

**SNF-rapport nr. 06/07**

**BESLUTNINGER PÅ SVAKT  
INFORMASJONSGRUNNLAG**

**TILNÆRMINGER OG  
UTFORDRINGER I TIDLIG FASE**

**AV**

**KJELL J. SUNNEVÅG (RED.)**

**SNF-prosjekt nr. 2682  
Tidligfasevurderinger på svakt informasjonsgrunnlag**

**Prosjektet er finansiert av Finansdepartementet**

**SAMFUNNS- OG NÆRINGSLIVSFORSKNING AS  
BERGEN, MAI 2007**

© Dette eksemplar er fremstilt etter avtale  
med KOPINOR, Stenergate 1, 0050 Oslo.  
Ytterligere eksemplarfremstilling uten avtale  
og i strid med åndsverkloven er straffbart  
og kan medføre erstatningsansvar.

ISBN 978-82-491-0510-6 Trykt versjon  
ISBN 978-82-491-0511-3 Elektronisk versjon  
ISSN 0803-4036

# FORORD

Erfaringer fra en rekke store prosjekter har vist at deler av forutsetningene som har ligget til grunn for vedtak og igangsetting ikke er blitt oppfylt. Dette har medført problemer som store overskridelser på kostnadssiden, forsinkelser og lav måloppnåelse.

Mye tyder på at valget av konsept er minst like avgjørende for at prosjekter skal lykkes som at en har god styring av prosjektene igjennomføringsfasen. På denne bakgrunnen er det et stort behov for å se nærmere på prosessene og premissgivingen som ligger til grunn for konseptvalg i store prosjekter, med sikte på å forbedre det faglige grunnlaget, prosedyrer og praksis som kan bidra til kvalitativt fornuftige valg.

I denne rapporten presenteres en samling artikler som tar for seg ulike faglige utfordringer og faglige tilnærminger til å vurdere alternative konsept i en tidlig fase. Artikkelsamlingen er finansiert gjennom forskningsprogrammet Concept. Dette er et program som skal utvikle kunnskap som sikrer bedre ressursutnyttning og effekt av store statlige investeringer. Concept-programmet utføres ved NTNU Institutt for bygg, anlegg og transport. Oppdragsgiver for Concept-programmet er Finansdepartementet.

Kjell J. Sunnevåg, forskningssjef ved Samfunns- og næringslivsforskning AS har vært redaktør for artikkelsamlingen.



# INNHOOLD

Innledning og sammendrag ..... 1

## **Del 1 Generelt om tidligfaseutfordringer**

*Knut Samset*

1 Generelt om prosjekter og utfordringer i tidligfasen.....9

*Petter Næss*

2 Fra behov til konsept – om prosjekter og perspektiver i tidligfasen .....21

*Bent Flyvbjerg*

3 How Optimism Bias and Strategic Misrepresentation in Early Project Development Undermine Implementation .....41

## **Del 2 Generering av informasjon i en tidlig fase**

*Knut Samset*

4 Hvilke muligheter har vi til å forutsi i en tidlig fase? .....55

*Kjell J. Sunnevåg*

5 Ekspertvurderinger for å avdekke og strukturere kvalitativ informasjon .....69

*Bent Flyvbjerg*

6 Eliminating Bias In Early Project Development through Reference Class Forecasting and Good Governance .....90

*Jostein Lillestøl*

7 Dataanalyse og prosjektrisiko ved sparsom informasjon..... 111

## **Del 3 Kvalitetssikring av informasjon**

*Ole Jonny Klakegg*

8 Kvalitetssikring av informasjon i praksis ..... 141

*Geir Kirkebøen*

9 Skjevheter i fagfolks skjønn: Hvordan kan beslutningstaking forbedres?..... 174

*Knut Samset*

10 Tiltak for å bedre kvaliteten på kvalitativ informasjon – semantikk og presisjon.....200

## Del 4 Anvendelse av informasjon

### ***Bjørn Andersen***

11 En verktøykasse for analyse i prosjekters tidligfase .....217

### ***Kjell Austeng***

12 Fuzzy logikk i tidligfasevurderinger .....239

### ***Kjell J. Sunnevåg***

13 Verdi og konsekvenser av ny informasjon.....270

### ***Kåre P. Hagen***

14 Fra analyse til beslutning - samfunnsøkonomisk lønnsomhetsvurdering .....289

## INNLEDNING OG SAMMENDRAG

Med ”tidligfasen” forstås gjerne tiden fra det identifiseres et gitt samfunnsmessig behov, til man tar beslutningen om hvilket konsept som best ivaretar det identifiserte behovet. Det kan f.eks. være et gitt transportbehov i samferdselssektoren, og de ulike løsningene kan være jernbane eller ny vei, og da satt opp mot et realistisk nullalternativ. I denne fasen vil det altså kunne foreligge flere ulike alternative løsninger som i større eller mindre grad ivaretar det gitte behovet. En sentral utfordring er knyttet til å skaffe et informasjonsgrunnlag som kan bidra til å avdekke de ulike alternativene, og som i tillegg tillater en meningsfull og konsistent sammenligning mellom disse.

Det kan være begrenset informasjon tilgjengelig om de ulike alternative løsningene. Selv om det foreligger informasjon i kvantitativ eller kvalitativ form om de alternative løsninger, kan informasjonsgrunnlaget være svakt og beheftet med stor grad av usikkerhet. Informasjon om andre lignende prosjekter kan også være begrenset eller ha begrenset overføringsverdi. Det vil altså kunne foreligge tildels ”råttan” og lite formalisert eller strukturert informasjon om ulike alternative løsningene, eller konseptene.

Siktemålet med prosessen er å identifisere den løsningen som ivaretar det samfunnsmessige behovet på en best mulig måte, ut fra behovstilfredsstillelse og ressursbruk. Den løsningen som identifiseres som den mest gunstige ut fra disse kriteriene vil senere bli underlagt en nærmere vurdering som til dels krever andre faglige tilnærminger enn det som vil være fokus her. Vi må altså hele tiden ha klart for oss vi befinner oss i en beslutningssituasjon.

Denne artikkelsamlingen handler om beslutningsgrunnlag og beslutninger i den tidligste fasen av et prosjekt. Det vil si på det tidspunktet der de største avgjørelsene skal tas, men samtidig da usikkerheten er størst, muligheten for påvirkning er størst, og informasjonsgrunnlaget er mest begrenset. Formålet med artikkelsamlingen er for det første å drøfte noen av de faglige utfordringer man står overfor. Dette omfatter både kunnskap om denne delen av beslutningsprosessen og de aktører som er involvert, så vel som å forstå de begrensninger og utfordringer som et svakt informasjonsgrunnlag innebærer. Det andre formålet er å bidra til et bedre beslutningsgrunnlag gjennom en systematisk og anvendt orientert beskrivelse av ulike teknikker som kan benyttes for å finne og systematisere struktur i de kvantitative og kvalitative data som er tilgjengelige i en tidlig fase. For det tredje vil vi se på hvordan dette informasjonsgrunnlaget kan håndteres videre i beslutningsprosessen på en slik måte at vi får et bedre utgangspunkt for å rangere alternative løsninger i forhold til hverandre på en faglig holdbar måte.

Artikkelforfatterne går altså inn på beslutningssituasjonen og de ulike faglige utfordringene og tilnærmingene som kan benyttes i en tidlig fase. Til en viss grad vil dette være beskrivelse av vel etablerte tilnærminger, mens andre vil være nyere, og kanskje fremstå som mer esoteriske. Et siktemål med artikkelsamlingen har vært å få frem muligheter og begrensninger ved både eksisterende og ny metodikk for å erverve og håndtere informasjon i en tidlig fase.

Artikkelsamlingen er delt inn i fire deler:

Del 1 Generelt om tidligfaseutfordringer

Del 2 Generering av informasjon i en tidlig fase

Del 3 Kvalitetssikring av informasjon

Del 4 Anvendelse av informasjon - vurderinger på svakt informasjonsgrunnlag

Den første delen innledes med tre kapitler skrevet av henholdsvis Knut Samset, Petter Næss og Bent Flyvbjerg. I **kapittel 1** skriver *Knut Samset* generelt om de utfordringer man står overfor i en tidlig fase. Han peker på at selv om man intuitivt vil tenke at mangel på informasjon må være det største problemet i en slik situasjon, er dette bare til en viss grad riktig. Det avgjørende er *hvilken* type informasjon en har behov for. På dette innledende tidspunktet i prosessen er det først og fremst selve *problemet* og de *behovene* prosjektet er ment å bidra til å løse som må belyses grundig.

I **kapittel 2** peker *Petter Næss* på at erfaringer fra en rekke store offentlige investeringsprosjekter har vist at deler av forutsetningene som har ligget til grunn for vedtak og igangsetting ikke er blitt oppfylt. Dette har medført problemer som store overskridelser på kostnadssiden, forsinkelser og lav måloppnåelse. For å unngå slike uønskede resultater, må planleggingen i den tidlige fasen av prosjektarbeidet styrkes. På bakgrunn av identifiserte svakheter og mangler ved planleggingen av en rekke tidligere store offentlige investeringsprosjekter, gir Næss konkrete anbefalinger om behovsanalyser, målformuleringer, utforming av løsningskonsepter og effektvurderinger i den tidlige fasen av planarbeidet, med særlig fokus på behovsanalyser og effektvurderinger. Kapitlet bygger i stor grad på forfatterens forskning på disse temaene innenfor Concept-programmet.

Også *Bent Flyvbjerg* baserer seg på omfattende forskningsbasert erfaring når han i **kapittel 3** oppsummerer viktig kunnskap fra store offentlige investeringsprosjekt internasjonalt. Flyvbjerg fokuserer på prosjekter i transportsektoren, og et viktig poeng er at et stort problem i planleggingen av slike prosjekter er den store grad av misvisende informasjon om kostnader og nytteeffekter som produseres i de tidlige stadier av prosjektutviklingen, og den risiko som denne misvisende informasjonen medfører. Flyvbjerg ser også på årsakene til at dette problemet oppstår; ikke minst skjevheter forårsaket av optimisme og strategisk feilrepresentasjon.

Den andre delen av artikkelsamlingen fokuserer på de utfordringer som er knyttet til å generere informasjon i en tidlig fase. Denne delen innledes med en artikkel av *Knut Samset*, som ser på hvilke muligheter vi har til å forutsi en tidlig fase. I **kapittel 4** peker han på at en står overfor formidable utfordringer innledningsvis som blir større desto lenger tidligfasen varer. Videre vil vurderingene i den tidligste fasen av et prosjekt i stor grad måtte baseres på *antakelser*. Imidlertid vil evnen og muligheten for å forutsi være avgjørende for beslutningenes godhet. Muligheten for å skaffe seg informasjon om tilsvarende og liknende prosjekter og erfaringene med disse vil bedre muligheten for å forutsi på et tidlig tidspunkt. Samset hevder at problemet i praksis ikke så mye ligger i at slik informasjon ikke er lett tilgjengelig. I de fleste tilfellene er det mulig å finne mange kilder til slik informasjon. Bøynen er heller ganske enkelt at en ikke har en tradisjon for å bruke ressurser på dette, slik at det ofte ikke blir gjort i tilstrekkelig grad.



I **kapittel 5** ser *Kjell J. Sunnevig* nærmere på i hvilken grad ekspertvurderinger kan benyttes til å avdekke den informasjon som er nødvendig for å kunne ta gode beslutninger. Når det er knapt med tilgjengelige eller sammenlignbare kvantitative (statistiske) erfaringsdata, kan den best tilgjengelige informasjon være erfaringsbaserte og subjektive oppfatninger hos eksperter innenfor det aktuelle området. Ekspertene kan også benyttes til å strukturere et problem, dvs. bestemme hvilke data og variable som er relevante for analysen, hvilke analytiske metoder som er hensiktsmessige, samt hva som er relevante forutsetninger. I artikkelen ser Sunnevig nærmere på ulike tilnæringer til å avdekke ekspertinformasjon, så vel som noen av de fallgrubene en må passe seg for.

I **kapittel 6** følger *Bent Flyvbjerg* opp noen av sine tanker fra kapittel 3, med forslag til flere tiltak som tar sikte på å realisere bedre planlegging og beslutninger. Ett viktig tiltak i så måte er bedre metodikk for å kunne gi anslag på prosjektets kostnader og nytteeffekter ("Reference class forecasting"). Et annet tiltak retter seg inn mot insentivstrukturen for de som planlegger prosjektet.

Mens kapittel 5 fokuserte på avdekking av ekspertbasert informasjon, tar *Jostein Lillestøl* i **kapittel 7** for seg teknikker for å behandle og analysere kvantitativ informasjon. Kapitlet inneholder tre hovedavsnitt som kan leses uavhengig av hverandre: i) Kvalitetssikring av datagrunnlaget for prosjekter; ii) dataanalyse av usikkerhet og risiko og iii) bayesiansk sannsynlighetsvurdering. Mens de to første delene vil være velkjent stoff for en som er vel bevandret i statistikkens verden, vil det være nyttig oppfriskings- og referensemateriale for de som ikke er like bevandret. Den siste delen derimot, tar for seg statistiske tilnæringer som er av nyere art. Lillestøl kommer inn på tema som Bayesiansk klassifisering, Bayes nett, og informasjon og entropi. Dette er tema som er tatt med nettopp fordi det representerer faglig interessante vinklinger i en tidlig fase, dvs. når informasjonsgrunnlaget er svakt.

I den tredje delen av artikkelsamlingen fokuserer vi på kvalitetssikring av informasjon. I **kapittel 8** ser *Ole Jonny Klakegg* nærmere på hvordan dette gjøres i praksis. Første del omhandler ulike rammeverk for kvalitetssikring av underlag for beslutninger, slik de er introdusert i et par utvalgte land og store private bedrifter. Andre del identifiserer fallgruver som er aktuelle trusler mot god kvalitetssikring i praksis. Videre ses det på erfaringer med kvalitetssikring av beslutningsgrunnlaget slik det faktisk er utført og bør utføres i store statlige investeringsprosjekt. Dette er i hovedsak basert seg på dokumentert praksis fra utførte kvalitetssikringer i perioden 2000-2006. Målet er å fastslå hvordan kvalitetssikring av informasjon faktisk er utført i store offentlige prosjekt i denne perioden, og spesielt innenfor ordningen med eksternt kvalitetssikring. Tredje del av dette kapitlet viser elementer av hvordan en beste praksis på dette området vil kunne se ut i dag. I denne delen er det vist gode prinsipper fra flere fagfelt så vel som andre sammenhenger for å supplere bildet. Til slutt presenterer Klakegg en løsning som vil forbedre evnen til å håndtere de fleste av fallgrubene presentert i andre del, samt tilføre en rekke andre fordeler.

