomisk analyse som synonymt med nytte-kostnadsanalyse, vil en ofte oppleve at en ikke har de nødvendige data – og dersom analysen gjøres vil den kunne komme ut som noe snever og ufullstendig. Her er det forskjeller mellom prosjekter og sektorer, men det er uansett fornuftig å ikke ha for store ambisjoner med hensyn til detaljnivå og verdsetting i en slik overordnet evaluering. Derimot bør evaluator ha som ambisjon å sammenstille informasjonen på en slik måte at det kan gjøres en samlet vurdering av hvorvidt prosjektet fremstår som lønnsomt eller ikke.

### ***Samtidig bruk av modellene***

Som drøftet i kapittel 2 dekker modellen for målorientert evaluering i prinsippet alle sider av prosjektets vellykkethet. Når vi foreslår å supplere med samfunnsøkonomisk lønnsomhet, er det både fordi det i dag er krav om at dette bregnes *ex ante* for store investeringsprosjekter, og fordi en kan argumentere for at samfunnsøkonomisk effektivitet alltid er et mål ved statlige prosjekter (jevnfør Økonomiregelverket) uansett hva målene ellers består i. Med utgangspunkt i det som gjøres i en målorientert evaluering, bør en i prinsippet kunne gjøre en analyse av samfunnsøkonomisk lønnsomhet relativt enkelt, da mye av dataunderlaget allerede vil være innsamlet under de andre fem evalueringskriteriene.

Motsatt kan det også argumenteres for at en korrekt gjennomført samfunnsøkonomisk analyse langt på vei er tilstrekkelig for å vurdere prosjektets vellykkethet, og at den målorienterte modellen dermed kan betraktes som et supplement. Det kan argumenteres for at dersom en besvarer den samfunnsøkonomiske analysen *først*, vil en ha en god oversikt over sentrale nytte- og kostnadsvirkninger på ulike nivå og i ulike tidsperspektiv. Den supplerende analysen for å besvare den målorienterte evalueringen kan dermed gjøres mindre omfattende. Men igjen, dette krever en god og tilstrekkelig bred samfunnsøkonomisk analyse (se Finansdepartementets og DFØs veiledning). I vårt tilfelle kan ingen av evaluatorene sies å ha gjennomført dette, og det fungerte derfor best å bruke motsatt rekkefølge, altså den samfunnsøkonomiske analysen som supplement til målorientert modell.

### ***Konklusjon***

Modellene er altså dels komplementære, men er også dels svar på ulike spørsmål, og erfaringen fra pilotevalueringene er at kombinasjonen av dem vil være nyttig i en evaluering. Metodene bygger på til dels det samme informasjonsunderlaget og dette reduserer faren for at det blir «dobbeltarbeid» å benytte begge. Tvert imot kan en si at de to ulike måtene å sammenstille informasjon på representerer en form for metodemessig triangulering. Likevel ser vi at det kan være nyttig å tillate en viss fleksibilitet når det gjelder hvor grundig en skal besvare hver av de to modellene. Som antydnet over, kan dette blant annet avhenge av hvilken rekkefølge en velger på evalueringskriteriene, noe som i sin tur vil påvirkes av sektor, tradisjoner og datatilgang. Det er likevel essensielt at alle de seks evalueringskriteriene blir besvart som del av evalueringen.

En lærdom er uansett at evalueringsteamet bør ha kompetanse både på samfunnsøkonomisk analyse og målorientert evaluering. Nye evaluatorene bør få en gjennomgang og eksemplifisering av bruken av begge metoder – og kombinasjonen av dem. Kanskje bør det både utvikles en veileder eller kort håndbok og et kursopplegg.

## 4.5. Mangler ved prosjektlogikk og målstruktur

Et viktig element i den felles metodikken som ble lagt til grunn (jevnfør kap. 2) er at evaluator skal starte med å vurdere prosjektets logikk og målhierarki. Utfordringen er at det formelt avtalte kan være mangelfullt. Teamet må også forstå prosjektets logikk før en kan starte evalueringen. Det har vist seg å være nyttig men også krevende, og i flere tilfeller var det nødvendig som nevnt i kapittel 4.4 å etablere en helt ny logikk.

Generelt hadde alle prosjektene mangelfull målstruktur og logikk i utgangspunktet. Dette er problemer som ventes å være mindre i fremtiden når en skal evaluere prosjekter som har vært igjennom KS1, dog er det neppe realistisk at en helt slipper en slik kritisk vurdering av sammenlikningsgrunnlaget.

Evaluator av Svinesund-prosjektet gikk lengst i å omformulere de eksisterende målene. En fant rett og slett at de eksisterende målene var misvisende, motsetningsfulle, ulogiske eller urealistiske og derfor ikke egnet som grunnlag for evaluering. Også to av de andre teamene måtte gå til underlagsdokumentene å trekke ut hva den egentlige hensikten med prosjektet var. Her vil det være en fordel å konsultere sentrale personer i berørte departement og etater, for å få en anerkjennelse og forankring av det en vil foreslå som utgangspunkt for evalueringen. På den annen side kan det være ulike syn på målene avhengig av hvem man spør. En bør derfor holde seg strengt til skriftlig dokumentasjon fra tidligfasen som referanse og klargjøre hvorfor det formelt avtalte ikke kan legges til grunn for evalueringen. Det er for eksempel åpenbart at dersom det formelt avtalte målet er helt urealistisk vil evaluator måtte konkludere at prosjektet har feilet i forhold til realiseringen av målet.

### **Konklusjon**

En evaluering er en formell, overordnet analyse. Prinsippet er at det er det formelt avtalte som skal ligge til grunn for vurderingen av måloppnåelse. Erfaringen er at det formelt avtalte, også det som har passert Stortinget og regjeringen, ikke alltid reflekterer virkeligheten i tilstrekkelig grad. Det kan være urealistiske eller trivielle mål, feilaktige estimater og antakelser, etc. Praksis må i disse tilfellene være at evaluator påpeker slike forhold, og legger til grunn forutsetninger som bygger på logikk, sannsynlighets- eller rimelighetsvurderinger der det er nødvendig. Videre at evaluator begrunner anvendelsen av disse eksplisitt. For mer informasjon se Samset (2008), kapittel 11 og 12.

## 4.6. Informasjonsgrunnlaget og behovet for triangulering

De to samferdselsevalueringene har fokusert mye på kvantitative indikatorer og man har hovedsakelig søkt etter ”objektive” og faktabaserte data i form av tall eller informasjon fra skriftlige kilder. En har vært grunnleggende skeptisk til å la ulike

parters subjektive vurderinger få for stor plass i evalueringen. Faktabasert informasjon må stå i en særstilling ifølge disse evaluatorene, ikke minst i lys av utfordringene med å sikre legitimitet for evaluators vurderinger. I disse tilfellene gjelder også det at samferdselsprosjekter nettopp er eksempler på at det finnes – eller burde finnes – et betydelig tallmateriale som grunnlag for evalueringen, for eksempel om trafikkstrømmer. De to andre evaluatorene har i langt større grad benyttet kvalitativ informasjon fra dokumentanalyse, observasjon og intervjuer, noe som blant annet har sammenheng med at tallgrunnlaget er begrenset. Skjold-evalueringen er i særstilling ved å ha brukt intervju som hovedkilde til informasjon.