I **kapittel 9** gir *Geir Kirkebøen* et innblikk i beslutningspsykologiens forståelse av mennesket som beslutningstaker. Han forklarer ulike typer skjevheter som systematisk reduserer beslutningers kvalitet. Videre gir han interessante eksempler på effektive strategier for å forbedre fagfolks beslutningstaking. Skjevhetene i folks beslutningstaking, hvordan vi avviker fra normative standarder og hvorfor, er godt kartlagt i faget beslutningspsykologi. Kirkebøen starter med å gi noen glimt fra beslutningspsykologiens (for)historie. Dette kapitlet kan sees i sammenheng med kapittel 5 om avdekking av ekspertinformasjon. Kirkebøens peker på at

feilene i fagfolks skjønnsmessige bedømminger og beslutninger ikke er tilfeldige. Fagfolks skjønn avviker *systematisk* fra normative, ”rasjonelle” standarder. Systematiske skjevheter gir pr definisjon redusert beslutningskvalitet og kan følgelig ha alvorlige konsekvenser. Men det at skjevhetene er systematiske gir også håp. Et viktig budskap er at så lenge vi kan identifisere skjevhetene og deres årsaker, bør det også være mulig å motvirke dem.

*Knut Samset* ser i **kapittel 10** nærmere på betydningen av semantikk og presisjon i forhold til å bedre kvaliteten på kvalitativ informasjon. Mye av informasjonen vi bruker er basert på underliggende systematisk analyse av fakta, ofte kvantitativ informasjon. Andre deler av informasjonen bygger på antakelse og skjønn. Mye kvalitativ informasjon fremstår i form av komplekse, sammensatte utsagn. Analyse av slik informasjon forutsetter at den relevante teksten brytes ned i sine enkelte meningsbærende elementer. Mye av analysen består i *kategorisering*, i å identifisere eller bygge *struktur*, eller å vurdere *egenskap*. Samset peker i denne sammenheng på betydningen av å vurdere kausalitet og sannsynlighet som viktige hjelpemidler.

De foregående delene av artikkelsamlingen har tatt for seg utfordringer i prosjekters tidligfase mer generelt, og hvordan man kan fremskaffe informasjon av god kvalitet i denne fasen. Denne informasjonen blir først nyttig om den anvendes som datagrunnlag for ulike analyser, som i sin tur utgjør grunnlaget for å fatte beslutninger. Hensikten med **kapittel 11** er å presentere ulike analyseteknikker som kan anvendes til dette formålet. *Bjørn Andersen* peker på at det finnes et stort antall ulike analyseteknikker tilgjengelig. Disse teknikkene har typisk vært utviklet for ulike formål, deretter tatt i bruk i andre sammenhenger, gjerne i noe tilpasset form. Det finnes ingen verktøy utviklet spesifikt for prosjekters tidligfase. Det må dermed gjøres et utvalg av egnede analyseverktøy fra andre fagområder, og Andersen presenterer ulike teknikker tilpasset i) behov knyttet til å forutse effekter (positive og negative) av ulike beslutningsalternativer; ii) evaluere/sammenlikne ulike beslutningsalternativer; iii) vurdere klimaet i prosjektet og dets omgivelser og krefter for og imot gjennomføring av ulike beslutningsalternativer; så vel som iv) å analysere alternativets effekt på de ønskede mål/effekter prosjektet skal oppnå.

I **kapittel 12** går *Kjell Austeng* nærmere inn på en spesifikk teknikk, nemlig *fuzzy logic*. Dette er en fellesbetegnelse på flere beslutnings- eller styringsmodeller basert på teorien om diffuse mengde (fuzzy sets). Felles for alle disse modellene er at de er laget for å håndtere situasjoner i virkelighetens verden, og at de bevisst søker å utnytte den menneskelige vurderingsevne. Fuzzy logic er å betrakte som et ekspertsystem, men til forskjell fra tradisjonell tenkning hvor et utsagn er vurdert som enten riktig eller galt, så er fuzzy logic i stand til å ta vare på utsagn som "kanskje riktig" eller "delvis galt". Felles for modeller basert på fuzzy logic er at de er tilpasset menneskelig måte å tenke på. De gir rom for omtrentlighet, usikkerhet i anslagene basert på subjektive vurderinger, og de kan brukes i etablering av beslutningsstøtte. Dette gjør metodikken særlig interessant i en tidlig fase.

**Kapittel 13** er også orientert mot beslutningsmetodikk. Fokus i dette kapittelet, skrevet av *Kjell J. Sunnevig*, er på hvordan man metodisk håndterer *ny* informasjon som kommer etter hvert som de ulike beslutningspunktene nærmer seg, og informasjon som først vil avsløres *etter* at beslutning om investering er tatt. Sentralt i den første del av drøftingen står bayesiansk metodikk. Den andre del av kapittelet er en gjennomgang av hvordan informasjonsavsløring som skjer etter at prosjektet er igangsatt kan påvirke både innfasing og valg av prosjektkonsept. Perspektivet her er en realopsjonstilnærming til prosjektvurdering.

I siste kapittel, **kapittel 14**, trekker Kåre P. Hagen trådene sammen. Han skriver om samfunnsøkonomisk lønnsomhetsanalyse av prosjekter. I den siste delen av kapitlet ser han spesielt på noen av de utfordringene som oppstår ved å bruke slike analyser i prosjektets tidlige fase. I en tidlig fase i planleggingen og lønnsomhetsevalueringen av et prosjekt vil en typisk stå overfor et uttrykt behov eller etterspørsel etter en nærmere bestemt type tjenester, f.eks. et transportbehov, og problemstillingen er å finne frem til et prosjektkonsept som tilfredsstillende dette behovet på en mest mulig effektiv måte. I noen tilfelle kan dette ta form av en kostnadseffektivitetsanalyse som går på å finne frem til en prosjektutforming som realiserer et bestemt mål til lavest mulig kostnad. En slik effektivitetsanalyse vil imidlertid kompliseres av at den fremtidige nytte- og kostnadssiden til prosjektet kan variere på en ikke fullt ut predikerbar måte sett fra analysetidspunktet. Om forventet kostnad legges til grunn, kan det skjule det forholdet at prosjektkonseptet kan føre til svært høye kostnader i noen scenarier. Hagen ser spesielt på diskonteringsrenten og betydningen av ulike former for risiko i denne sammenhengen.



# Del 1

## Generelt om tidligfaseutfordringer



# 1 GENERELT OM PROSJEKTER OG UTFORDRINGER I TIDLIGFASEN

KNUT SAMSET\*

## 1.1 Problemer og utfordringer i store investeringsprosjekter

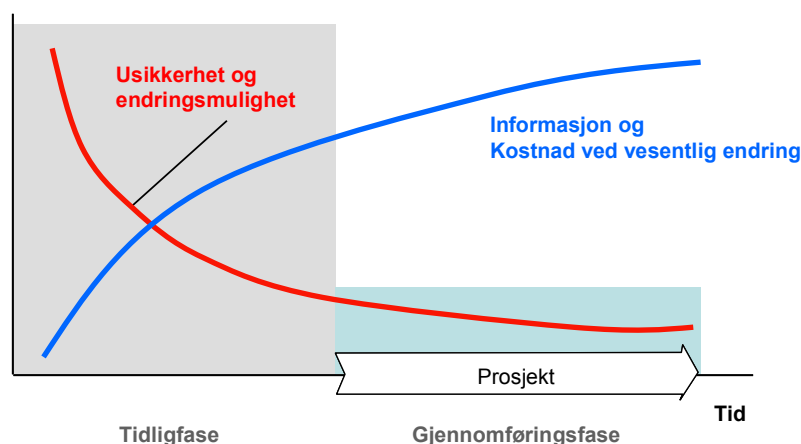
Et prosjekt kjennetegnes ved at det er et avgrenset tiltak med et definert mål som gjennomføres innenfor en avtalt tidsramme, budsjettramme, med en midlertidig organisasjon. Det skal resultere i en leveranse i den hensikt å løse et problem eller tilfredsstillende et behov. Prosjekter anses som unike i den forstand at de alltid i større eller mindre grad vil være forskjellige, selv om målet og leveransen er den samme, på grunn av at omgivelsene vil være forskjellige og den usikkerheten de opererer under vil variere.

Prosjekter er mer eller mindre vellykket. Problemene som oppstår i et prosjekt kan være mer eller mindre alvorlige og konsekvensene kan være mer eller mindre langvarige. To av de vanligste problemene, som samtidig ofte er av de minst alvorlige, er at prosjektet forsinkes eller blir dyrere enn antatt. Paradoksalt nok er dette allikevel de problemene som blir mest omtalt i media. Kostnadsoverskridelsen kan være stor i forhold til budsjettet, fra noen få prosent til flere hundre prosent. Men oftest er det slik at overskridelsen er liten i forhold til fremtidige driftsinntekter og derfor påvirker lønnsomheten i begrenset grad på lang sikt. Det nye Rikshospitalet som sto ferdig år 2000 er et eksempel. Budsjettoverskridelsen som var nesten 20 prosent tilsvarte bare om lag ett års driftsbudsjettet.

Mer alvorlig blir det når selve konseptvalget viser seg å være feil, og prosjektet på sikt, etter at det er bygget, blir en fiasko. Rikshospitalet er i så måte vellykket. Men i noen tilfeller viser det seg at prosjektet er feil løsning på det problemet en står overfor, eller at det løser bare en del av problemet. Det kan også hende at det forårsaker nye problemer som overskygger forventet nytte. I enkelte tilfeller kan det til og med hende at det problemet som prosjektet skal løse ikke lenger eksisterer når prosjektet er gjennomført. Et eksempel på dette er Forsvarets torpedobatteri i Malangen som ble bygget for å kunne beskytte fiendtlige skip fra land. På det tidspunktet anlegget stod ferdig i 2004 var det politiske trusselbildet endret i en slik grad – og prosjektet innhentet av den teknologiske utviklingen på området – slik at Stortinget vedtok å stenge anlegget som kostet 250 millioner kroner bare en uke etter overlevering. Det er åpenbart lite sannsynlig at fiendtlige skip under en konflikt vil oppsøke farvannet utenfor batteriet for å bli beskyttet.

---

\* Knut Samset er professor ved NTNU, Institutt for bygg, anlegg og transport.



*Figur 1-1 Betydningen av tidligfasen. En antar at det er en sammenheng mellom usikkerhet og informasjonstilfang over tid. Muligheten for å redusere usikkerhet er størst i tidligfasen og kostnadene ved å gjøre vesentlige endringer minst. I gjennomføringsfasen er situasjonen i stor grad låst, vesentlige endringer i prosjektet vil medføre store kostnader*

Usikkerheten knyttet til beslutninger om et prosjekt endres over tid. Det er vanlig å anta at usikkerheten er størst på det tidligste tidspunktet når ideen unnfanges, og avtar raskt etter hvert som en skaffer seg nødvendig informasjon. Usikkerheten antas å flate ut mot et minimum under gjennomføringen av prosjektet, se figur 1-1. Nyttan av å skaffe til veie mer informasjon er størst innledningsvis. Muligheten for å gjøre vesentlige endringer i prosjektet kan illustreres med en liknende kurve. På ide- og konseptstadiet kan en fritt sjonglere med ulike strategiske løsninger på et gitt problem, men etter hvert vil det være slik at bordet fanger. Det blir vanskeligere, og i økende grad dyrere å gjøre vesentlige endringer i et prosjekt desto senere i prosessen en befinner seg. Eller sagt på en annen måte: De store og fundamentale spørsmålene som gjelder konseptvalget og det å finne den mest effektive løsningen på et gitt problem, er noe en må ta tak i så tidlig som mulig. Senere er det for sent. Mindre grunnleggende problemer, som for eksempel kostnadsoverskridelser, kan en nok ta tak i senere og finne en avklaring på, for eksempel når det endelige budsjettet skal fastsettes.

## 1.2 Usikkerhet og informasjonsbehov

Denne studien handler om beslutninger i den tidligste fasen av et prosjekt. Det vil si på det tidspunktet der de største avgjørelsene skal tas, men samtidig da usikkerheten er størst, muligheten for påvirkning er størst, og informasjonsgrunnlaget er mest begrenset. Intuitivt vil en tenke at mangel på informasjon må være det største problemet i en slik situasjon. Til en viss grad er dette riktig, men bare til en viss grad. Det avgjørende er hvilken type informasjon en har behov for. På dette innledende tidspunktet i prosessen er det først og fremst selve problemet og de behovene prosjektet er ment å bidra til å løse som må belyses grundig. Innledningsvis har en i mindre grad behov for detaljkunnskap om de alternative løsningene på problemet. Dette illustrerer noe som er et dilemma i alt for mange prosjekter. Det er nemlig slik at de fleste prosjekter oppstår som én konkret løsning på et problem. Selve problemet blir viet lite oppmerksomhet. Mye av informasjonsgenereringen vil være knyttet til denne konkrete løsningen. Faktisk er det slik i mange tilfeller at mengden av konkret

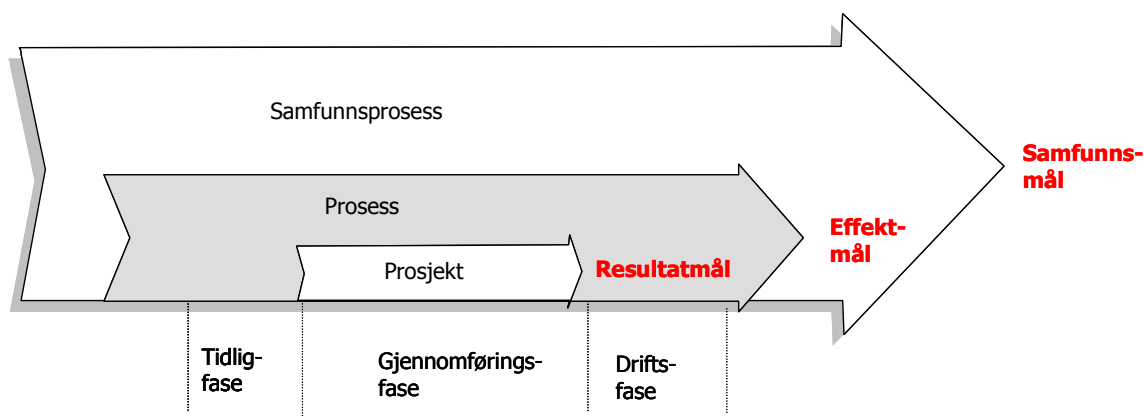


detaljinformasjon bidrar til å låse fast det opprinnelige konseptvalget, i den grad at dette i de fleste tilfeller også viser seg å bli det alternativ som tilslutt realiseres. I alt for få tilfeller har en identifisert og realitetsbehandlet andre mulige prosjektkonsepter. Spissformulert kan man si at de fleste prosjekter blir til ”i hodet på en gammel gubbe”, kanskje mer basert på intuisjon og erfaring enn på grundig innledende analyse av problemer, behov, rammebetingelser osv.

Når det gjelder informasjonsbehovet står en her overfor noe at et paradoks. En trenger mindre informasjon innledningsvis – i alle fall mindre av den typen informasjon som vanligvis produseres. Erfaring tilsier at i den tidligste delen av prosessen handler det om å forstå problemet en står overfor og problemets kontekst, hvilke behov og interesser de berørte parter har, for å finne et fornuftig strategisk grep og vurdere muligheter og risiko. Behovet innledningsvis er derfor først og fremst helhetstenkning og det å etablere perspektivet, der intuisjon, kreativitet og fantasi kan være mer verdifullt enn eksakt informasjon. En kan kanskje snu problemstillingen og hevde at mangel på informasjon i den tidligste fasen ikke bare er et problem men kan også være en styrke. For i den innledende analysen kan det være en fordel å operere mest mulig med kvalitative størrelser og bare i begrenset grad med kvantitativ informasjon.

### 1.3 Varigheten av tidligfasen og informasjonens gyldighet over tid

Investeringsprosjekter settes i gang for å løse et problem eller tilfredsstille et behov i samfunnet. Prosjektet har en leveranse som for eksempel kan være en vei, en flyplass eller et datasystem. Det er ment å ha en effekt, for eksempel i form av bedre trafikkavvikling, eller høyere produktivitet. Dette skal igjen gi samfunnsmessige gevinster for eksempel i form av økonomisk vekst. Dette er illustrert i figur 1-2 som viser at en her står overfor tre planleggingsperspektiver.

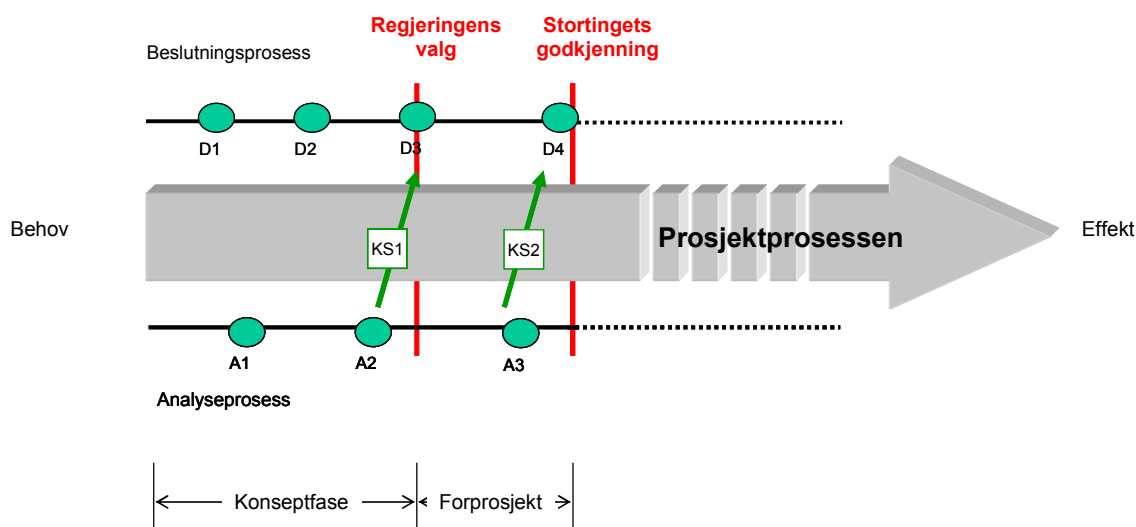


Figur 1-2 Prosjekter som en del av en større prosess vurdert i et bruker- og samfunnsperspektiv

Prosjektet er en fokusert og tidsavgrenset innsats som inngår som et første ledd i en ønsket årsak-virkning kjede: Leveransen medfører at *resultatmålet* realiseres. Dette skal medvirke til realiseringen av *effekt-målet*, som igjen skal gi et bidrag i retning av *samfunns-målet*. På bakgrunn av det som er drøftet ovenfor er det forholdsvis selvsagt at disse perspektivene må

komme inn i vurderingen på det tidligste tidspunktet i tidligfasen. Figuren viser en meget enkel inndeling i tre faser: Tidligfase, gjennomføringsfase og driftsfase. En tenker seg at tidligfasen starter med den første ideen eller det første initiativet som fører frem til prosjektkonseptet – og avsluttes med den endelige beslutningen om å finansiere dette. Dette utløser detaljert planlegging og gjennomføring. Gjennomføringsfasen avsluttes ved overlevering av leveransen.

Ser en nærmere på tidligfasen så kan det være nyttig å beskrive denne i forhold til Finansdepartementets kvalitetssikringsordning for store offentlige investeringsprosjekter. Som en ser av figur 1-3 innebærer denne at det er lagt inn to beslutningspunkter med forutgående kvalitetssikring av beslutningsunderlaget, som er utarbeidet av ansvarlig departement. Kvalitetssikringen av dette er markert med betegnelsen KS1 og KS2. Tidligfasen deles på den måten i to: En konseptfase og en forprosjektfase. Konseptfasen fører frem til valg av hovedkonsept. Dette skal skje på et så tidlig tidspunkt at valget mellom reelle konseptalternativer fremdeles står åpent. Det er selvsagt vanskelig å konkretisere og tidfeste dette nærmere i generelle termer. I norske prosjekter som omfattes av kvalitetssikringsordningen er ansvarlig departement forpliktet til å utrede minst to konseptuelle løsninger på problemet i tillegg til det såkalte nullalternativet (ikke gjøre noen). Det er regjeringen som formelt avgjør konseptvalget. Dette utløser en forprosjektfase der det foretrukne konseptalternativet utredes videre og planlegges, før Stortinget fatter sitt endelige vedtak om finansiering. Denne studien fokuserer på konseptfasen og spesielt på den tidlige delen av denne.



Figur 1-3 Finansdepartementets regime for kvalitetssikring av store offentlige investeringsprosjekter

Store investeringsprosjekter som gjelder for eksempel bygging av vei, flyplass, sykehus, anskaffelse av forsvarsmateriell, etc. har en tidligfase som strekker seg over mange år, i noen tilfeller over flere tiår. Denne perioden er karakterisert ved at den er lite forutsigbar og at utviklingen mot den endelige beslutningen skjer stegvis, som antydnet i figur 1-3. Den drives frem av en beslutningsprosess og en analytisk prosess som går parallelt, med påvirkning fra ulike interessenter, politisk tautrekking, medieeksponering, offentlig debatt, osv. En slik åpen prosess er i samsvar med våre demokratiske prinsipper og det er en vanlig oppfatning at en ikke ønsker en mer stringent, teknokratisk, ekspertdrevet prosess bygget utelukkende på rasjonalistisk planlegging. En innser samtidig at kompleksiteten og utforutsigbarheten i

prosessen også gjør at rasjonalistisk planlegging på et tidlig tidspunkt vil ha begrenset verdi over tid.

Men samtidig ser en ut fra det som er nevnt ovenfor behovet for en idéfase der intuisjon og erfaring i større grad er supplert med rasjonell analyse og vurdering. I tillegg er det åpenbart et behov for et best mulig beslutningsunderlag på de avgjørende punktene i prosessen der de viktigste beslutningene tas. Det er nærliggende å tenke seg at den mest avgjørende beslutningen gjelder selve konseptvalget, fordi dette mer enn noe avgjør prosjektets lønnsomhet og nytte over tid. Her er det et stort potensial for forbedring av dagens praksis. Imidlertid vet vi også at det opprinnelige konseptvalget kan innhentes av utviklingen og miste sin relevans i løpet av en lang tidligfase. I så fall kreves ny konseptvurdering og ny beslutning. Dette har vi heller ikke en god tradisjon på. Tvert imot vet vi at beslutningsprosessen i tidligfasen i mange tilfeller har holdt fast ved et konsept som en kunne forutsett ikke var det relevant. Torpedobatteriet i Malangen kan være et eksempel på det. Det er god grunn til å tro at begge deler kan forbedres ved å sikre et fornuftig strategisk grep på et tidligst mulig tidspunkt, samt å problematisere og eventuelt revidere dette gjennom tidligfasen før endelig beslutning om igangsetting.

Når det gjelder informasjonens gyldighet gjennom tidligfasen er den ennå mer utsatt enn selve konseptvalget. Det er nokså opplagt, og erfaring bekrefter dette, at jo mer presis informasjon, desto raskere utdateres denne over tid. En kan være fristet til å bruke et begrep som informasjonens halveringstid. Informasjon om etterspørsel i et marked i sterk utvikling kan for eksempel ha redusert sin verdi som planleggingsunderlag i løpet av dager. Samtidig er det generelt slik at gyldigheten av kvalitative vurderinger ofte har en langt bedre holdbarhet enn presis kvantitativ informasjon. Tenk for eksempel på grunnleggende oppfatninger om en brukergruppes behov. Vi vet ikke noe eksakt om hvor sterkt dette behovet er men kan være sikker på at det vil være der i lang tid fremover, og derfor kan ta det med i betraktningen i en innledende vurdering.

Det som er beskrevet ovenfor medfører ikke nødvendigvis et problem i tidligfasen. Behovet for presisjon og detaljinformasjon øker etter hvert som prosessen skrider frem, og en nærmer seg tidspunktet for detaljplanlegging og gjennomføring. Dette er i samsvar med vanlig praksis. Problemet oppstår når beslutningsprosessen druknes i detaljinformasjon på et for tidlig tidspunkt. Det kalles populært for *analysis paralysis*. Snur en om på resonnetet kan en si at nytten av eksakt informasjon vil være mindre jo lenger tilbake en går i tid.<sup>1</sup> Dette er ennå et argument for i størst mulig grad å unngå å drukne den innledende prosessen med detaljer og kvantitativ informasjon.

## 1.4 Beslutning på et svakt informasjonsgrunnlag

Det er en allmenn oppfatning at beslutningsunderlagets godhet er avgjørende for beslutninger. Imidlertid kan det være stor uenighet om hva en mener med beslutningsunderlagets godhet. Erfaring tilsier at beslutningstakere ofte stiller for små krav til beslutningsunderlaget og at viljen til å ta beslutninger er mindre påvirket av dette enn hva en skulle tro er ønskelig. Studier viser at selv der det foreligger et godt beslutningsunderlag vil mange beslutninger i

---

<sup>1</sup> Her må en selvsagt skille mellom mer eller mindre bestandig informasjon, som for eksempel fysiske data på den ene siden og økonomiske data på den annen.

større grad påvirkes av beslutningstakers intuisjon og egne eller politiske preferanser enn fakta og analyse (Henden 2004, Minzberg 2000). De praktiske implikasjonene av dette er ikke nødvendigvis katastrofale, en rekke studier har vist at intuisjon i mange situasjoner er å foretrekke fremfor rasjonell analyse, særlig der en står overfor komplekse beslutningssituasjoner slik tilfellet ofte er i store investeringsprosjekter. Forutsetningen er at intuisjon bygger på erfaring og trening. Dersom beslutningstakeren har de nødvendige forutsetninger på disse områdene kan intuisjon gi gode beslutninger langt raskere enn rasjonell analyse. Imidlertid er det slik at vi ikke kan ha en effektiv intuisjon uten å først ha en inngående kunnskap om det aktuelle temaet som behandles, noe som ofte innebærer mange års erfaring og en brukbar analyse i kombinasjon.

Når det gjelder politiske preferanser, så innebærer disse at enkelte interesser eller verdier prioriteres fremfor andre. Det er vanskelig å si noe generelt om i hvilken grad et analytisk beslutningsunderlag bør ha forrang fremfor politiske preferanser, det vil blant annet avhenge av konsekvensene av de ulike valgene i hvert enkelt tilfelle. Men det vil også være legitimt å trekke inn i vurderingen de politiske følgene av det valget som gjøres.

Ett av de avgjørende spørsmålene her er derfor i hvilken grad beslutningsunderlaget faktisk blir brukt. Ett problem i formelle beslutninger er at omfanget, kompleksitet og detaljeringsgrad i beslutningsunderlaget ofte er så stort at det i seg selv er til hinder for at det blir brukt som forventet. Sent i prosjektprosessen kan det være nødvendig med mye og detaljert informasjon. Tidlig i prosessen imidlertid, som drøftet tidligere, kan dette være en ulempe. Det som er det essensielle i den tidligste fasen er å få etablert en best mulig virkelighetsforståelse som grunnlag for å identifisere en hensiktsmessig strategi. Deretter, å få identifisert overordnede rammebetingelser som bør være førende for senere beslutninger om konseptvalg. På dette grunnlaget kan en avklare de nødvendige krav som bør være styrende for planlegging og utforming av prosjektet, det vil si den konkrete løsningen på problemet, eller det prosjektet skal levere. I en slik stegvis prosess er en tjent med å velge en tilnærming med tilsvarende tiltakende økning i detaljeringsgrad og presisjonsnivå på informasjonen. Her vil utfordringen ligge i å få tak i det essensielle, og begrense mengden av det som skal kommuniseres. Det vil gjøre kommunikasjonen bedre og øke sannsynligheten for at beslutningsunderlaget vil bli brukt, samtidig som en opprettholder den nødvendige fleksibiliteten som er nødvendig for å ta hensyn til de demokratiske prosessene og den teknologiske utviklingen. Pareto-prinsippet, som også populært kalles 80-20 regelen kan anvendes for å illustrere informasjonsbehovet. Tanken er at i et årsak-virkningsforhold er det bare en liten del av årsakene som forårsaker mesteparten av konsekvensene. Å tydeliggjøre disse blir den sentrale utfordringen.

Med de begrensningene som er nevnt ovenfor med hensyn til beslutningstakernes evne og vilje til å ta såkalt rasjonelle eller rasjonalistisk underbygde beslutninger, er det selvsagt viktig å bestrebe seg på å få frem et best mulig beslutningsunderlag på de viktigste beslutningspunktene. Det finnes åpenbart ikke én generell formel for hva som er et godt beslutningsunderlag. Imidlertid finnes det en svært enkel og generisk modell som anvendes på prosjekter og som har fått meget stor internasjonal utbredelse innenfor FN, EU og OECD, som kan illustrere hvordan informasjon kan fokuseres for å gi underlag for beslutning. Modellen identifiserer fem overordnede parametere eller beslutningskriterier som samlet skal gi et overordnet svar på et prosjekts status og prognose, og som vil utgjøre øverste lag i et hierarkisk system av informasjon for å underbygge hva en vet på de fem områdene (Samset 2002). Modellen innebærer at beslutningsunderlaget skal (1) gi en tilstrekkelig beskrivelse av

prosjektets forventete *produktivitet*, det vil si omfang og kvalitet av leveransen sett i relasjon til de ressursene inngår. Dernest (2) skal det beskrive prosjektets *måloppnåelse*, der en ikke fokuserer på leveransen men på første ordens effekt av prosjektet, se figur 1-2. Dernest (3) skal en gi en beskrivelse av prosjektets *relevans*, som er et mål på selve strategiens godhet, det vil si hvorvidt de målene en har definert reflekterer behov og prioriteringer hos brukerne og samfunnet for øvrig. I tillegg skal en (4) så langt det er mulig drøfte alle *virksomheter* som prosjektet kan tenkes å forårsake, både positive og negative, forventete og ikke-forventete virksomheter, for å kunne vurdere mulige effekter av prosjektet i et bredt perspektiv. Det siste kriteriet er (5) prosjektets *levedyktighet*, med andre ord videreføringen av prosjektet etter at leveransen er på plass og resultatmålet realisert. Dette må bygge på en finansiell vurdering av etterspørsel, inntekter, vedlikeholdskostnader, finansiell bæreevne, levetid, etc.