Som nevnt i kapittel 2 trenger vi alltid en kombinasjon av kvantitativ og kvalitativ informasjon. Dette er også de fire evaluatorenes anbefaling etter gjennomførte pilotevalueringer. Kvalitative vurderinger hjelper oss til å beskrive helheten, mens kvantitativ informasjon gir beskrivelsen presisjon. Vår ofte overdrevne trang til å vektlegge verdien av tallbasert informasjon gjør at det er lett å overse de viktigste fordelene ved å bruke tekstlig informasjon:

- Det er dette som er vårt viktigste grunnlag for kommunikasjon
- Denne typen informasjon er grunnleggende nødvendig for å gi en helhetsbeskrivelse av komplekse forhold der mange påvirkningsfaktorer inngår
- Mye informasjon kan ganske enkelt ikke kvantifiseres
- Dersom den tallbaserte informasjonen ikke er pålitelig, kan det være mer hensiktsmessig å benytte kvalitative uttrykk inntil pålitelig informasjon kan skaffes
- Det er et kjent fenomen at et stort tallbasert datamateriale eller data på høyt aggregeringsnivå ofte kan bidra til å tilsløre heller enn å klargjøre et forhold.

På den annen side, å bruke intervju og befaring som viktigste informasjonskilde gir mulighet for synsing fra både intervjuobjektene og evaluator, og evalueringen vil kunne kritiseres for manglende presisjon og legitimitet.

Et relevant spørsmål er hva en bør gjøre i tilfeller hvor faktabasert informasjon enten ikke finnes eller ikke kan brukes fordi den er konfidensiell. Dette var tilfelle i Skjold-evalueringen, og evaluator påpeker selv dette som en svakhet ved evalueringen – spesielt ettersom prosjektet av mange anses som kontroversielt. Når man allikevel i stor grad baserte seg på intervju med sentrale informanter var det ut fra en vurdering av at dette var den eneste muligheten en hadde. Evaluator har dokumentert hvilke spørsmål som ble tatt opp i intervjuene, hvem som ble intervjuet, og det er lagt vekt på stor bredde i sammensetningen av informanter slik at en oppnår en trianguleringseffekt. I tillegg har en supplert med det som finnes av skriftlige kilder til beskrivelse av prosjektet ex ante, ulike forsvarsstudier etc., og

det er gjennomført befarings. Dette har resultert i en evalueringsrapport som vi mener inngir tillit. Likevel kan en diskutere om evaluator burde hatt tilgang for eksempel til studier som Forsvarets forskningsinstitutt har gjort av forsvarsstrukturen og mulige konfliktscenarier, som beskriver de ulike våpenplattformenes plass i forhold til hverandre.

Ved vurderingen av effekter av et tiltak er spørsmålet om kontroll for eksterne faktorer relevant. I evalueringene har man på ulike måter søkt å belyse dette med statistikk og annen informasjon om generelle trender og om prosjektets kontekst, for deretter å gjøre en samlet vurdering av sannsynlig effekt av investeringen. I Svinesund-evalueringen har en for eksempel tatt utgangspunkt i utviklingen i beslag av ulovlige varer på grensen, før og etter nytt kontrollområde. En har da lagt vekt på å dokumentere *andre endringer* som fant sted i samme periode som også kan ha hatt innvirkning på kontrolleffektiviteten (økt bemanning, innføring av elektronisk overvåkingssystem, nye administrative rutiner, annen ny teknologi etc.) og en har sammenliknet med utviklingen i beslag ved *andre kontrollstasjoner* i Norge. Presisjonsnivået er ikke høyt, men gir etter vår vurdering en evaluering av akseptabel kvalitet.

### **Konklusjon**

Erfaringen er at i spørsmålet om kvantitativ eller kvalitativ informasjon og analyse er det ikke snakk om enten eller, men både og. Intervju med sentrale informanter er en av de viktigste kildene en har til informasjon under en evaluering, men ukritisk bruk av slik informasjon kan skape problemer med troverdighet og validitet. Det viktigste virkemiddelet for å motarbeide dette er å triangulere informasjon, det vil si at en sjekker opplysninger fra flere kilder, bruker forskjellig type informasjon for å belyse samme fenomen, etc. Faktabaserte data bør ligge i bunn der slike kan innhentes med rimelig ressursbruk.

## **4.7. Tilgang til eksisterende data**

En evaluering er en begrenset studie både i omfang og dybde. Det er ressurskrevende å samle inn data og budsjettet setter klare begrensninger. Utfordringen er å avgrense og samtidig fokusere på den mest sentrale informasjonen. Tilgang til allerede innsamlet materiale i sektorene, samt generell statistikk, er viktig for at evaluator skal kunne gjøre jobben effektivt. Pilotevalueringene viser at det ikke alltid er vilje eller mulighet til å stille slik informasjon til rådighet. Det er også ofte en utfordring å få tilstrekkelig gode tidsseriedata fordi innsamlingsrutiner, datasystemer osv. er endret over tid.

Evaluator av Asker-Sandvika har påpekt problemet med mangel på data, og store deler av rapporten handler om utfordringene med datatilgang om dette jernbaneprosjektet. Dette gjelder trolig for jernbaneprosjekter mer generelt. Endringer i passasjervolumet og eventuelt mengde gods per strekning er en viktig

indikator på nytten av en jernbaneinvestering. Det er også en variabel som det rutinemessig legges inn forventningsverdier for, når det gjøres samfunnsøkonomiske analyser ex ante. Da virker det urimelig at det ikke skal være mulig å etterprøve effekten i etterkant fordi aktørene som sitter på dataene (her NSB) ikke gir ut slik informasjon. Svenske VTI som har vært evaluator av Asker-Sandvika, påpeker at situasjonen er nokså lik i Sverige, og at det der skyldes en uvilje fra kommersielle operatører mot å offentliggjøre det de anser som konkurransesensitiv informasjon. I Norge, med bare en operatør på persontransporten, burde dette være mulig.

Også i de andre evalueringene hadde man visse utfordringer med å finne frem i eksisterende data. I tilfellet Momarken-Sekkelsten hadde en av deltakerne i evalueringsteamet mangeårig bakgrunn fra Statens Vegvesen og hadde slik sett en betydelig fordel. For eksterne evaluatører vil det trolig være vanskeligere å finne frem til relevante arkiver, systemer og andre kilder.

En minst like viktig erfaring er at det i mange tilfeller finnes lite informasjon om hva nullalternativet består i og hvilke forutsetninger, ikke minst prognoser for ulike variablers videre utvikling, som er lagt til grunn ex ante. Det gjelder særlig for den samfunnsøkonomiske analysen. Når en i tillegg vet at det skjer hyppige oppdateringer og endringer i etatens modellverktøy, er det åpenbart ikke tilstrekkelig for evaluator å få oppgitt størrelsen på en nytte-kostnadsbrøk som ble beregnet ex ante. For å kunne sammenlikne forventet effekt og dermed lønnsomhet med det som ble realisert må en vite hvilke forutsetninger som lå bak, for eksempel om befolkningsutvikling, inntekter, relative priser, kvaliteten på togtilbudet etc. Dette er informasjon som åpenbart burde finnes og gjøres tilgjengelig for evaluator, men det krever at etatene har gode rutiner for å dokumentere underlaget for ex ante-analysen, og å ta vare på slike data. I Sverige har man nylig strammet opp retningslinjene for behandling av trafikkprognoser, blant annet fordi en hadde observert at prognosene ikke alltid ble dokumentert slik at de var tilgjengelige for oppfølging i ettertid (Trafikverket, 2012). Prognosene er i sin tur basert på en rekke forutsetninger, og både disse og selve prognosemodellen må være dokumentert dersom det skal være mulig å vurdere i ettertid.