Om en ser disse beslutningskriteriene i sammenheng, så vil kravet til beslutningsunderlag være at det skal gi et kvalifisert svar på om prosjektet antas å være kostnadseffektivt, om målet peker i riktig retning og vil kunne realiseres, om virksomhetene forøvrig er akseptable, og om lønnsomhet og nytte på sikt kan forsvare investeringen.

For å anskueliggjøre dette kan en anvende modellen på de to prosjektene som ble nevnt innledningsvis: Torpedobatteriet i Malangen og Rikshospitalet i Oslo. Det første prosjektet hadde en lang tidligfase, men da beslutning om gjennomføring var tatt, ble byggingen gjennomført som planlagt innenfor budsjettet og med en kostnad som var innenfor hva en kunne forvente i forhold til prisen på liknende anlegg andre steder. *Produktiviteten* var derfor god. *Effekt målet*, som her er knyttet til forsvarsevne, vil ikke bli realisert ettersom anlegget aldri kommer i drift. Årsaken til dette er at prosjektet ikke lenger er *relevant* fordi trusselbildet har endret seg underveis. Prosjektet har hatt kortvarige positive *virksomheter* lokalt i form av arbeidsplasser og næringsvirksomhet, men det er usikkert hvilke virksomheter det vil ha i fremtiden. Prosjektet er ikke *levedyktig* ettersom Stortinget ikke er villig til å opprettholde driften. Anlegget må eventuelt omsettes til en pris som dekker investeringene og/eller utnyttes på en måte som gir lønnsom drift. Dette er lite sannsynlig. De fem vurderingene tatt i betraktning samlet har en et godt grunnlag for å konkludere at prosjektet er meget mislykket.

Når det gjelder Rikshospitalet, så var det her en betydelig kostnadsoverskridelse under byggingen og forsinket levering. *Produktiviteten* var altså lavere enn planlagt. Imidlertid er det liten tvil om at det er behov for sykehuset, at det produserer helsetjenester som forventet, og har hatt stort sett positive virksomheter, ikke minst på byutviklingen i det fraflyttede området sentralt i Oslo. Prosjektet scorer derfor høyt både på *måloppnåelse*, *relevans* og *virksomheter*. Når det gjelder *levedyktigheten* så er det all grunn til å tro at finansieringen er sikret i overskuelig fremtid. En hadde riktignok en kostnadsoverskridelse under byggingen, men som nevnt tidligere utgjorde den bare en mindre del av driftskostnadene. Den overordnede vurderingen av prosjektet er derfor at det er meget vellykket – til tross for begrenset produktivitet.

Denne beslutningsmodellen vil også kunne anvendes i den tidligste ide- og konseptfasen. Imidlertid er det åpenbart at ikke alle beslutningskriteriene kan anvendes i den tidligste fasen. Mye av vurderingene må bygges på antakelser og til dels på et meget svakt og usikkert informasjonsgrunnlag. Mye av den informasjonen som trenges finnes ikke og muligheten for å forutsi er begrenset eller ikke tilstede. Dette gjelder særlig tre av beslutningskriteriene, det vil si produktivitet, måloppnåelse og virksomheter. Det som imidlertid er viktigst på det aller tidligste tidspunktet er å gjøre en skikkelig vurdering av prosjektets *relevans* og *levedyktighet*. Dette er spørsmål som har å gjøre med behov, strategi, etterspørsel og bruk, og der en i mange

tilfeller har gode muligheter for å i alle fall gjøre gode kvalitative vurderinger på et tidlig tidspunkt. Det omhandler så grunnleggende spørsmål at en også kan anta at informasjonen kan ha en viss bestandighet.

Ser en på torpedobatteriet i Malangen så var problemet der først og fremst at tiltaket ikke er relevant. Derfor ble det nedlagt og derfor vil det ikke nå målet. Vurderingen av prosjektets relevans har først og fremst med trusselbildet å gjøre. Dette har endret seg i løpet av en lang tidligfase. Men en har fått mye relevant informasjon underveis og det er grunn til å tro at problemet kunne vært forutsett - om ikke på det tidligste tidspunktet da ideen ble unnfanget – så i alle fall senere dersom en hadde hatt en rutine for å vurdere planlagte prosjekters relevans med jevne mellomrom. Da hadde prosjektet blitt stoppet før det kom til utførelse. Når det gjelder Rikshospitalet, så er det forholdsvis opplagt av dette måtte vurderes som relevant og levedyktig på det tidligste tidspunktet og at denne vurderingen ikke endret seg frem til endelig vedtak om finansiering. Dette ga et godt grunnlag for å gå videre med prosjektideen.

## 1.5 Kvalitativ versus kvantitativ informasjon

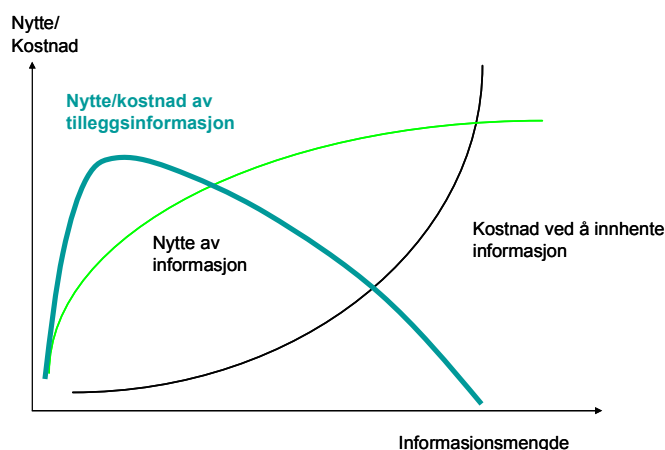
Det interessante i disse eksemplene er ikke konklusjonene, men at den modellen som anvendes representerer et opplegg for vurdering på et kvalifisert grunnlag der en i dette tilfellet kun benytter kvalitativ informasjon. Selvsagt er dette bare øverste nivå i et informasjonshierarki. Skal konklusjonene være gyldige og troverdige så må de underbygges med mer informasjon på et underliggende nivå. I praksis vil dette si at en må identifisere flere underliggende parametere eller indikatorer og skaffe informasjon for å dokumentere disse. Flere indikatorer på underliggende nivå bidrar til bedre validitet på aggregert nivå forutsatt at indikatorene er valide i forhold til det de skal måle. For å vurdere realiseringen av effektmålet i prosjektet Rikshospitalet vil det være naturlig å se på indikatorer som karakteriserer produksjon av helsetjenester, for eksempel sykehusets kapasitet og kapasitetsutnyttelse, bredden i tilbudet, godheten av helsetjenestene, etc. Lengre ned i informasjonshierarkiet kan disse indikatorene splittes opp i mer detaljerte indikatorer som til sammen gir et mer presist bilde av situasjonen. Kapasitetsutnyttelse, for eksempel, kan karakteriseres i forhold til sengekapasiteten, poliklinisk kapasitet, bruk av ulike typer utstyr, medisinsk personell, etc.

Det som karakteriserer et slikt informasjonshierarki er at det inneholder en blanding av kvalitativ og kvantitativ informasjon. Gradienten går mot mer kvalitativ informasjon jo høyere en kommer, og mer kvantitativ informasjon nedover i hierarkiet. Om en tenker denne modellen brukt langs en tidsakse, det vil i dette tilfellet si på forskjellige tidspunkter i et prosjekts tidligfase, så er det åpenbart slik at en innledningsvis i den aller tidligste fasen vil måtte operere med et hierarki som ikke stikker særlig dypt, ganske enkelt fordi informasjonen ikke finnes. Den vil måtte bygge på antakelse og derfor ha begrenset nytte. Etter hvert som prosessen skrider frem vil det være naturlig å bygge ut informasjonshierarkiet og skaffe frem mer detaljert informasjon, som i økende grad vil være kvantitativ - og i økende grad bygge på fakta og ikke bare på antakelser.

## 1.6 Informasjonens kostnad og nytte

Den informasjonen en samler inn for å belyse spørsmål om fremtidige hendelser vil alltid være ufullstendig. En mye brukt definisjon på begrepet usikkerhet som kopler fenomenet til beslutningssituasjonen, er at usikkerhet er et uttrykk for mangel på informasjon for å kunne ta den rette beslutningen. Det som først og fremst motiverer for generering av informasjon er

derfor ønsket om å redusere usikkerheten (Galbraight, 1979). Sett i forhold til prosjektprosessen antar en som vist i figur 1.1 at usikkerheten er størst innledningsvis. Dette har sammenheng med mangel på tilgjengelig og pålitelig informasjon, noe som innebærer at usikkerheten kan reduseres betydelig og raskt ved å skaffe tilveie relevant informasjon. Samtidig er det generelt slik at nytten av ny informasjon er størst på det tidspunkt der en har minst kunnskap. Dette er illustrert i figur 1.4 (Jessen, 1995). Det besnærende er at kostnaden ved å tilegne seg ny informasjon samtidig er lav relativt sett. Erfaring tilsier at kostnaden ved å skaffe ny informasjon øker nærmest eksponentielt etter hvert. Dette har sammenheng med tilgjengelighet – en må grave dypere for å finne ny informasjon – samtidig som en normalt stiller større krav til presisjon, reliabilitet og validitet etter hvert.



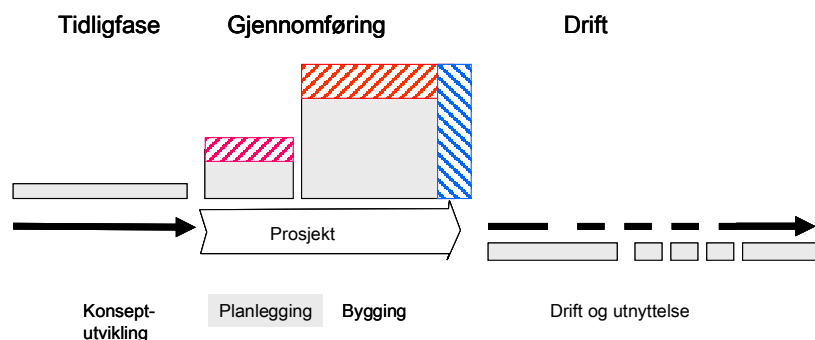
Figur 1-4. Kostnad og nytte av tilleggsinformasjon.

Men bildet er mer dramatisk enn som så, fordi erfaring viser også at nytten av ny tilleggsinformasjon avtar raskt med mengden av informasjon som allerede er tilgjengelig. Dette betyr at nytte/kostnads-forholdet er størst innledningsvis. Videre at det er gode grunner for å ikke bruke ressurser på ny informasjon når nytten blir for lav i forhold til kostnaden. Sett samlet er disse forholdene åpenbart en meget sterk motivasjon for å sikre et best mulig informasjonsunderlag på et tidlig tidspunkt i en prosjektprosess. Men det er samtidig et sterkt signal om at det er hensiktsmessig å begrense informasjonstilfanget og unngå å bruke ressurser på å skaffe frem detaljkunnskap og informasjon med for høy presisjon.

## 1.7 Tidligfasevurderingens kostnad og nytte

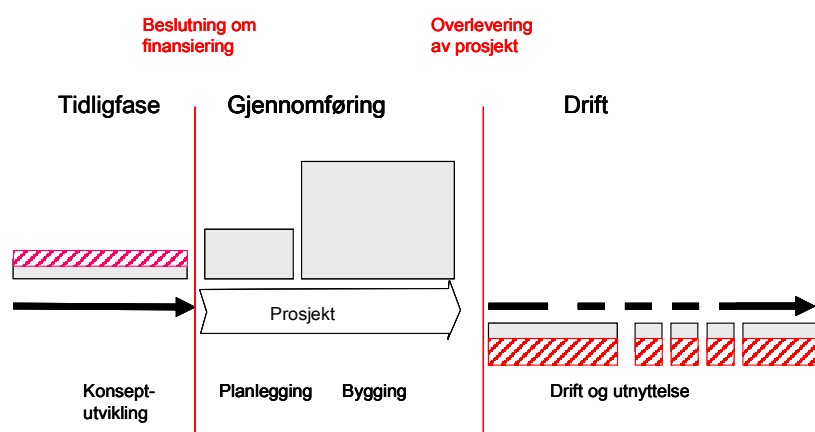
Om en nå løfter blikket fra informasjonens nytte/kostnadsfunksjon generelt, til betydningen av tidligfasevurdering sett i prosjekters livsløpsperspektiv, så kan dette illustreres prinsipielt i et tenkt eksempel i form av kontantstrømdiagrammet i figur 1-5. Som nevnt ovenfor har erfaring vist at prosjekter ofte unnfanges som én konkret løsning på et problem uten at det er gjort en forutgående systematisk analyse og vurdering av problemet, mulige strategier og alternative løsningskonsepter. De største ressursene legges inn i detaljplanleggingen i etterkant av tidligfasen når beslutning om finansiering er avklart. Formålet med denne planleggingsinnsatsen er imidlertid først og fremst å redusere usikkerheten knyttet til prosjektets gjennomføring og sikre leveransen av dette med hensyn til kostnad, kvalitet og leveringstidspunkt, som vist med de skraverte flatene i diagrammet. Dette er det mest

avgrensede perspektivet en kan ha på et prosjekt, også kalt leverandørperspektivet. Erfaringen er at i det fleste prosjekter vil nytten av den betydelige innsatsen som legges inn for å ha kontroll med kostnad og tid i gjennomføringen, være begrenset om en betrakter prosjektet i et livsløpsperspektiv.



Figur 1-5 Leverandørperspektivet på et prosjekt. Bedre planlegging for å sikre gjennomføring og leveranse

Å bryte dette mønsteret vil innebære at en må gjennomføre en mer åpen og systematisk prosess innledningsvis som fører frem til en reell vurdering av løsningskonsepter før en foretar et valg. Dette er viktig fordi det innledende konseptet oftest viser seg å være det som blir lagt til grunn senere ved planlegging og gjennomføring av prosjektet. Vurderingen vil måtte være knyttet mer til langsiktig lønnsomhet og nytte i driftsfasen enn til gjennomføring og leveranse, som vist i figur 1-6.



Figur 1-6 Finansierende parts perspektiv. Rett konseptvalg for å sikre lønnsomhet og nytte på sikt

De skraverte arealene i figuren illustrerer at en økt ressursinnsats i tidligfasen for å identifisere og velge det mest hensiktsmessige prosjektkonsept kan få en betydelig effekt i driftsfasen i form av bedret lønnsomhet og nytte. Videre er det grunn til å tro at terskelen for forbedring er lav også kostnadmessig, mens potensialet for forbedring kan være stort. Nytte/kostnads-



forholdet av tilleggsnytt sett i relasjon til tilleggsinnsatsen i tidligfasen kan derfor bli meget stort i de tilfellene der en lykkes.