### **Konklusjon**

Erfaringene fra pilotevalueringene viser at det finnes – eller burde finnes – sentral informasjon som evaluator ikke har fått tilgang til. Det vil være en utfordring i fremtidige evalueringer å sikre at evaluator får tilgang til slik informasjon. Dette er av helt avgjørende betydning for om en skal kunne gjennomføre en riktig og troverdig evaluering av offentlige investeringstiltak. Databehovet for effektiv vurdering bør være rimelig godt definert gjennom gode målformuleringer for tiltaket. På flere sektorområder kan det være behov for klarere regler som støtter opp om innsamling, lagring og tilgjengeliggjøring av data.

## 4.8. Kompetanse og teamsammensetning

### *Uavhengighet*

Erfaringene fra pilotevalueringene bekrefter langt på vei det en vet om hva som er anbefalt praksis vedrørende kompetanse og sammensetning av evalueringsteam. For det første legges det stor vekt på at en skal benytte *uavhengig ekspertise*. Et særtilfelle er læringsevalueringer der en også bør ha med interne medarbeidere for å sikre at læringen tilføres den organisasjonen som er ansvarlig for tiltaket som evalueres.

Situasjonen i pilotevalueringene var spesiell ved at læringsarenaen i tillegg til de ansvarlige departementer og etater også omfatter Concept-programmet. Dette følgeforskningsmiljøet har som oppgave blant annet å samle og analysere resultatene fra evalueringene i årene som kommer med sikte på å trekke ut generisk kunnskap og gi råd til medvirkende og berørte parter på grunnlag av dette. Concept-programmet hadde derfor med minst en av sine egne medarbeidere i hvert team.

Sett under ett ble uavhengigheten sikret ved at alle de ansvarlige evaluatorene var uavhengige i forhold til berørte etater og departementer. Tre team ble ledet av eksterne leverandører på oppdrag fra Concept-programmet, men med forskere fra Concept involvert i mer begrensede roller. En evaluering ble ledet av Concept-programmet internt (Momarken-Sekkelsten). Erfaringene med forskermedvirkningen var gode, det ga vesentlig læringsutbytte både i form av kunnskap om de prosjektene som ble evaluert, om det metodiske opplegget, og om tilrettelegging og styring av evalueringer, noe som vil være nyttig i tiden som kommer. Erfaringene med å benytte eksterne leverandører var også i hovedsak god (Skjold, Svinesund, Asker-Sandvika), ikke minst når det gjaldt fremdriften og det å levere på tid.

### *Evalueringskompetanse*

Evalueringer er overordnede vurderinger av komplekse saksforhold som krever spesialisert *evalueringskompetanse*. Evalueringsfaget og omfanget av evalueringsvirksomheten i samfunnet har etter hvert utviklet seg så mye at vi har fått en egen profesjon av evalueringseksperter.

To av teamene hadde med og ble ledet av personer med betydelig kompetanse og erfaring på dette området, ikke minst når det gjelder den målorienterte evalueringsmetodikken som var valgt (Svinesund og Skjold). Dette kom til uttrykk både i opplegget og gjennomføringen, som i begge tilfellene foretok en bred analyse som særlig gikk grundig inn på de strategiske aspektene av prosjektene. Evalueringenes konklusjoner var i utpreget grad basert på en kombinasjon av foreliggende informasjon og data, selvstendig observasjon og refleksjon, og

intervjuer med alle sentrale aktører og berørte parter. Dette innebærer en betydelig trianguleringseffekt som bidrar til å sikre etterretteligheten i det en er kommet frem til. De andre to teamene (Asker/Sandvika og Momarken/Sekkelsten) manglet evalueringsekspertise men hadde i større grad sektorkompetanse, og kompetanse innenfor samfunnsøkonomisk analyse. De valgte et opplegg med større fokus på de operasjonelle og taktiske aspektene og å tallfeste de kvantifiserbare samfunnsøkonomiske virkningene. Dette har også sammenheng med egenskaper ved prosjektene og tradisjoner i transportsektoren, hvor en normalt vurderer effektene av en investering med utgangspunkt i trafikk tall. Bredere og mer strategiske vurderinger ble i disse tilfellene mer sekundære.

### **Sektorkompetanse**

Når det gjelder *sektor- og prosjektkompetanse* er erfaringene generelt at dette er nyttig, ikke minst for å vite hvilke eksisterende data en bør etterspørre. Men det kan også være avgrensende, særlig for den delen av evalueringen som gjelder det strategiske perspektivet. I en bredere analyse vil prosjektet som sådan til en viss grad ses på som en black-box. Det vi her ønsket å finne ut av var primært hva som gikk inn i og kom ut av boksen - og hvilke implikasjoner det fikk på omgivelsene. Selve gjennomføringen av prosjektet var mindre sentralt å få belyst. Det er en etablert erfaring at resultatet ofte blir mangelfullt dersom en benytter team som *utelukkende* har sektorekspertise, for eksempel leger som evaluerer et sykehusprosjekt eller prosjektledere som evaluerer et prosjekt.

I pilotevalueringene fikk en testet ut begge deler. I Skjold og Svinesund hadde en ingen sektorkompetanse. Det kan være en delforklaring på det som er nevnt over, at en her la stor vekt på å drøfte strategiske spørsmål. Manglende sektorkompetanse later ikke til å ha skapt store problemer for gjennomføring og vurderinger, men en har gitt uttrykk for (Skjold) at det var en utfordring at en ikke hadde den nødvendige kunnskap for å kunne vurdere forholdet mellom Skjold og andre plattformer og andre ulike forsvarsmessige vurderinger.

I samferdselsprosjektene var sektorkompetanse hovedtyngden i teamet. I evalueringen av Momarken-Sekkelsten fikk en detaljerte vurderinger av selve prosjektet og sektorinterne virkninger som trafikkstrømmer etc. Det var et sterkt fokus på de operasjonelle og taktiske aspektene, på bekostning av de strategiske. En burde i større grad lagt vekt på å dokumentere sentrale kontekstuelle forhold ved etter-situasjonen og drøfte strategiske spørsmål knyttet til den typen partiell investering som prosjektene representerer. Erfaringene i dette tilfellet er derfor at det er særdeles viktig å benytte tverrfaglige team, og at disse ikke domineres av sektorekspertise.

### **Samfunnsøkonomisk kompetanse**

Spørsmålet om samfunnsøkonomisk lønnsomhet står sentralt i tilretteleggingen og kvalitetssikringen av KS-prosjekter, og vil av mange oppfattes som den endelige



testen på om et investeringstiltak er vellykket eller ikke. Tre av teamene hadde med *samfunnsøkonomiske ekspertise*, dvs. Svinesund, Momarken-Sekkelsten og Asker-Sandvika. I det første tilfellet fikk en frem en systematisk kartlegging av nyttesiden og behandling av hva som kan karakteriseres som prosjektets samfunnsøkonomiske nytte, men der nyttesiden i stor grad bestod av ikke-prissatte konsekvenser. I samferdselsprosjektene bestod evalueringsteamene av samfunnsøkonomer med mye erfaring fra transportsektoren, og her hadde en større fokus på å kvantifisere og verdsette alle virkninger, slik det er tradisjon for i samfunnsøkonomiske analyser av samferdselsprosjekter. I det ene prosjektet greide en likevel ikke å gjennomføre analysen på grunn av dataproblemer. I den siste evalueringen (Skjold), som fremdeles er i produksjonsfasen ble det ikke foretatt noen samfunnsøkonomisk analyse.

Erfaringen på dette området er at en fullstendig samfunnsøkonomisk analyse i mange tilfeller er krevende men allikevel bør få en sentral plass i fremtidige evalueringer. For å sikre at dette lykkes vil en også her være tjent med faglig bredde i teamet.

### ***Bredde og triangulering***

En siste erfaring som må fremheves er hvor viktig det er at teamet utnytter den muligheten en har til å fremskaffe og triangulere informasjon av forskjellig art og fra forskjellige kilder for å *sikre bredden i perspektivene og validiteten* i det en kommer frem til. I særlig én av evalueringene (Svinesund) lykkes en med dette ved at det ble gjennomført grundig befaring på stedet og intervjuer med involverte og berørte på flere nivåer, kombinert med at en benyttet et bredt tilfang av tilgjengelig statistikk og andre kvantitative data.

Det er viktig at evalueringen ikke blir en ren deskstudie, også selv om en har tilgang til faktabaserte opplysninger og statistikk, og spesielt transportprosjektene kunne i større grad ha supplert med kvalitative datainnsamlingsmetoder.

Det må selvsagt være en balanse mellom tallfestet og kvalitativ informasjon av hensyn til evalueringens troverdighet. I så måte representerer evalueringen av Skjold den andre ytterligheten, der en i stor grad var avhengig av intervjuer for å fange ulike aktørers vurderinger av prosjektet. Dette trekkes også frem som en svakhet av evaluator selv.

### ***Konklusjon***

Konklusjonen på dette blir at i den typen evalueringer av store statlig finansierte investeringsprosjekter som det nå legges opp til, er det viktig at en får belyst alle sentrale sider av ettersituasjonen. En legger vekt på å dokumentere og vurdere de overordnede aspektene som evalueringsmetodikken legger opp til. Det er viktig at en benytter team som (1) ledes av folk med evalueringsekspertise både når det gjelder målorientert evaluering og samfunnsøkonomisk analyse, at de (2)





hovedsakelig representerer uavhengig ekspertise i forhold til involverte departementer og etater, at de (3) er tverrfaglig sammensatt, og at en (4) tillater intern forskermedvirkning der det er hensiktsmessig.

## 4.9. KS-ordningens betydning

En underliggende ambisjon for Concept-programmet som driver følgeforskning på KS-ordningen, er å kunne si noe om KS-ordningens betydning for graden av vellykkethet i prosjektene. Dette spørsmålet har likevel ikke fått veldig mye fokus i de fire pilotevalueringene. Her har en vært opptatt av å vurdere prosjektenes *overordnede* vellykkethet, mens de fire prosjektene kun hadde gjennomgått en kvalitetssikring av kostnadsramme og styringsunderlag (KS2) i forkant.

Kun når det gjelder prosjektenes *produktivitet* har vi derfor noen betraktninger rundt kvalitetssikringens betydning, i lys av at de fire prosjektene har gjennomgått KS2. Tre av prosjektene hadde en sluttkostnad som var innenfor kostnadsrammen, og det kan tyde på at kostnadsestimatene var relativt troverdige og realistiske. I figuren under sammenholdes prosjektenes sluttkostnad også med styringsrammen for prosjektet, som er lavere da kostnadsrammen inneholder en avsetning for usikkerhet. Bare Svinesund-prosjektet ble levert innenfor styringsrammen.

**Tabell 4.2.** Økonomiske rammer og kostnader i de fire prosjektene

Rammer og kostnader (millioner kroner)	Kontrollområde Svinesund	Dobbeltspor Sandvika - Asker	E18 Momarken - Sekkelsten	MTB klasse Skjold til Sjøforsvaret
Sluttkostnad	257	3714	643	5000
Kostnadsramme	305	4165	611	5658
Styringsramme	289	3314	562	n/a
Samlet vurdering				

De usikkerhetsfaktorene som slo ut var stort sett påpekt i KS2-rapportene, selv om også etatene i enkelte tilfeller selv hadde identifisert disse risikoene i forkant. Evaluatør av Skjold-anskaffelsen antyder at kvalitetssikrer ikke fanget opp den mest fundamentale usikkerheten i dette prosjektet, nemlig teknologien og markedet for fremdriftsløsningen, som skulle vise seg å bli et betydelig problem. Derimot var KS2 opptatt av usikkerheten knyttet til kontraktstruktur, hvilket i dette tilfellet ikke ble et problem.

I Asker-Sandvika-prosjektet hadde man sterkt fokus på kostnadskontroll i gjennomføringsfasen, og fulgte blant annet forslaget til kuttliste i KS2-rapporten.

Vi har ikke med dette grunnlag for å si at resultatene var en effekt av KS2. Til det burde vi hatt et langt større utvalg og en kontrollgruppe uten KS2. De gode resultatene hva gjelder sluttkostnader understøtter imidlertid foreløpige resultater fra en pågående studie av kostnadsutviklingen i et større utvalg av KS2-prosjekter, som viser at de fleste prosjektene (ca. 80 prosent) gjennomføres innenfor kostnadsrammen og at sluttkostnaden for alle prosjektene i porteføljen har en symmetrisk spredning om forventningsverdien, som altså ligger et stykke lavere enn kostnadsrammen. Dette er meget bra, sammenliknet med hva en tidligere kunne forvente.

Prosjektene taktiske og strategiske vellykkethet er vurdert ex post selv om de fire prosjektene ikke har vært gjennom en kvalitetssikring av konseptvalget (KS1). To av prosjektene ble av evaluatorene vurderet som strategisk vellykkete (Svinesund og Momarken-Sekkelsten). En hypotese kan være at de prosjektene som senere skal evalueres og som har vært gjennom KS1 i større grad vil lykkes fordi konseptvalget har vært bedre utredet. Dette kan vi selvsagt ikke få dokumentert i dag. Det vi kan si med sikkerhet er at eksistensen av en KS1-rapport ville gjort ex post evalueringen langt enklere. Det ville også gitt en helt annet mulighet for å skille mellom ulike forklaringer på vellykkethet: ”tilfeldigheter” versus ”god styring og utslag av kjent usikkerhet”.

### ***Konklusjon***

KS-ordningens betydning for i hvilken grad investeringstiltakene lykkes på sikt vil inntil videre måtte bygge på andre studier i tillegg til fremtidige evalueringer av KS-prosjekter. Særlig viktig vil det være å komme i gang med evaluering av prosjekter som har vært gjennom KS1, noe som etter prognosen ennå vil måtte vente en tid til de første er kommet i driftsfasen som vist i figur 1.3.

## 5. Råd og anbefalinger om evaluering av investeringsprosjekter

Dette kapitlet presenterer råd og anbefalinger vedrørende ex post evaluering av investeringsprosjekter, basert på erfaringene fra de fire pilotevalueringene. Vi mener generelt at det i langt større grad enn i dag bør gjennomføres ex post evalueringer av store investeringsprosjekter. Den primære hensikten må være læring – både internt i statlige virksomheter og på tvers i staten, slik at fremtidige prosjekter i stor grad skal bli vellykkete. Ikke minst bør en styrke fokuset på strategisk vellykkethet.

Hensikten med denne studien var først og fremst å teste ut et rammeverk med målorientert evaluering og samfunnsøkonomisk analyse, for å vurdere investeringsprosjekters vellykkethet. Erfaringene med dette rammeverket er relativt gode, og særlig de strategiske evalueringskriteriene var bevisstgjørende for evaluatorene. Fire prosjekter er selvsagt ikke nok til å kunne trekke bastante konklusjoner, men vi mener likevel å ha fått et godt bilde av rammeverkets egnethet og om utfordringer en bør være oppmerksom på. Dette er sammenfattet nedenfor.