Det er selvsagt her ikke et spørsmål om enten/eller, men om både/og. Det handler om å sikre at prosjektet lykkes både strategisk og taktisk. Ressursinnsatsen i tidligfasen er først og fremst en investering for å sikre prosjektets strategiske grep og langsiktige nytte. Detaljplanleggingen i gjennomføringsfasen skal sikre at prosjektet lykkes taktisk. Det som skiller de to prosjekt eksempelne som er drøftet i dette kapitlet er at Rikshospitalet er vellykket strategisk sett med mindre vellykket taktisk. Torpedobatteriet i Malangen var mislykket strategisk sett, og da hjelper det ikke om den taktiske gjennomføringen var vellykket.

## 1.8 Referanser

- Bazerman, Max, 1994, *Judgment in Managerial Decision Making*, John Wiley & Sons, Inc., New York
- Clemen, Robert, Reilly, Terence, 2001, *Making Hard Decisions with Decision Tools*, Duxbury Brooks/Cole, USA
- Galbraith, J.R., 1979, *Designing Complex Organizations*. Addison-Wesley, Reading, MA.
- Goodwin, Paul and Wright, George, 1996, *Decision Analysis for Management Judgment*, John Wiley & Sons, Inc., London
- Gullvåg, Ingemund, 1990, *Rasjonalitet, forståelse og forklaring. Innføring i argumentasjonsteori, logikk og vitenskapsfilosofi*, Tapir Akademisk Forlag, Trondheim
- Hansen, Sven Ove, 2003, *Konsten att vara vetenskaplig*, Kompendium, Filosofienheten, Kungliga Tekniska Högskolan, Stockholm
- Henden Gisle, 2004, *"Intuition and its Role in Strategic Thinking"*, PhD Thesis, Norwegian School of Management, Oslo
- Husby Otto, Kilde Halvard, Klakegg, Ole Jonny, Torp Olav, Berntsen Stein, Samset Knut, 1999, *Usikkerhet som gevinst. Styring av usikkerhet i prosjekter*, Norsk Senter for Prosjektledelse, Trondheim, 1999
- Klakegg, Ole Jonny, 2004, *Målformulering i store statlige investeringsprosjekt*, Concept rapport Nr. 6, Concept-programmet, NTNU, Trondheim
- Løwendahl Bente, Wenstøp Fred, 2002, *Grunnbok i strategi*, NKS-forlaget, Oslo
- Mintzberg, H. 1994, *The Rise and Fall of Strategic Planning*, Prentice Hall International, Hemel Hempstead/Englewood Cliffs, NJ.
- Mintzberg Henry, Lampel Joseph, Ahlstrand Bruce, 2005, *Strategy Bites Back*, *Financial Times/Prentice Hall*, UK, 2005

- Næss, Petter m.fl., 2004, *Bedre utforming av store offentlige investeringsprosjekter. Vurdering av behov, mål og effekt i tidligfasen*, Concept rapport Nr. 9, Concept-programmet, NTNU, Trondheim
- Samset, Knut, 1999, *The Logical Framework Approach (LFA), Handbook for objectives-oriented planning*, 4th edition, Norad, Oslo
- Samset, Knut, 2001, *Prosjektvurdering i tidligfase, Fokus på konseptet*, Tapir Akademisk Forlag, Trondheim
- Samset, Knut, 2003, *Project Evaluation. Making Investments Succeed*, Tapir Academic Press, Trondheim
- Samset, Knut, Berg, Peder Andreas, Klakegg, Ole Jonny, 2006, "Front-end Governance of Major Public Projects", paper presented at the 6<sup>th</sup> Annual Conference of the European Management Academy, Oslo
- Samset, Knut, 2006: Design of high-uncertainty projects in international aid, paper presentert på konferansen "PROMAC 2006", Sydney
- Wright, George, 2001, *Strategic Decision Making. A Best Practice Blueprint*, John Wiley & Sons, UK, 2001

## 2 FRA BEHOV TIL KONSEPT – OM PROSJEKTER OG PERSPEKTIVER I TIDLIGFASEN

PETTER NÆSS\*

### 2.1 Innledning

Dette kapitlet handler om behovsanalyse, målformulering og effektvurdering i den tidlige fasen av store offentlige investeringsprosjekter. Sammen med utvikling av løsningsalternativer utgjør disse aktivitetene sentrale elementer i en rasjonell planleggingsprosess. De kan utføres i ulike faser av et prosjekt. Det er imidlertid i den tidlige fasen valget av overordnet løsning – *konseptvalget* – blir foretatt. Det er som regel i den tidlige fasen mulighetene er størst for å påvirke prosjektets behovsoppfyllelse, økonomi og omfanget av sideeffekter.

Erfaringer fra en rekke store offentlige investeringsprosjekter har vist at deler av forutsetningene som har ligget til grunn for vedtak og igangsetting ikke er blitt oppfylt. Dette har medført problemer som store overskridelser på kostnadssiden, forsinkelser og lav måloppnåelse. For å unngå slike uønskede resultater, må planleggingen i den tidlige fasen av prosjektarbeidet styrkes.

På bakgrunn av identifiserte svakheter og mangler ved planleggingen av en rekke tidligere store offentlige investeringsprosjekter, inneholder dette kapitlet konkrete anbefalinger om behovsanalyser, målformuleringer, utforming av løsningskonsepter og effektvurderinger i den tidlige fasen av planarbeidet, med særlig fokus på behovsanalyser og effektvurderinger. Kapitlet bygger i stor grad på forfatterens forskning på disse temaene innenfor Concept-programmet. Lesere som ønskes mer utdypende drøftinger og analyser, henvises til rapportene fra dette arbeidet (Næss, 2004a og b).

Nedenfor følger definisjon av noen av de sentrale begrepene i kapitlet. Begrepene presenteres først med en generell ordboksdefinisjon (basert på Kunnskapsforlaget, 1992), etterfulgt av en utdyping av hvordan begrepet brukes i vår sammenheng.

**Behov:** Det å behøve, ha bruk for. Omfatter bl.a. biologiske, psykologiske, sosiale og økonomiske behov. Sosiale og økonomiske behov som ikke tilfredsstilles gjennom den enkeltes produksjon eller kjøp av varer og tjenester, men som dekkes for hele samfunnet eller store grupper i felleskap, betegnes som kollektive behov (Kunnskapsforlaget, 1996-98). De fleste store offentlige investeringsprosjekter skal tilfredsstille både kollektive og individuelle behov, ofte over flere generasjoner. Prosjektene dekker som oftest behovskomplekser sammensatt av enkeltbehov.

**Mål:** Noe som skal nås. I vår sammenheng er et mål noe en beslutningstaker søker å oppnå gjennom sin beslutning. Målsetting er et synonymt begrep, mens ambisjon, ideal, og motiv er

---

\* Petter Næss er professor ved Norsk institutt for by- og regionforskning (NIBR) og Aalborg Universitet

beslektede begreper. Et mål kan formuleres mer eller mindre generelt, kan være kvantitativt eller ikke kvantifisert, og inngår ofte som del av et hierarki av mål. Kan både omfatte prosjektets formål/hensikt og hvilke sidevirkninger av prosjektet man ønsker å oppnå eller unngå/ redusere.

**Formål:** Hensikt, det som begrunner prosjektets eksistens. Beskrives kvalitativt og kan uttrykkes som en visjon. Formålsbegrepet er snevrere enn begrepet mål, som også kan omfatte sidevirkninger av prosjektet.

**Konsept:** Grunnleggende idé. Overordnet løsning for å ivareta et sett av behov og problemstillinger som er spesifisert i form av prosjektets formål og øvrige mål.

**Fleksibilitet:** Smidighet, lettforanderlighet. Fleksibilitet er en mulig strategi for å møte usikkerhet i prosjekter, og kan dreie seg både om muligheten for endring av prosjektets utforming, kopling av alternative delprosjekter og frihet til valg av igangsettingstidspunkt for hele eller deler av prosjektet.

**Effekt:** Virkning eller følge; noe som er fremkalt av noe annet. Konsekvens er et synonym. I vår sammenheng benyttes effektbegrepet om virkningen av store offentlige investeringsprosjekter på omgivelsene. Effektene kan være både positive og negative, forventede og ikke forventede. Effekten av et prosjekt må ikke forveksles med *resultatet*, som er det ferdigstilte produktet, ressursforbruket, framdriften og gjennomføringsprosessen fram til prosjektets avslutning.

**Nytte:** Fordel, gagn, gevinst. De positive effektene av et offentlig investeringsprosjekt, sett fra ”bestillerens” (dvs. i prinsippet Stortinget og Regjeringen ved fagdepartementene) synspunkt.

**Sidevirkninger:** Sekundære virkninger av den løsningen som velges for å oppnå prosjektets hovedformål. Omfatter både ønskede (positive) og uønskede (negative) virkninger. Bivirkninger brukes av og til som synonym, særlig om uønskede virkninger. *Ringvirkninger* (virkninger som brer seg i stadig videre kretser, som ringer i vann) er indirekte virkninger av et prosjekt og brukes særlig om positive effekter for sysselsetting og økonomi.

## 2.2 Behovsanalyse, målformulering og effektvurdering som ledd i planprosessen

Behovsanalyse, målformulering og effektvurdering er selvstendige aktiviteter, men samtidig i stor grad avhengige av hverandre. Når vi sier at det er behov for et tiltak, underforstår vi at tiltaket vil ha bestemte effekter, nemlig slike som fører til behovsoppfyllelse. Når vi formulerer mål, tar vi stilling til hvilke av ulike identifiserte behov man skal forsøke å dekke. Målformulering forutsetter dessuten at tiltak med ønskede effekter kan finnes. Endelig utgjør målsettinger og identifiserte behov premisser for hvilke konsekvenser effektvurderingen skal fokusere på. Disse forholdene understreker nødvendigheten av tette koplinger mellom behovsanalyser, målformuleringer og effektvurderinger i prosjektplanleggingen. Ved utvikling av løsningskonsepter, som er avgjørende for å kunne omsette behov og mål til tiltak med ønskede effekter, er det dessuten mulighet til i større eller mindre grad å ivareta hensynet til fleksibilitet.

Ifølge den klassiske, rasjonalistiske planleggingsmodellen (Banfield, 1959) bør planleggingsarbeidet starte med en problemanalyse (trinn 1). Med denne som grunnlag formuleres mål for hvilken utvikling man ønsker på det feltet planleggingen dreier seg om (trinn 2). Deretter (trinn 3) søker man etter løsningsalternativer som kan bidra til å oppfylle målene. Her bør i prinsippet alle mulige midler for å nå målene identifiseres, ikke bare ett eller noen få. Neste skritt (trinn 4) er å forsøke å forutsi hvilke konsekvenser hvert av alternativene kan tenkes å få, både med hensyn til å oppfylle de målene som planforslaget umiddelbart er et svar på, og i forhold til andre verdier og hensyn. Tilsvarende som for løsningsalternativer, foreskriver den rasjonalistiske modellen at alle relevante konsekvenser belyses, ikke bare noen få. På bakgrunn av konsekvensanalysen foretas så (trinn 5) en mest mulig systematisk og grundig sammenlikning av alternativene. Det blir så beslutningstakernes ansvar å gjøre et valg mellom alternativene (trinn 6) og sette dem i verk (trinn 7). Helst bør det etter at planen er gjennomført, lages en etterprøving for å se om den har fått de konsekvensene man regnet med (trinn 8). Hensikten her er å vinne erfaringer som kan gi mer presise konsekvensanalyser når neste plan skal utarbeides.

I forhold til den rasjonalistiske planmodellens ulike trinn, inngår behovsanalyse som en del av arbeidet i den fasen Banfield omtaler som problemanalyse (trinn 1). Målformulering utgjør trinn 2, mens effektvurdering i prosjektets tidlige fase utgjør trinn 4 (og eventuelt også trinn 5, hvis den inkluderer formalisert sammenlikning av ulike løsningsalternativer). De effektvurderingene som gjøres i tidligfasen av store investeringsprosjekter vil alltid være *forhåndsvurderinger* av de effekter forskjellige løsningskonsepter kan forventes å få, dvs. såkalte *ex-ante* effektvurderinger. Effektvurdering kan også skje i ettertid (*ex-post*), men slike vurderinger kan naturlig nok ikke foregå i tidligfasen av prosjektarbeidet. De hører derimot hjemme i det åttende og siste av trinnene ovenfor (ved etterprøving).

Den rasjonalistiske planmodellen beskriver en omfattende og idealisert planprosess som det i praksis vil være umulig å følge fullt ut. I praksis vil man være henvist til *begrenset rasjonalitet* (Simon, 1959), bl.a. fordi tiden og ressursene som står til rådighet i planleggingsarbeidet ikke er ubegrensede. Det kan også være nødvendig å gå fram og tilbake mellom de forskjellige trinnene i stedet for å følge en lineær sekvens. Disse forbeholdene betyr likevel ikke at vi bør gi avkall på idealet om rasjonalitet i planleggingen. Selv om dette idealet ikke kan nås fullt ut, bør vi forsøke å strekke oss etter det – både for å få en god utnytting av de investeringene som foretas i store prosjekter, og for å sikre at investeringene harmonerer med de politiske målene for samfunnsutviklingen.

For å skape grunnlag for kollektiv handling er det nødvendig å forhandle, bygge enighet og skape en prosess med rom for offentlig og demokratisk drøfting av framtidige, kollektive handlinger. Befolkningsmedvirkning i planprosessene er derfor et viktig demokratisk prinsipp (Sager, 1991). Medvirkning gjør at problemene blir mer allsidig belyst og kan også gi bedre muligheter til rettferdig avveining mellom forskjellige grupperes behov og interesser. For å kunne avveie, må man vite hvilke behov de forskjellige gruppene har, og de må derfor gis mulighet til å komme fram med sine ønsker. En planprosess preget av samråd og kommunikasjon kan også ha stor betydning når det gjelder å skape bedre samarbeid og samordning innad i det offentlige byråkratiet.

For tiltak etter plan- og bygningsloven er bred offentlighet og kunngjøring om oppstart av planarbeidet et lovpålagt krav. For store statlige investeringsprosjekter bør plan- og bygningslovens formulering "fra et tidlig tidspunkt" fortolkes slik at den også omfatter behovsanalyse- og målformuleringsfasene, og ikke bare fasene der man arbeider med konkrete

løsningsalternativer og konsekvensanalyser av disse. Reell befolkningsmedvirkning er dessuten avhengig av at plandokumentene utformes på en slik måte at også ikke-eksperter kan forstå dem.

## 2.3 Planlegging på konseptnivå versus planlegging av konkrete prosjekter

Behov, mål og effekter kan defineres på ulike nivå. Jo mindre generelt behovet defineres, desto sterkere føringer vil det ligge i retning av spesifikke typer av løsninger. Figur 1.1 på neste side illustrerer hvordan både behovsidentifikasjon, målformulering, utforming av løsningsforslag og effektvurderinger kan utføres på forskjellige konkretiseringsnivå – fra overordnede samfunnsmessige vurderinger til mer detaljerte analyser i forbindelse med avgrensede prosjekter.

Denne rapporten handler om behovsvurderinger, målformuleringer og effektvurderinger *i den tidlige fasen* av store investeringsprosjekter. Dette innebærer vurdering av behov, mål og effekter på konseptnivå (dvs. *før* det er valgt et hovedkonsept for løsning), ikke på prosjektnivå (innenfor rammene av et allerede valgt hovedkonsept). Betegnelsen ”strategisk nivå” brukes ofte synonymt med det vi her har kalt konseptnivå. Forskjellen mellom prosjektnivå og strategisk nivå står mellom å analysere behovet for en bestemt teknisk løsning (et konkret prosjekt) og formulere mål for den konkrete utformingen av dette prosjektet, og å analysere de behovene og målene den tekniske løsningen har som formål å oppfylle (men som også kan tenkes oppfylt ved hjelp av andre løsninger).

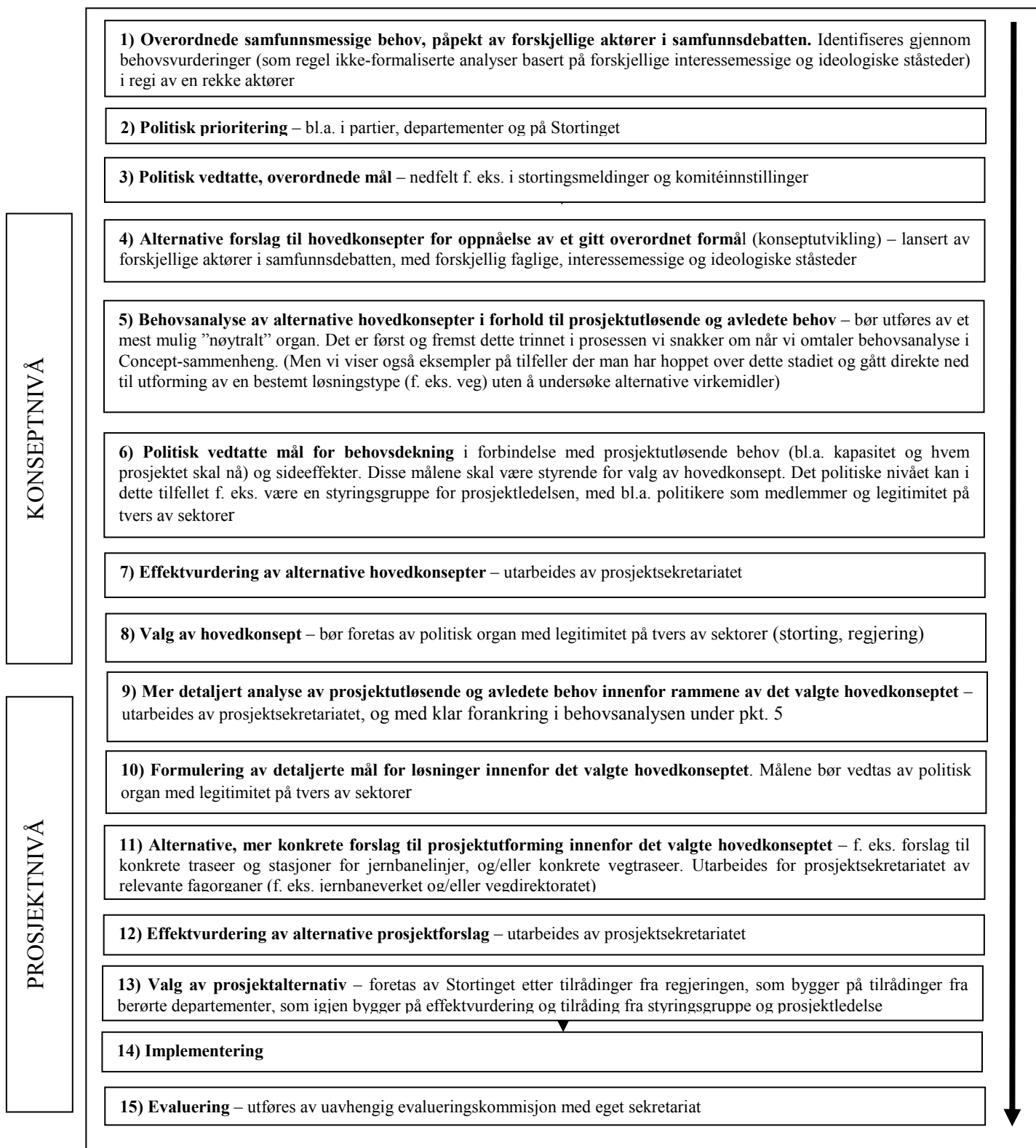
Hva dette betyr i praksis, burde framgå klart av følgende eksempel fra planlegging av et samferdselsprosjekt i et byområde med trengselsproblemer på hovedvegene: På konseptnivå kan behovene f. eks. dreie seg om lettere og mindre tidkrevende transport mellom forskjellige deler av byområdet, behov for en mer miljøvennlig transportmiddelfordeling, og behov for å oppnå et mindre transportskapende og bilavhengig utbyggingsmønster, samt om å oppnå eventuelle positive sidevirkninger og unngå/redusere negative sideeffekter av de valgte løsningene. Målene på dette nivået må avspeile disse behovene, og effektene av forskjellige løsningskonsepter (og aktuelle kombinasjoner av disse) må vurderes.

Når et gitt hovedkonsept (f. eks. en bybane) er valgt, vil behovsanalyser, målformuleringer og effektvurderinger derimot dreie seg om å sikre at dette hovedkonseptet blir utformet og gjennomført på en samfunnsmessig best mulig måte. Behovene og målene på dette nivået kan f. eks. dreie seg om å oppnå høye passasjertall, økonomisk gunstig og miljøvennlig linjeføring, og å bidra (gjennom plasseringen av stasjoner) til byutvikling i områder der dette er ønsket.

Hvis behovsanalyser, målformuleringer og effektvurderinger på det strategiske nivået utelates, og man i stedet går direkte i gang med planlegging på prosjektnivå innenfor rammene av et gitt løsningskonsept, er det fare for at prosjektinitiativtakernes behov forveksles med samfunnets behov, og at særinteressers ønske om økonomisk gevinst, prestisje eller ideologisk foretrukne løsninger får forrang framfor overordnede politiske målsettinger og bredere samfunnsgruppers behov.

Store investeringsprosjekter har dessuten som regel effekter og berører behov ut over de behovene og effektene som i utgangspunktet utløser ideen til prosjektet. Analysene må ikke

begrenses til de *prosjektuløsende* behovene og effektene, men også ta med sidevirkninger og behov som blir berørt av disse.



Figur 2-1 Behovsidentifikasjon (trinn 1, 5 og 9), målformulering (trinn 3, 6 og 10), løsningsforslag (trinn 4 og 11) og effektvurdering (trinn 7 og 12) på ulike konkretiseringsnivåer. De planaktivitetene det fokuseres på i dette kapitlet, er på strategisk plannivå og omfatter trinnene 5, 6 og 7. Trinn 8 og 11 er beslutninger om hhv. konseptvalg og prosjektvalg. Skjemaet er en stilisert framstilling av en "rasjonell" plan- og beslutningsprosess og avspeiler ikke nødvendigvis den faktiske saksgangen i forbindelse med store, statlige investeringsprosjekter.



## 2.4 Hvorfor kommer store investeringsprosjekter ofte skjevt ut?

Nedenfor vil vi se på forskjellige typer feil, mangler og fordreininger som kan opptre i planleggingen i den tidlige fasen av store offentlige investeringsprosjekter: Fokus på prosjektnivå i stedet for på konseptnivå, manglende oppmerksomhet omkring sideeffekter, neglisjering av overordnede politiske målsettinger, feilaktige forutsigelser av vurderte effekter og manglende åpenhet i prosessen.

**Fokus på prosjektnivå i stedet for på konseptnivå.** Som nevnt ovenfor, er det viktig at planleggingen i den tidlige fasen av store investeringsprosjekter ikke starter på et for lavt nivå i kjeden av behov, mål og midler. Da risikerer man at store investeringsprosjekter ikke bidrar til å oppfylle de prioriterte samfunnsmessige målene. Det finnes imidlertid en rekke norske og utenlandske eksempler på at behovsanalysene framstiller en bestemt teknisk løsning som behovet, og at målene og effektvurderingene innsnevres til å dreie seg om gjennomføringen av et gitt hovedkonsept.

Moen og Strands (2000) undersøkelse av praktiseringen av pkt. 3.5 i de rikspolitiske retningslinjene for samordnet areal- og transportplanlegging (RPR-SAT) viser f. eks. en rekke eksempler på at en bestemt løsning (vegbygging) er blitt framstilt som behovet, i stedet for behovet for god tilgjengelighet mellom de funksjonene og aktivitetene som forbindes av de foreslåtte vegparsellene. I disse tilfellene er den snevre fokuseringen på vegbygging direkte i strid med de rikspolitiske retningslinjene, som fastslår at:

Når kapasitetsproblemer i vegsystemet oppstår, skal andre alternativer enn økt vegkapasitet vurderes på lik linje, f. eks. regulering av trafikk, forbedring av kollektivtransporttilbudet. (RPR-SAT, pkt. 3.5, 2. ledd)

Av 11 undersøkte vegprosjekter i geografiske områder som omfattes av de rikspolitiske retningslinjene, fant Moen og Strand at planoppgavene i fem av tilfellene ble avgrenset til prosjektnivå til tross for at de ifølge de rikspolitiske retningslinjene for samordnet areal- og transportplanlegging skulle vært definert på strategisk nivå. Å starte planleggingen av samferdselsprosjekter på prosjektnivå uten innledende, bredere analyser på strategisk nivå, er ikke et fenomen begrenset til vegprosjekter. Analysene som skulle belyse behovet for Gardermobanen fokuserte f. eks. ensidig på tog som det kollektive transportmidlet, og neglisjerte busstransporten. Resultatet var en trafikkprognose som overdrev det framtidige passasjertallet på flytoget med 67 %, og som dermed også gav et overdrevet inntrykk av behovet for Gardermobanen. Andre alternativer for å styrke kollektivtransporten mellom Oslo og lufthavnen, f. eks. styrket ekspressbusstrafikk med lav billettpris, ble ikke utredet (NOU 1999:28).

Tilsvarende viser Klakeggs (2004) gjennomgang av målformuleringer i store statlige prosjekter etter at kvalitetssikringsordningen ble innført i 2000, at målformuleringene i 5 av 8 prosjekter begrenset seg til å handle om hvordan et allerede fastlagt hovedkonsept best kunne gjennomføres. I noen av prosjektene var målformuleringene knapt mer enn en beskrivelse av den valgte løsningen. Blant de tre prosjektene der målformuleringen også var relevant for andre tenkelige løsninger enn den valgte, og således rent formelt tilfredsstilte kriteriene for en målformulering på strategisk nivå, ble det dessuten i minst ett tilfelle ikke utformet og effektvurdert mer enn ett løsningskonsept.

**Manglende oppmerksomhet omkring sideeffekter.** I forbindelse med store investeringsprosjekter er det viktig å skille mellom de behovene som i utgangspunktet utløser det foreslåtte tiltaket, og andre behov og effekter som det samtidig vil være nødvendig å ta hensyn til hvis prosjektet realiseres. Et tiltak som springer ut av et gitt behov, kan ha positive så vel som negative sideeffekter. Befolkningsgruppene som berøres av slike effekter, er ikke alltid de samme som har nytte av prosjektets hovedformål. I et samfunnsmessig perspektiv vil det være behov for å oppnå så store positive og så små negative sideeffekter som mulig, samtidig som det behovet prosjektet i utgangspunktet er begrunnet med, oppfylles på en tilfredsstillende måte. Denne flersidigheten må fanges opp både i prosjektets behovsanalyser, målformuleringer og i effektvurderingene av alternative løsningskonsepter.

I planleggingen av store offentlige investeringsprosjekter blir likevel andre behov, mål og effekter enn dem som knytter seg til prosjektets hovedformål, ofte neglisjert. Klakeggs (2004) gjennomgang av målformuleringer i store statlige prosjekter etter at kvalitetssikringsordningen ble innført i 2000, viser f. eks. at det for fire av åtte undersøkte prosjekter overhodet ikke var formulert mål i tilknytning til sideeffekter. For tre andre prosjekter omfattet målformuleringene positive sideeffekter (bl.a. reduksjon av støy og ulykker), mens begrensning av negative sideeffekter ikke inngikk blant målene. I de tilfellene der miljøkonsekvenser av prosjektene tas med i målformuleringer og effektvurderinger, er disse ofte begrenset til lokale og umiddelbare virkninger på miljøet (Røpke, 1993; Nielsen, 2000). Indirekte miljøkonsekvenser, f. eks. som følge av at nye veger ofte skaper mer trafikk, blir ofte neglisjert i konsekvensanalyser av infrastrukturplaner, til tross for at nødvendigheten av å inkludere slike konsekvenser som regel framheves i litteraturen om strategiske miljøvurderinger (se bl.a. European Commission, 2001). Jonsson (2004) har undersøkt konsekvensanalysene i 24 regionale infrastrukturplaner og funnet at bare tre av disse oppfylte kriteriet om akseptabel dekning og analyse av indirekte effekter.

Det forekommer også at positive sidevirkninger av prosjektene overvurderes. I forbindelse med mange transportinfrastrukturprosjekter i utkantområder argumenteres det f. eks. med prosjektenes positive virkninger for lokal økonomi og sysselsetting. En rekke studier har imidlertid vist at slike forventninger sjelden er velbegrunnede (Flyvbjerg et al., 2003:65-72; Engebretsen, Lian & Strand, 1998). Derimot forekommer det at mulige positive sidevirkninger av alternativer til den foreslåtte løsningen neglisjeres i planleggingen (Næss, 2004a).

**Neglisjering av overordnede politiske målsettinger.** Store offentlige investeringsprosjekter gjennomføres innenfor sektorer der det som regel alltid er formulert overordnede politiske mål for samfunnsutviklingen. Ofte vil dessuten sektorovergripende mål, f. eks. når det gjelder miljøvern og bærekraftig utvikling, være relevante for prosjektene. Det skjer likevel forholdsvis ofte at behovsanalyser og målformuleringer utelater viktige og relevante samfunnsmessige behov og målsettinger. Spesielt er dette uheldig i situasjoner der det er politiske mål om en annen utvikling enn den dagens trender og markedsetterspørsel peker mot.

Det finnes en rekke norske og utenlandske eksempler på at utformingen av transportinfrastrukturprosjekter bygger på trafikkmodellberegninger som tar utgangspunkt i den forventede etterspørselen etter plass på hovedvegene (Moen & Strand, 2000; Tennøy, 2003; Næss, 2004a). De gjeldende transportpolitiske målsettingene i Norge og en rekke andre OECD-land (bl.a. Storbritannia) handler imidlertid om at etterspørselen etter mer biltrafikk i de største byene skal dempes ("predict and prevent" i stedet for "predict and provide"). Det

framgår bl.a. av stortingsmeldingene om Bedre miljø i byer og tettsteder og Bedre kollektivtransport (St. meld. nr. 23 og 26, 2001-2002). Ved å basere planleggingen på etterspørselsbaserte modellberegninger og trendforlengelser i slike situasjoner, undergraves de politiske målene for samfunnsutviklingen.