### Standardisering av evalueringsmodell og overordnede kriterier...

1. Det er viktig å finne en god balanse mellom et standardisert opplegg og evaluators frihet til å velge evalueringsspørsmål og indikatorer i det enkelte tilfelle. Den evalueringsmodellen som er valgt er fleksibel og generisk, med noen få overordnede kriterier som hver kan vurderes med det antall indikatorer og de datakilder som måtte passe. Den har i hovedsak fungert godt på fire prosjekter i helt forskjellige sektorer. Det kan imidlertid være krevende å skille mellom måloppnåelse og virkninger da ikke alle mål er eksplisitt formulert i forkant – fortrinnsvis bør det prinsipielle skillet mellom tilsiktede og utilsiktede virkninger synliggjøres, men det viktigste er at en inkluderer alle virkninger i evalueringen.
2. En av de viktigste fordelene med OECD-modellen er at den tvinger evaluators til å ivareta den strategiske dimensjonen. Strategisk vellykkethet er krevende å analysere fordi det uunngåelig krever kvalitative og til dels skjønnsmessige vurderinger. Det vanlige er derfor at man forsøker å unngå slik analyse ved å snevre inn evalueringen til det som lett kan måles. Følgelig er det de

---

operasjonelle aspektene som får oppmerksomhet. Da blir evalueringen mye mindre interessant og nyttig, derfor bør dette unngås.

3. Samfunnsøkonomisk lønnsomhet bør inngå som et selvstendig evalueringskriterium. Det er allerede krav om slik analyse ex ante og da gir det også god mening å følge opp analysen i ettertid. Grunnlaget for å gjennomføre en samfunnsøkonomisk analyse er ofte mangelfullt men vil i fremtiden, etter hvert som flere prosjekter har gjennomgått KS1, forhåpentligvis bli bedre. En bør likevel ikke ha for høye ambisjoner om dette, da det har vist seg enda vanskeligere å dokumentere alle effekter ex post enn å simulere dem ex ante.
4. Vi tror det kan være nyttig å standardisere ex post evalueringer. I særdeleshet gjelder det for KS-prosjekter. Da vil det i etterkant være mulig å trekke ut samme type informasjon for flere prosjekter fra de respektive evalueringene over tid (metaevalueringer), noe som gir enda bedre muligheter for læring på tvers av prosjekter og virksomheter. Følgende felles elementer bør være til stede i alle overordnede evalueringer av store prosjekter:
  - Evalueringen starter med en kritisk gjennomgang og eventuelt justering av prosjektets logikk og målstruktur. Denne er nesten alltid mangelfull og må justeres.
  - Seks kriterier benyttes for å vurdere ulike aspekter ved prosjektets vellykkethet: *produktivitet, måloppnåelse, virkninger, relevans, levedyktighet* samt *samfunnsøkonomisk lønnsomhet*. Det strategiske aspektet (og da særlig relevans og levedyktighet) er særlig viktig men også særlig krevende, og bør få mye fokus i evalueringen.
  - De seks evalueringskriteriene brytes ned til mer konkrete evalueringsspørsmål og indikatorer som er særlig relevante for evalueringsobjektet.

### **... men ikke standardisering av konkrete indikatorer**

5. Valget av konkrete evalueringsspørsmål, indikatorer og datainnsamlings- og analysemetoder må tilpasses den enkelte evaluering. Den fleksibiliteten evaluator har innenfor det overordnede rammeverket for evaluering er nyttig og nødvendig. Vi er skeptiske til å foreslå ytterligere standardisering, for eksempel i form av faste evalueringsspørsmål og indikatorer, eller faste måleskalaer for å kvantifisere de kvalitative vurderingene, selv om fremtidige evalueringer kanskje kunne ønske dette. Evalueringer – både ex ante og ex post – bør gjennomføres av personer med erfaring, kreativitet og evne til å tenke

selv, og et opplegg basert på standardiserte skjema eller regneark som skal fylles inn er åpenbart ingen god ide.

6. Selv om vi ikke anbefaler standardisering av indikatorer, ser vi ikke bort fra at det kan være hensiktsmessig å utvikle veiledning og sjekklister for hvilke indikatorer som *ofte* vil være relevant under det enkelte kriterium for prosjekter med stor grad av fellestrekk. Dette kan være en utfordring til den enkelte sektor.
7. Generelt er det å si at evaluering ikke er forskning, og mangelen på data vil trolig *alltid* være et problem gitt de kostnadsrammene som er satt. En må derfor fokusere på hva som er ”godt nok”, men likevel sette høye kvalitetskrav og redegjøre for metodevalg og begrensninger. Det innebærer å ha et særlig fokus på validitet, slik at de indikatorene som velges er klart relevante for å besvare evalueringsspørsmålene. Triangulering er svært viktig, og en bør normalt bruke både kvantitativ og kvalitativ informasjon og analyse, flere informasjonskilder, respondenter, etc.

### **Relevante nøkkeltall bør uansett være lett tilgjengelig – også om nullalternativet**

8. En må sikre at evaluator får lett tilgang til relevante data som allikevel samles inn i sektoren. Videre bør det finnes gode rutiner for å dokumentere prognosemodeller og øvrig underlag for analyser som gjøres *ex ante*, dvs. at forutsetninger som der er lagt lagres for ettertiden. Jernbanesektoren trekker frem som særlig utfordrende. Her hadde evaluator problemer med å få tilgang til informasjon om *ex ante*-analysen og sentrale data om trafikktviklingen på den aktuelle strekningen. Det gjør at effektiviteten i evalueringsarbeidet, og gjerne også kvaliteten på sluttresultatet, reduseres betraktelig.
9. Vi mener at det først og fremst er departementene som har ansvar for å forsikre seg om at data som burde finnes, gjøres tilgjengelig for senere evaluatorene. Underliggende etater, operatører med statsstøtte, Statistisk sentralbyrå, eventuelt andre som kan fremskaffe data, må gis det nødvendige oppdrag slik at dette kommer på plass. Hvor mye ressurser som skal brukes på datainnsamling bør vurderes i den enkelte sektor, men et minimum av informasjon om helt sentrale effekter bør være tilgjengelig. I tillegg må en selvsagt gjøre nødvendige tiltak for å sikre at konkurransesensitiv informasjon ikke offentliggjøres i evalueringsrapporten i detaljert form.<sup>16</sup>

---

<sup>16</sup> Det kan nevnes at Storbritannia har en annen organisering, her har tilsynsmyndigheten (Office of Rail Regulation) detaljert innsyn i både kostnader og inntekter av den trafikk

- 
10. Når det gjelder forsvarsanskaffelser er det som nevnt et tilleggsproblem at viktig informasjon er konfidensiell. Utfordringen her er først og fremst at en må skaffe de nødvendige klareringer i forkant for å få tilgang til informasjon, noe som er mulig men kan ta tid. I tillegg må en ta høyde for en klareringsprosess av hva som kan omtales i rapporten i etterkant når utkast til sluttrapport foreligger. Det er altså mulig at man kan få tilgang til materiale som bidrar til analysen uten at dette kan offentliggjøres i detalj.

## Tidspunkt for evaluering og avgrensning

11. Erfaringen generelt er at evaluering 3-5 år ut i driftsfasen (eller senere) kan være passende for ex post evaluering av KS-prosjekter, jevnfør anbefaling i Andersen m.fl. (2007). Våre pilotevalueringer støtter opp om denne anbefalingen. Hva som er et optimalt tidspunkt må uansett vurderes fra prosjekt til prosjekt, avhengig av hvilke typer av effekter det skal generere og hva annet disse effektene avhenger av. Dette er viktigere enn å velge et felles tidspunkt for alle prosjektene.
12. De aller fleste investeringsprosjekter vil, i større eller mindre grad, være del av en større sammenheng. Systemgrenser bør alltid drøftes, og det kan være argumenter for ulike valg. Ideelt sett burde man ha evaluert alle tiltak med samme eller liknende formål som én pakke. Dette er imidlertid sjelden realistisk eller håndterbart, og det vil gjøre evalueringen svært kompleks.
13. Erfaringen er at avgrensningen i det store og hele må være det enkelte prosjektet (ofte slik det er definert i KS2), først og fremst begrunnet med å redusere arbeidsomfanget i evalueringen. Men prosjektet må studeres i sin kontekst, noe som medfører at man også må se det i sammenheng med relaterte tiltak og eventuelle følgeprosjekter. Her finnes det ikke noe enkelt fasitsvar på hvor langt en bør gå. Det sentrale er hvilke effekter en ønsker å studere, og dette må konkretiseres i hvert enkelt tilfelle. Det essensielle er at denne avklaringen skjer så tidlig som mulig i evalueringsprosessen.
14. Dersom en har flere prosjekter som er del av et større konsept, for eksempel vei- eller jernbaneparseller som følger etter hverandre i tid, kan en løsning være å evaluere prosjekt for prosjekt, men med en langsiktig plan om å se evalueringene i sammenheng når alle prosjektene er implementert.

---

som bedrives på sporet – og dermed også om antall reisende. Tilsynsmyndigheten offentliggjør ikke informasjonen, men har selv mulighet til å gjøre granskinger.

## Hvem skal evaluere – viktige krav til evaluator

15. Evalueringsprosessen bør være upartisk og uavhengig av partene som har medvirket til kvalitetssikringen, planleggingen og gjennomføringen av tiltaket. Dette er viktig for å unngå interessekonflikter og sikre en viss grad av objektivitet. (Så lenge formålet kun er avgrenset til læring, er det likevel ikke helt utenkelig at for eksempel en etat kan evaluere seg selv, men normalt vil også kontrollhensyn være relevante).
16. Evalueringsteamet må ha god evalueringskompetanse. Evaluering er hovedsakelig en samfunnsfaglig analyse, hvor samfunnsøkonomi er en av flere relevante fagområder. Generelt er det en fordel at evalueringsteamet har erfaring med både kvalitative og kvantitative datainnsamlings- og analysemetoder. En erfaring er at teamet bør være tverrfaglig og at en bør unngå at et team bare består av én faggruppe, for eksempel bare samfunnsvitere, økonomer eller ingeniører.
17. I tillegg er det alltid en fordel om evaluator har noe fag- og sektorkompetanse. Det er derimot ikke nødvendigvis en fordel å ha stor detaljkunnskap om selve prosjektet og det tekniske innholdet, da det kan hindre en i å heve blikket og stille de riktige spørsmålene om taktisk og strategisk vellykkethet.
18. Vi tror det er behov for en solid første gangs innføring og veiledning i metodikken for de som skal gjennomføre evalueringene. Kan hende bør det utarbeides både mer spesifikk veiledning (håndbok) og et enkelt kursopplegg.

## Standardisering av prosess og omfang

19. Hva så med standardisering av ambisjonsnivå, omfang og ressursbruk for den enkelte evaluering? Omfanget på våre pilotevalueringer lå på ca. 3-4 månedsverk per evaluering og tok ca. 6-9 mnd. Dette var piloter hvor også metodeutviklingen og innsamling av erfaringer var en viktig del av studien. Vi tror at evalueringer ofte kan gjennomføres enklere og raskere enn dette.
20. Det er generelt viktig at ambisjonsnivået ikke settes for høyt. Med et for stort budsjett blir det lett til at evaluator leter etter problemstillinger av mer marginal betydning. Videre kan det være en fordel om gjennomføringstiden er kort, 2-3 måneder er ofte tilstrekkelig om evalueringen prioriteres og teamet jobber intensivt og informasjon og informanter gjøres lett tilgjengelige.
21. Spesielt bør det tidlig lages et detaljert opplegg med indikatorer og plan for datainnsamling og analyse, og en bør raskt begynne å skrive ut tentative



---

konklusjoner, som selvsagt justeres etter hvert som en samler data og konklusjonene modnes. En sluttrapport bør sjelden være på mer enn 50 sider (vedlegg kan komme i tillegg).

22. Ut over dette må en selvsagt avstemme ambisjonsnivå, omfang og prosess fra gang til gang, avhengig av egenskaper ved prosjektet og dets samfunnsmessige viktighet, evaluators erfaring og kompetanse med evaluering av liknende tiltak, om det gjelder sektorspesifikke tiltak som veiparseller, etc.

## **Innpassing i eksisterende monitorerings- og evalueringsevirsomhet**

23. Enkelte statlige virksomheter har allerede etablert opplegg for systematisk evaluering av utvalgte prosjekter på sine områder. Spesielt kan nevnes Statens Vegvesen som har gjennomført etterprøving av den samfunnsøkonomiske analysen for et utvalg prosjekter i flere år. Dette er bra, og det er ikke vår hensikt å foreslå et konkurrerende opplegg. Samtidig er det en generell erfaring at slike evalueringer ofte er smale og avgrenser seg til en oppfølging av noen få, kvantitative indikatorer. I slike tilfeller bør etatenes eksisterende evalueringsopplegg kunne suppleres slik at også det strategiske perspektivet ivaretas.
24. Departementene og etatene er viktige aktører når det gjelder å sørge for at det fremskaffes statistikk og andre nødvendige data til bruk i evalueringene. Det gjelder både kostnadssiden og også nyttesiden i tilfeller hvor det er relevant å kvantifisere indikatorer på nytte. Den enkelte sektor bør vurdere behov for mer systematisk innsamling og tilgjengeliggjøring av statistikk.

## **Hvordan sikre at evalueringen kommer til nytte**

25. Til syvende og sist er verdien av evalueringen avhengig av i hvilken grad resultatene tas i bruk – direkte eller indirekte – og bidrar til forbedringer av fremtidige prosjekter. Noen enkle tiltak for å sikre dette kan være: Brukere, eiere og eventuelt andre sentrale interessenter involveres innledningsvis for å sikre at det er konsensus om hensikten og suksesskriteriene for prosjektet samt viktige evalueringsspørsmål.
26. Videre er det viktig med en viss involvering av både direkte og indirekte berørte parter, og resultatene må presenteres på en forståelig måte. Evalueringens troverdig og kvalitet er også kritisk viktig. Dette er avgjørende for om aktørene vil ha tillit til funnene, og er ofte i seg selv et godt argument for å bruke eksterne evaluatorene.

Det må presiseres at vi i denne rapporten har fokusert på innholdet i evaluering av enkeltprosjekter og ikke på viktige forhold når evalueringer skal komme i bruk i den enkelte virksomhet. Herunder hvordan sikre god forankring, hvordan integrere resultatene fra evalueringene av investeringsprosjekter med annen styringsinformasjon, og hvilke styringsmessige prosesser evalueringene bør inngå i slik at en sikrer bedre læring og styring på et institusjonelt nivå, både for virksomheten og på tvers av statlige virksomheter. Bruk av evaluering i styringen av virksomheter omtales i DFØ (2011).

---

## Litteratur

Andersen, B., S. Bråthen, T. Fagerhaug, O. Nafstad, P. Næss og N. Olsson, 2007, *Effektivrering av store statlige investeringsprosjekter*, Concept rapport nr. 19, NTNU, 2007

Argyris, C. og D.A. Schön (1978): *Organizational Learning*. San Fransisco: Jossey-Bass.