Blant de forskjellige kategoriene av store offentlige investeringsprosjekter, er problemet med neglisjering av overordnede politiske målsettinger ikke minst tydelig i forbindelse med samferdselsprosjekter. Dette er prosjekter innenfor en samfunnssektor der de politiske målene ofte peker i en annen retning enn trendutviklingen, og der en stor gruppe av prosjektene har en institusjonell forankring som medfører en tendens til å dreie fokuset mot én bestemt type løsning. I den grad samfunnsmessige behov reduseres til individenes etterspørsel eller betalingsvillighet ved planlegging av andre typer store offentlige investeringer, vil imidlertid tilsvarende – eller enda større – problemer kunne melde seg. Det ville f. eks. neppe blitt gjennomført særlig mange store investeringer i militært materiell hvis disse investeringene skulle gjøres avhengige av dokumentert betalingsvillighet blant befolkningen.

**Feilaktige forutsigelser av vurderte effekter.** I forbindelse med planleggingen av store offentlige investeringsprosjekter har det ofte vist seg i ettertid at behovsanalyser, kostnadsoverslag, finansielle analyser, miljøkonsekvensanalyser og analyser av sosiale virkninger var beheftet med alvorlige feil. Dette framgår bl.a. av en omfattende studie av over 200 internasjonale store transportinfrastrukturprosjekter (Flyvbjerg et al., 2003). Flyvbjerg og hans kolleger fant generelt store avvik mellom de prognostiserte og de faktiske trafikkmengdene. Særlig var avviket stort for jernbaneanlegg. For mer enn 9 av 10 slike anlegg overvurderte prognosene for passasjertall det faktiske antallet passasjerer. Gjennomsnittlig var det faktiske passasjerbelegget under halvparten av det prognostiserte. Også for vegprosjekter var trafikkprognosene lite treffsikre, idet halvparten av prosjektene hadde et avvik på mer enn 20 % mellom prognostisert og faktisk trafikkvolum, med omtrent like mange tilfeller av overvurdering som undervurdering av de framtidige trafikkmengdene. For en mindre gruppe av prosjektene ble også effektvurderingenes forutsigelser om miljøkonsekvenser sammenliknet med den faktiske utviklingen for de miljøparametrene som inngikk i analysene. Det var her en klar tendens til at miljøkonsekvensene ble mer omfattende enn forutsagt (Nielsen, 2000). Enkelte viktige typer miljøkonsekvenser var dessuten utelatt i analysene. Et interessant og alarmerende funn i Flyvbjerg og hans kollegers materiale er at et stort antall av de prosjektene der etterspørselsanalysene var overoptimistiske, også hadde overoptimistiske kostnadsoverslag. Særlig var dette typisk for bybaneprosjekter. Blant de 12 bybaneprosjektene i undersøkelsen var det gjennomsnittlige, faktiske passasjertallet 48 % lavere enn prognostisert, mens kostnadene gjennomsnittlig lå 40 % over de tallene nytte-kostnadsanalysene bygde på. Samlet innebar dette at forholdet mellom passasjertall og kostnader for de undersøkte bybaneprosjektene gjennomsnittlig ble feilvurdert med en faktor på mer enn 2,5.

Noen tilsvarende omfattende undersøkelse som gir grunnlag for statistisk analyse av prognostetreffsikkerhet i norske prosjekter foreligger ikke. Enkeltstående evalueringer av forskjellige store norske investeringsprosjekter peker imidlertid i samme retning. Gardermobanen, etterbruken av idrettsanlegg fra Lillehammer-OL og en rekke IKT-prosjekter i statlig så vel som fylkeskommunal og kommunal regi er eksempler på prosjekter der de implementerte løsningene bygde på analyser som i større eller mindre grad var mangelfulle.

Manglende treffsikkerhet i analysene kan i en del tilfeller være et resultat av dårlig faglig arbeid, f. eks bruk av mangelfulle analysemodeller eller sviktende kunnskap om relevante

årsak-virkningsmekanismer. Ifølge Flyvbjerg et al. (2003) forekommer det imidlertid urovekkende ofte at planleggerne lager prognoser som systematisk framstiller prosjekter i et mer fordelaktig lys enn det er saklig grunnlag for. I noen tilfeller skyldes slik ”vurderingsoptimisme” en allmennmenneskelig tendens til ønsketenkning. Flyvbjerg et al.’s materiale tyder imidlertid på at prognoseunøyaktighet i forbindelse med store infrastrukturprosjekter i minst like stor grad skyldes partiskhet i konsulentenes eller initiativtakernes vurderinger. Wachs (1989) hevder at planleggere ikke alltid er opptatt av å lage så nøyaktige og pålitelige prognoser som mulig, men i stedet av å få prosjekter finansiert og bygd. Realistiske prognoser er ofte ikke noe effektivt middel til å oppnå det siste.

**Manglende åpenhet i prosessen.** I mange situasjoner vil det ikke være tilstrekkelig å identifisere behov, formulere mål og vurdere effekter ”for samfunnet sett under ett”, men være nødvendig å inkludere flere typer av behov (basert på forskjellige verdimesseige utgangspunkter og forskjellige interessegruppers ståsted). De mange konfliktsakene man har opplevd i forbindelse med store investeringsprosjekter, kan tyde på at planleggingen i prosjektenes tidlige fase har vært for lukket, slik at man i for liten grad har vært klar over og i stand til å ta hensyn til forskjellige interesseparters behov. Dette gjelder særlig ved planlegging av utbygging av transportinfrastruktur (se bl.a. Medalen, 1987), kraftutbygging (se bl.a. Berntsen, 1994) og skyte- og øvingsfelter (se bl.a. Aftenposten, 2002; NRK Hordaland, 2003), men også i forbindelse med andre store statlige investeringsprosjekter (f. eks. Statoils administrasjonsbygg på Rotvoll i Trondheim). ”Lukkede” planprosesser risikerer ikke bare å skape unødige store konflikter i forhold til interessegrupper i sivilsamfunnet eller næringslivet, men kan også skape unødvendige motsetninger innad i den offentlige forvaltningen.

En åpning av planprosessen så den allerede i den tidlige prosjektfasen omfatter aktører utenfor den etaten som har hovedansvaret for planleggingen, er imidlertid ikke tilstrekkelig til å unngå neglisjering av viktige gruppers interesser, hvis medvirkningen begrenses til å omfatte de sterkeste og mest artikulerte ”stakeholder-gruppene”. For å sikre at behovsanalysen fanger opp alle relevante gruppers behov, må også underprivilegerte og mindre taleføre grupper gis innpass i prosessen. Mange studier (bl.a. Flyvbjerg et al., 2003; Medalen, 1987; Gaardmand, 1996) viser imidlertid at planarbeidet i store investeringsprosjekters tidlige fase ofte har sviktet på disse punktene.

En annen type åpenhet dreier seg om utforming av plandokumentene slik at det framgår klart hvilke forutsetninger de anbefalte løsningene bygger på. Erfaringen fra Gardermobanen (Riksrevisjonen, 1998) viser imidlertid at det kan være vesentlig forskjell på hvordan prognoser presenteres opprinnelig og hvordan de blir brukt i det endelige beslutningsgrunnlaget. I tilfellet med Gardermobanen utførte en verifiseringsgruppe en peer review av prognosene. Verifiseringsgruppen markerte kraftig at det var stor usikkerhet i prognosene. Dette ble ikke omtalt i beslutningsgrunnlaget for Gardermobanen, St prp nr 90 (1991-92).

## 2.5 Nytte-kostnadsanalyser egner seg dårlig for prosjekter der bådenyttefaktorene og kostnadene er usikre og omstridte

Ulike sammenlikningsmetoder legger i forskjellig grad opp til å aggregere løsningskonseptenes egenskaper til et begrenset antall indikatorer. Lengst i så måte går nytte-kostnadsanalysen, der prosjektets ulike positive og negative egenskaper sammenfattes i netto nåverdi, nytte-kostnadsbrøk eller en annen indikator på samfunnsøkonomisk lønnsomhet. Fordelen med en høy grad av aggregering er at det gir en oversiktlig presentasjon av analysens hovedkonklusjon. Ulempen er at mange nyanser og detaljer går tapt. Særlig i tilfeller der fordelene og ulempene berører ulike befolkningsgrupper, eller der verdsettingen av en del av effektene er etisk vanskelig og/eller omstridt, kan aggregering til en enkelt eller noen få indikatorer være problematisk. I slike situasjoner vil høy grad av aggregering kunne føre til maskering av viktige spørsmål som burde være gjenstand for politisk diskusjon i stedet for å fremstå som en del av et ”nøytralt” faglig regnestykke.

Blant de forskjellige typene av store investeringsprosjekter er nytte-kostnadsanalyser særlig utbredt i samferdselssektoren. Slike analyser er imidlertid av flere grunner dårlig egnet for denne typen prosjekter, både fordi de krever en høyere grad presisjon i forutsigelsene av effekter enn det er mulig å oppnå, og fordi det i høy grad er disputt om hvordan de kvantifiserte størrelsene skal ”oversettes” til kroner og øre og omregnes til nåverdi og nyttekostnadsbrøk. Jeg har diskutert dette mer omfattende i en engelsk tidsskriftartikkel som fokuserer på vitenskapsteoretisk uholdbare sider ved dagens bruk av nytte-kostnadsanalyser i transportsektoren (Næss, 2006). Nedenfor vil jeg bare stikkordsmessig nevne noen av de problematiske sidene ved slike analyser.

I nytte-kostnadsanalyser reduseres samfunnsmessig verdi til et aggregat av individuelle forbrukeres betalingsvillighet. Men mennesker er ikke fullt informerte forbrukere som bare er opptatt av å maksimere nytte. Det som er viktig og godt på samfunnsnivå, er noe annet og mer enn summen av individuelle preferanser slik disse kommer til uttrykk gjennom betalingsvillighet. Hvis ikke, ville det for eksempel ikke vært nødvendig med offentlig planlegging og styring.

Analysene krever presise forutsigelser av effekter, men i praksis er forutsigelsene ofte meget upresise, ikke minst på transportområdet der den viktigste nytten ofte er summen av de minuttene eller sekundene i reisetid som en ny transportåre antas å spare trafikantene for. Disse forutsigelsene skjer som regel ved hjelp av transportmodeller. Slike modeller bygger imidlertid ofte på ufullstendige og til dels misvisende antakelser om relevante årsakssammenhenger som påvirker trafikkutviklingen (Kenworthy, 1990; Sactra, 1999; Arge et al., 2000; Tennøy, 2003).

Prosjektene økonomiske nytteverdi er enda vanskeligere å fastslå. Anslag over effekter f. eks. i form av endret transportaktivitet (bl.a. endret trafikkmengde og endrede reisetider) vil måtte bygge på eksisterende teoretisk og empirisk kunnskap om aktuelle årsakssammenhenger, og bør utstyres med feilmarginer pga. kunnskapens ikke-eksakte natur og forskjeller mellom den lokale konteksten og de situasjonene de benyttede empiriske forskningsresultatene stammer fra. Det vil deretter være nødvendig å ”oversette” anslagene over trafikkendringer til pengeverdier, bl.a. basert på verdien av spart tid. Det er imidlertid kontroversielt hvordan verdien av tid bør beregnes. Med mindre man ønsker å favorisere en bestemt oppfatning, må

beregningen av tidsverdi skje ut fra ulike måter å regne på som avspeiler uenigheten. Anslagene over trafikale konsekvenser måtte så kombineres med tidsverdiene til et økonomisk anslag over prosjektets direkte nytteeffekt. På grunn av usikkerheten som knytter seg både til trafikkanslagene og tidsverdiene ville det bli snakk om et anslag med temmelig store feilmarginer.

I enda større grad vil usikkerhet og uenighet kunne gjøre seg gjeldende ved økonomisk verdsetting av prosjektets sideeffekter, f. eks. miljøskader. Dette skyldes ikke minst uenighet om hvorvidt det er rimelig å neddiskontere langsiktige, irreversible og alvorlige miljøskader, og hvilken kalkulasjonsrente det i så fall er rimelig å benytte.

Ved å basere vurderingen av hvorvidt et prosjekt er gunstig for samfunnet på betalingsvillighet og neddiskontering av framtidige effekter, kan nytte-kostnadsanalyser generelt sies å fremme en markedsbasert utvikling (Ackerman & Heinzerling, 2004). Dette er problematisk i situasjoner der det oppstår viktige ”eksternaliteter” og prosjektene påvirker fordelingen av goder og ulemper mellom ulike befolkningsgrupper. Ikke minst er denne innvendingen aktuell for prosjekter innenfor samferdselssektoren, der det som nevnt er formulert offentlige mål om å erstatte den hittil dominerende ”predict and provide”-tilnærmingen med andre strategier enn å dekke markedsetterspørselen etter vegkapasitet.

## 2.6 Hva kan gjøres for å forbedre situasjonen?

På bakgrunn av identifiserte svakheter og mangler ved en rekke tidligere store prosjekter, inneholder Concept-rapporten om planlegging i den tidlige fasen av store offentlige investeringer (Næss, 2004b) konkrete anbefalinger om behovsanalyser, målformuleringer, utforming av løsningskonsepter og effektvurderinger. Noen hovedpunkter fra disse anbefalingene er gjengitt nedenfor.

### Generelle anbefalinger

Store offentlige investeringsprosjekter gjennomføres innenfor sektorer der det som regel alltid er formulert overordnede politiske mål for samfunnsutviklingen. Ofte vil dessuten sektorovergripende mål, f. eks. når det gjelder miljøvern og bærekraftig utvikling, være relevante for prosjektene. Et tiltak for å unngå neglisjering av andre behov, mål og effekter enn dem som knytter seg til prosjektets hovedformål, kan være å kreve at behovsanalyser, målformuleringer og effektvurderinger struktureres med separate avsnitt om behov, mål og vurdering av effekter i tilknytning til henholdsvis prosjektets hovedformål, negative sidevirkninger og positive sidevirkninger.

Behov, mål og effekter kan defineres på ulike nivå. Jo mindre generelt behovet defineres, desto sterkere føringer vil det ligge i retning av spesifikke typer av løsninger. Det er viktig at planleggingen i den tidlige fasen av store investeringsprosjekter ikke starter på et for lavt nivå i kjeden av behov, mål og midler. Man risikerer i så fall at prosjektene ikke bidrar til å oppfylle de prioriterte samfunnsmessige målene.

### Bedre behovsanalyser

For å sikre at store investeringsprosjekter bidrar til å oppfylle de prioriterte samfunnsmessige behovene, og ikke bare behovene slik de oppfattes innenfor en avgrenset sektor eller blant

tilhengerne av en bestemt teknisk løsning, må det utføres en behovsanalyse i den tidlige fasen av prosjektforløpet, *før* det tas beslutning om valg av hovedkonsept.