Berg, P. m.fl. (1999): *Styring av statlige investeringer. Sluttrapport fra styringsgruppen* (prosjektet for styring av statlige investeringer, med deltakelse fra Finansdepartementet, Samferdselsdepartementet, Forsvarsdepartementet og Arbeids- og administrasjonsdepartementet)

Direktoratet for økonomistyring (2007): *Veileder. Evaluering av statlige tilskuddsordninger*

Direktoratet for økonomistyring (2008): *Sjekkliste for planlegging, gjennomføring og oppfølging av evalueringer*

Direktoratet for økonomistyring (2010): *Håndbok for samfunnsøkonomiske analyser*

Direktoratet for økonomistyring (2011): *Strategisk og systematisk bruk av evalueringer i styringen*

Finansdepartementet (2005a): *Veileder i samfunnsøkonomiske analyser*

Finansdepartementet (2005b): *Veileder til gjennomføring av evalueringer*, Statens forvaltningstjeneste

Finansdepartementet (2010): *Konkurransgrunnlag med innarbeidet kravspesifikasjon og kontraktspesifikasjoner til rammeavtale om konsulent tjenester vedrørende kvalitets sikring av konseptvalg, samt styringsunderlag og kostnadsoverslag for valgt prosjektalternativ*

Finne m.fl. (2012): *Klarering, kontroll og kø på grensen. Evaluering av felles statlig kontrollområde for toll, vegvesen og politi på Svinesund*, SINTEF, Trondheim

Fornyings-, administrasjons- og kirke departementet (2005): *Utredningsinstruksen. Instruks om utredning av konsekvenser, foreleggelse og høring ved arbeidet med offentlige utredninger, forskrifter, proposisjoner og meldinger til Stortinget.*

- Grønn, E. (2003): *Forelesninger i offentlig økonomi*, Cappelen Akademisk Forlag
- Heinzerling, L. og F. Ackerman (2002): *Pricing the priceless – Cost Benefit Analysis of Environmental Protection*. Georgetown Environmental Law and Policy Institute, Georgetown University Law Centre, Washington, USA
- Kehlert, R. (2012): *Impact evaluation using RCT. Reflections on the EES stance*, Evaluation Connections, nr. 2/2012
- Lædre, O. m.fl. (2012): *Levedyktighet og investeringstiltak*, Concept-rapport nr. 29
- NOU 1997: 27 *Nytte-kostnadsanalyser. Prinsipper for lønnsombetsvurderinger i offentlig sektor*
- NOU 2012: 16 *Samfunnsøkonomiske analyser*
- Nyborg og Spangen (1996): *Politiske beslutninger om investeringer i veier. Intervjuer med medlemmene i Stortingets samferdselskomite*. TØI-notat 1026
- Nyborg, K. (2012): *The Ethics and Politics of Environmental Cost-Benefit Analysis*, Routledge
- Nyström m.fl. (2012): *Ny järnväg mellan Asker och Sandvika. En utvärdering*, Rapport fra VII
- Næss, P. (2004a): *Bedre behovsanalyser. Erfaringer og anbefalinger om behovsanalyser i store offentlige investeringsprosjekt*, Concept-rapport nr. 5, NTNU
- Næss, P. (2004b): *Bedre utforming av store offentlige investeringsprosjekter* *Vurdering av behov, mål og effekt i tidligfasen*, Concept-rapport nr. 9, NTNU
- OECD (1999): *Improving Evaluation Practices. Best Practice Guidelines for Evaluation and Background Paper*
- OECD (2002): *Glossary of key terms in evaluation and results based management*. Paris: OECD.
- Olsson, N. (2005): *Hvordan tror vi at det blir? Effektvurderinger av store offentlige prosjekter*, Concept-rapport nr. 7, Concept/NTNU
- Olsson N. og Veiseth M. (2010), *Jernbanetraffikk*, Tapir Akademisk forlag, Trondheim

- 
- Rosenberg, L.J. m.fl. (1970): *Project evaluation and the project appraisal reporting system*. Fry Consultants Inc.
- Samset, K. (2003): *Project Evaluation – Making Investments Succeed*, Tapir Akademisk Forlag
- Samset, K. 2008, *Prosjekt i tidligfasen. Valg av konsept*, Tapir Akademisk Forlag, Trondheim 2008
- Samset, K. m.fl. (2009): *Sykebus, fregatter og skipstunnel: Logisk minimalisme, rasjonalitet – og de avgjørende valg*, Concept-rapport nr. 21, NTNU
- Samset, K. og T. Christensen (2012): *Evaluering av prosjekter ex ante og ex post – og beslutningsprosessenes kompleksitet og betydning*, kapittel i *Evaluering. Tradisjoner, praksis, mangfold*, Fagbokforlaget, Oslo, 2013
- Samset, K. (2012): *Lenge før og langt senere. Et tilbakeblikk på konseptvalget i historiske prosjekter*, Ex Ante akademisk forlag
- Statens Vegvesen (2006): *Veileder for etterprøving av store vegprosjekter*
- Sunnevåg, K.J., red. (2007): *Beslutninger på svakt informasjonsgrunnlag. Tilnærminger og utfordringer i prosjekters tidlige fase*, Concept rapport nr. 17, NTNU
- Trafikverket (2012): *Riktlinjer for framtagande av trafikprognoser*, rapport 2012:045 fra Trafikverket
- Whist m.fl. (2012): *Evaluering av 6300 Skjold-Klassen MTBer*, Rapport fra Scanteam
- Aass og Welde (2012): *Evaluering av E18 Momarken - Sekkelsten*, arbeidsrapport fra Concept-programmet

## Concept rapportserie

Papirtrykk: ISSN 0803-9763

Elektronisk utgave på internett: ISSN 0804-5585

Lastes ned fra: [www.concept.ntnu.no/publikasjoner/rapportserie](http://www.concept.ntnu.no/publikasjoner/rapportserie)

Rapport	Tittel	Forfatter
Nr. 1	Styring av prosjektporteføljer i staten. Usikkerhetsavsetning på porteføljenivå <i>Project Portfolio Management. Estimating Provisions for Uncertainty at Portfolio Level.</i>	Stein Berntsen og Thorleif Sunde
Nr. 2	Statlig styring av prosjektledelse. Empiri og økonomiske prinsipper. <i>Economic Incentives in Public Project Management</i>	Dag Morten Dalen, Ola Lædre og Christian Riis
Nr. 3	Beslutningsunderlag og beslutninger i store statlige investeringsprosjekt <i>Decisions and the Basis for Decisions in Major Public Investment Projects</i>	Stein V. Larsen, Eilif Holte og Sverre Haanæs
Nr. 4	Konseptutvikling og evaluering i store statlige investeringsprosjekt <i>Concept Development and Evaluation in Major Public Investment Projects</i>	Hege Gry Solheim, Erik Dammen, Håvard O. Skaldebø, Eystein Myking, Elisabeth K. Svendsen og Paul Torgersen
Nr. 5	Bedre behovsanalyser. Erfaringer og anbefalinger om behovsanalyser i store offentlige investeringsprosjekt <i>Needs Analysis in Major Public Investment Projects. Lessons and Recommendations</i>	Petter Næss
Nr. 6	Målformulering i store statlige investeringsprosjekt <i>Alignment of Objectives in Major Public Investment Projects</i>	Ole Jonny Klakegg
Nr. 7	Hvordan tror vi at det blir? Effektvurderinger av store offentlige prosjekt <i>Up-front Conjecture of Anticipated Effects of Major Public Investment Projects</i>	Nils Olsson
Nr. 8	Realopsjoner og fleksibilitet i store offentlige investeringsprosjekt <i>Real Options and Flexibility in Major Public Investment Projects</i>	Kjell Arne Brekke
Nr. 9	Bedre utforming av store offentlige investeringsprosjekter. Vurdering av behov, mål og effekt i tidligfasen <i>Improved Design of Public Investment Projects. Up-front Appraisal of Needs, Objectives and Effects</i>	Petter Næss med bidrag fra Kjell Arne Brekke, Nils Olsson og Ole Jonny Klakegg
Nr. 10	Usikkerhetsanalyse – Kontekst og grunnlag <i>Uncertainty Analysis – Context and Foundations</i>	Kjell Austeng, Olav Torp, Jon Terje Midtbø, Ingemund Jordanger, og Ole Morten Magnussen
Nr. 11	Usikkerhetsanalyse – Modellering, estimering og beregning <i>Uncertainty Analysis – Modeling, Estimation and Calculation</i>	Frode Drevland, Kjell Austeng og Olav Torp
Nr. 12	Metoder for usikkerhetsanalyse <i>Uncertainty Analysis – Methodology</i>	Kjell Austeng, Jon Terje Midtbø, Vidar Helland, Olav Torp og Ingemund Jordanger

## Concept rapportserie

Papirtrykk: ISSN 0803-9763

Elektronisk utgave på internett: ISSN 0804-5585

Lastes ned fra: [www.concept.ntnu.no/publikasjoner/rapportserie](http://www.concept.ntnu.no/publikasjoner/rapportserie)

Rapport	Tittel	Forfatter
Nr. 13	Usikkerhetsanalyse – Feilkilder i metode og beregning <i>Uncertainty Analysis – Methodological Errors in Data and Analysis</i>	Kjell Austeng, Vibeke Binz og Frode Drevland
Nr. 14	Positiv usikkerhet og økt verdiskaping <i>Positive Uncertainty and Increasing Return on Investments</i>	Ingemund Jordanger
Nr. 15	Kostnadsusikkerhet i store statlige investeringsprosjekter; Empiriske studier basert på KS2 <i>Cost Uncertainty in Large Public Investment Projects. Empirical Studies</i>	Olav Torp (red.), Ole Morten Magnussen, Nils Olsson og Ole Jonny Klakegg
Nr. 16	Kontrahering i prosjektets tidligfase. Forsvarets anskaffelser. <i>Procurement in a Project's Early Phases. Defense Aquisitions</i>	Erik N. Warberg
Nr. 17	Beslutninger på svakt informasjonsgrunnlag. Tilnærminger og utfordringer i prosjekters tidlige fase <i>Decisions Based on Scant Information. Challenges and Tools During the Front-end Phases of Projects</i>	Kjell Sunnevåg (red.)
Nr. 18	Flermålsanalyser i store statlige investeringsprosjekt <i>Multi-Criteria Decision Analysis In Major Public Investment Projects</i>	Ingemund Jordanger, Stein Malerud, Harald Minken, Arvid Strand
Nr. 19	Effektvurdering av store statlige investeringsprosjekter <i>Impact Assessment of Major Public Investment Projects</i>	Bjørn Andersen, Svein Bråthen, Tom Fagerhaug, Ola Nafstad, Petter Næss og Nils Olsson
Nr. 20	Investorers vurdering av prosjekters godhet <i>Investors' Appraisal of Project Feasibility</i>	Nils Olsson, Stein Frydenberg, Erik W. Jakobsen, Svein Arne Jessen, Roger Sørheim og Lillian Waagø
Nr. 21	Logisk minimalisme, rasjonalitet - og de avgjørende valg <i>Major Projects: Logical Minimalism, Rationality and Grand Choices</i>	Knut Samset, Arvid Strand og Vincent F. Hendricks
Nr. 22	Miljøøkonomi og samfunnsøkonomisk lønnsomhet <i>Environmental Economics and Economic Viability</i>	Kåre P. Hagen
Nr. 23	The Norwegian Front-End Governance Regime of Major Public Projects – A Theoretically Based Analysis and Evaluation	Tom Christensen
Nr. 24	Markedsorienterte styringsmetoder i miljøpolitikken <i>Market oriented approaches to environmental policy</i>	Kåre P. Hagen
Nr. 25	Regime for planlegging og beslutning i sykehusprosjekter <i>Planning and Decision Making in Hospital Projects. Lessons with the Norwegian Governance Scheme.</i>	Asmund Myrbostad, Tarald Rohde, Pål Martinussen og Marte Lauvsnes

## Concept rapportserie

Papirtrykk: ISSN 0803-9763

Elektronisk utgave på internett: ISSN 0804-5585

Lastes ned fra: [www.concept.ntnu.no/publikasjoner/rapportserie](http://www.concept.ntnu.no/publikasjoner/rapportserie)

Rapport	Tittel	Forfatter
Nr. 26	Politisk styring, lokal rasjonalitet og komplekse koalisjoner. Tidligfaseprosessen i store offentlige investeringsprosjekter <i>Political Control, Local Rationality and Complex Coalitions. Focus on the front-end of large public investment projects</i>	Erik Whist, Tom Christensen
Nr. 27	Verdsetting av fremtiden. Tidshorisont og diskonteringsrenter <i>Valuing the future. Time horizon and discount rates</i>	Kåre P. Hagen
Nr. 28	Fjorden, byen og operaen. En evaluering av Bjørvikautbyggingen i et beslutningsteoretisk perspektiv <i>The Fjord, the City and the Opera. An Evaluation of Bjørvika Urban Development</i>	Erik Whist, Tom Christensen
Nr. 29	Levedyktighet og investeringstiltak. Erfaringer fra kvalitetssikring av statlige investeringsprosjekter <i>Sustainability and Public Investments. Lessons from Major Public Investment Projects</i>	Ola Lædre, Gro Holst Volden, Tore Haavaldsen
Nr. 30	Ettrevaluering av statlige investeringsprosjekter. Konklusjoner, erfaringer og råd basert på pilotevaluering av fire prosjekter <i>Evaluating public investment projects. Lessons and advice from a meta-evaluation of four projects</i>	Gro Holst Volden og Knut Samset



Forskningsprogrammet Concept skal utvikle kunnskap som sikrer bedre ressursutnyttning og effekt av store, statlige investeringer. Programmet driver følgeforskning knyttet til de største statlige investeringsprosjektene over en rekke år. En skal trekke erfaringer fra disse som kan bedre utformingen og kvalitetssikringen av nye investeringsprosjekter før de settes i gang.

Concept er lokalisert ved Norges teknisk- naturvitenskapelige universitet i Trondheim (NTNU), ved Fakultet for ingeniørvitenskap og teknologi. Programmet samarbeider med ledende norske og internasjonale fagmiljøer og universiteter, og er finansiert av Finansdepartementet.

*The Concept research program aims to develop know-how to help make more efficient use of resources and improve the effect of major public investments. The Program is designed to follow up on the largest public projects over a period of several years, and help improve design and quality assurance of future public projects before they are formally approved.*

*The program is based at The Norwegian University of Science and Technology (NTNU), Faculty of Engineering Science and Technology. It cooperates with key Norwegian and international professional institutions and universities, and is financed by the Norwegian Ministry of Finance.*

**Address:**

The Concept Research Program  
Høgskoleringen 7A  
N-7491 NTNU  
Trondheim  
NORWAY

Tel.: +47 73594670

Fax.: +47 73597021

ISSN: 0803-9763

ISBN: 978-82-93253-09-9