Behovsanalyser kan utføres med utgangspunkt i forskjellige forståelser av behovsbegrepet. De kan også utføres innenfor mer eller mindre objektivistiske versus fortolkningsbaserte perspektiver (Hiebert & Smallwood, 1987). Situasjonen ved planlegging av store, statlige investeringsprosjekter preges ofte av tvetydighet, dynamiske omgivelser, uklarhet omkring behovenes innhold, mål og midler, og uklarhet omkring avgrensningen av analysen.

Metodene for behovsanalyser kan inndeles i tre hovedkategorier: Normative, etterspørselsorienterte og interessegruppebaserte (McKillip, 1987). Ved å kombinere flere metoder, kan en metodes mangler og svakheter bli kompensert av en annen metodes styrke. Metodene som velges, bør til sammen gjøre det mulig å identifisere alle relevante behov – både ”prosjektutløsende” behov og behov i forbindelse med sideeffekter – og vurdere disse behovene mht. viktighet og hvilke befolkningsgrupper de gjelder.

Behovsanalysen må fange opp alle relevante samfunnsmessige behov, og ikke bare behov som kommer til uttrykk gjennom individers betalingsvillighet. På grunn av den samfunnsmessige karakteren av de behovene det er snakk om i forbindelse med store, offentlige investeringsprosjekter, vil det være hensiktsmessig å ta utgangspunkt i overordnede, politisk vedtatte målsettinger på relevante områder. Vi vil imidlertid advare mot å basere behovsanalysen på normtall knyttet til bestemte tekniske løsninger. Både for å sikre tverrfaglighet, for å sikre ulike sektorer innflytelse på analysene og fordi behovene varierer mellom befolkningsgrupper, bør behovsanalysen organiseres med deltakelse fra forskjellige sektorer (og eventuelt forvaltningsnivåer) innenfor offentlig administrasjon, så vel som fra forskjellige befolkningsgrupper.

I behovsanalyser som tar utgangspunkt i overordnede offentlige målsettinger, bør prognoser basert på dagens utviklingstrekk (”forecasting”) ikke stå alene. Ved å se hvilken framtidssituasjon dagens utvikling leder til på et gitt saksområde, og sammenlikne denne tilstanden med en normativt ønskelig situasjon, kan man belyse behovet for tiltak som kan endre utviklingen (”backcasting”, se Dreborg, 1996).

I behovsanalyser for transportinfrastrukturprosjekter bør hensiktsmessigheten av å benytte transportmodellberegninger vurderes kritisk, se nærmere omtale nedenfor i avsnittet om effektvurderinger.

Ved presentasjonen av behovsanalysens hovedkonklusjoner må alle vesentlige behov tas med, både prosjektutløsende behov og behov knyttet til sideeffekter, og tallfestede så vel som ikke-kvantifiserte behov. Dokumentasjonsmaterialet må på en lett tilgjengelig måte gjøre rede for de kriterier, antagelser og vektlegginger som konklusjonene bygger på.

### **Bedre målformuleringer**

Mål i forbindelse med offentlige prosjekter kan inndeles i tre kategorier, forankret i ulike aktørgruppers perspektiv: *samfunns mål*, som dreier seg om overordnede, samfunnsmessige hensyn på tvers av ulike gruppers behov, *effekt mål*, som reflekterer brukernes perspektiv, og *resultat mål*, som er formulert ut fra leverandørens perspektiv (Samset, 2001). Målformuleringene i den tidlige prosjektfasen skal være styrende for valg av hovedkonsept og omfatter samfunns- og effekt mål. Målene skal reflektere de prosjektrelevante og politisk

prioriterte behovene som et foreslått investeringsprosjekt tar sikte på å tilfredsstille, og bør både omfatte mål knyttet til prosjektets hovedformål og mål knyttet til viktige sidevirkninger av prosjektet.

Store statlige investeringsprosjekter er komplekse og har gjerne mange mål som i større eller mindre grad er avhengige av hverandre. For at sammenhengene og avhengighetene skal komme klart fram, anbefales å framstille et *målhierarki*. Målhierarkiet utgjør en nedbryting av overordnede mål i delmål som blir stadig mer konkrete jo lavere en befinner seg i målhierarkiet. Målene må ikke være logisk uforenlige, men kan i varierende grad komme i konflikt med hverandre, avhengig av hvilke konkrete løsninger som velges. Målene på samme nivå i hierarkiet kan ha forskjellig viktighet. Dette kan uttrykkes ved å gi målene forskjellige tallmessige vekt. For at en slik innbyrdes prioritering mellom mål skal ha politisk legitimitet, må den foretas av et relevant politisk beslutningsorgan.

*Hovedmålene* for prosjektet er en konkretisering av hva formålet eller intensjonen betyr for samfunnet og brukerne, samt hvilke sideeffekter det er viktig å oppnå eller unngå/begrense. Hvert hovedmål må være knyttet til virkningen av det foreslåtte prosjektet, det må være relevant, og det må ha en viss bestandighet. Hovedmålene skal være ambisiøse, men realistisk oppnåelige og ha en formulering uten tvetydigheter eller forutinntatthet mht. løsning. De må dessuten i tilstrekkelig klargjøre grad hva som skal oppnås, hvor mye og når. (Klakegg, 2004.)

*Delmålene* er en konkretisering av hva som skal oppnås gjennom prosjektet. Alle delmålene skal være hensiktsmessige bidrag til realisering av prosjektets hovedmål. Delmålene bør oppfylle kravene i akronymet SMART, dvs. være spesifiserte, målbare, aksepterte, realistiske og tidsrelaterede (Klakegg, ibid.). Kravet om målbarhet betyr ikke at alle delmål nødvendigvis må være tallfestede, og må ikke føre til at viktige og relevante mål utelates.

### **Utvikling av alternative konseptløsninger**

Konseptutvikling innebærer å identifisere, utrede og konkretisere mulige løsningskonsepter. Arbeidet med konseptutviklingen bør skje i tverrfaglige team. Planlegging på konseptnivå innebærer at prinsipielt ulike måter å dekke behovene og oppfylle målene på, blir belyst. Det er derfor avgjørende at genuint forskjellige løsningskonsepter utvikles. Også nullalternativet må beskrives.

Prosessen med å omforme konseptidéer til konkrete løsningskonsepter bør fortsette helt til prosjektteamet anser at de har kommet fram til en utforming som gir en rimelig god realisering av hver hovedidé's potensial for måloppnåelse og behovsdekning. Det må advares mot å arbeide seriøst med utviklingen av bare ett av konseptene, mens konkurrerende konsepter behandles overflatisk. Løsningskonseptet ville i så fall i realiteten være valgt på forhånd, og arbeidet på strategisk plannivå ville bli redusert til en pro forma-aktivitet for å legitimere det foretrukne løsningskonseptet.

Usikkerhet omkring et prosjekts forutsetninger og konsekvenser innebærer at en beslutning om å gjennomføre prosjektet medfører en viss risiko, økonomisk så vel som mht. måloppnåelse. Fleksibilitet er en mulig strategi for å møte usikkerheten. Graden av fleksibilitet bør være noe som besluttes samtidig med prosjektets øvrige rammevilkår. En slik forhåndsplanlagt fleksibilitet kan oppnås ved å utforme konseptalternativer som er fleksible mht. framdriftsplan og sammensetning av komponenter. Dette kan for eksempel gjelde store utbyggingsprogrammer som inneholder flere mer eller mindre avhengige delprosjekter som



dels bygger på hverandre og dels er gjensidig utelukkende. Løsningene som velges, kan dessuten i varierende grad være fleksible mht. bruksmåter og muligheter for ombygging.

### **Bedre effektvurderinger**

Formålet med effektvurderinger i tidligfasen av prosjekter er å gi grunnlag for prioritering mellom forskjellige konseptalternativer. Effektene av et stort investeringsprosjekt omfatter både virkninger i forhold til prosjektets primære formål og sideeffekter. De omfatter både positive og negative, forventede og ikke forventede effekter som kan tilbakeføres til prosjektet, direkte eller indirekte. Effektene kan også slå ulikt ut for ulike befolkningsgrupper og dermed berøre viktige fordelingsproblemer.

Den kunnskapen det er aktuelt å trekke veksler på ved slike effektvurderinger, er sjelden kontekststøttet. Den må derfor tilpasses og modifiseres i forhold til den aktuelle situasjonen. Både av denne grunn og fordi samfunnet er i stadig endring, vil forutsigelser om investeringsprosjekteneffekter på menneskers handlingsmønstre aldri kunne bli særlig nøyaktige.

Det er nødvendig å ha kriterier for å velge mellom forskjellige handlingsalternativer og beslutte om et prosjekt skal realiseres. Disse kriteriene bør gjøres så eksplisitte som mulig, og ikke hemmeligholdes i de enkelte planleggernes og beslutningstakeres hoder. Behovet for formelle vurderingsmetoder innebærer imidlertid ikke at de metodene som i størst grad baserer seg på kvantifisering av effektene, nødvendigvis er å foretrekke.

Beslutningstre, kostnadseffektivitetsanalyse, kostnads-virkningsanalyse og nytte-kostnadsanalyse er eksempler på metoder for å sammenligne alternativer og belyse hvilket som er å anbefale. Hvilke metoder som bør benyttes, avhenger bl.a. av hvor betydningsfulle fordelingsvirkningene er, hvor mange konsekvenser vi finner det faglig forsvarlig å verdsette i kroner, og hvor stort innslag det er av etisk vanskelige spørsmål (Finansdepartementet, 2000). Ofte vil en kombinasjon av flere metoder være formålstjenlig. Effektvurderingen må imidlertid ikke innsnevres til de effektene som lettest lar seg måle eller som det med rimelighet kan settes prislapp på.

Hensikten med effektvurderinger må være å klargjøre, ikke å tilsløre. Litt grovt kan man spørre seg om nytten ved å bruke nytte-kostnadsanalyser i tidligfase-planleggingen av store offentlige investeringsprosjekter er større enn kostnadene i form av tilsløring av usikkerhet og en tendens til å overse virkninger som er vanskelige å prissette. I stedet for å streve etter å komme fram til ett enkelt talluttrykk for hvordan nytten av et prosjekt står i forhold til kostnadene, er det etter mitt syn bedre å presentere ulike hovedgrupper av konsekvenser hver for seg. De ulike effektene bør behandles på det målenivå som passer best for den enkelte effekt. Direkte utbyggingskostnader og nytteeffekter som det gir mening i å uttrykke i økonomiske termer, bør prissettes med angivelse av feilmarginer. Effekter som er vanskelige å måle, eller der det man måler ikke har entydig sammenheng med konsekvensen for de verdienne effekten berører, bør gjøres til gjenstand for mer kvalitative beskrivelser, eller vurderes ut fra en kombinasjon av kvalitative og kvantitative indikatorer. Også i slike tilfeller vil en viss grad av kvantifisering av den samlede effekten være mulig, f. eks. innplassering på en skala for måloppnåelse.

I effektvurderinger i forbindelse med transportinfrastrukturprosjekter (og i behovsvurderinger i forbindelse med slike prosjekter, jf. ovenfor) bør hensiktsmessigheten av å benytte

transportmodellberegninger vurderes nøye. Slike modellberegninger bør i alle tilfelle ikke stå alene, men benyttes som innspill til bredere analyser. Transportmodellberegninger bør etter mitt syn bare brukes hvis modellene er kvalitetssikret av uavhengige eksperter fra en fagkrets som spenner videre enn modellmakernes egen. Et alternativ kan være å basere trafikkprognosene på en enklere og mer transparent vurdering av forskjellige faktorer som kan bidra til å øke eller redusere trafikken. Særlig i den tidlige fasen av prosjektarbeidet bør dette være aktuelt. En slik vurdering, basert på foreliggende, tverrfaglig viten om relevante påvirkningsfaktorer (og hver påvirknings størrelsesorden), vil i mange tilfeller gi vel så pålitelige resultater som avanserte matematiske modellberegninger. Dette fordi noen relevante påvirkninger uansett vil måtte anslås temmelig grovt, også hvis de brukes som input i en trafikkmodellberegning. Jo flere grove anslag som inngår blant modellforutsetningene, dess mindre meningsfylt er det å foreta sofistikerte matematiske beregninger med sikte på størst mulig nøyaktighet for de lettest målbare påvirkningene.

Usikkerhets- og følsomhetsanalyser kan gi en indikasjon om hvordan handlingsalternativene vil slå ut under ulike omstendigheter for forskjellige interesseparter. Basert på slike analyser bør man utvikle strategier for hvordan ulike situasjoner – især de ugunstigste – kan møtes. Usikkerhet og risiko kan også reduseres ved å utsette irreversible beslutninger, f. eks. om iverksetting av hele eller deler av prosjektet. Realopsjonsanalyse er et hjelpemiddel til å vurdere hvilket gjennomføringstidspunkt som er gunstig i situasjoner med usikkerhet omkring viktige forutsetninger. Denne metoden kan også brukes til å vurdere opsjonsverdier som et foreslått prosjekt vil utløse (f. eks. muligheter for lønnsomme tilleggsinvesteringer) eller berøre negativt (f. eks. muligheter for naturopplevelse) hvis det blir gjennomført (Brekke, 2004).

### **Når planleggerne er en del av problemet – eller ”the dark side of planning”**

Som jeg allerede har nevnt, viser datamaterialet vi har hatt tilgang til at det ofte er en systematikk i de forskjellige typene feil og skjevheter i behovsanalyser og effektvurderinger. De bidrar som regel til å støtte opp om en bestemt konklusjon, nemlig at investering i den foreslåtte fysiske løsningen er ønskelig og nødvendig. Materialet styrker en mistanke om at det forholdsvis ofte skjer en mer eller mindre bevisst fordreining av analysene for å få foreslåtte prosjekter til å framstå i så fordelaktig lys som mulig.

En mulig forklaring på tendensen til at prognosene gir et mer positivt bilde enn det viser seg å være grunnlag for, kan være en allmennmenneskelig tendens til ønsketenkning. En analyse Flyvbjerg og hans medarbeidere har gjort av prosjektledernes og evalueringsforskernes oppfatninger om årsaker til unøyaktighet og skjevheter i prognosene i forbindelse med transportinvesteringsprosjekter (Flyvbjerg, Holm og Buhl, 2005), tyder imidlertid på at også partiskhet i konsulentenes eller initiativtakernes vurderinger kan være en forholdsvis utbredt årsak til prognosefeil, spesielt for jernbaneprosjekter. For hele 25 % av baneprosjektene nevner informantene *bevisst tendensiøse prognoser* som årsak til prognoseunøyaktigheten. Mistanken om bevisste fordreininger for å få gjennomført prosjekter som ellers ikke ville oppnådd politisk støtte, styrkes av at overvurderte prognoser for passasjerbelegg ofte ledsages av kraftig undervurderte kostnadsoverslag. For vegprosjekter følger avvikene mellom prognostisert og faktisk trafikk et mer komplisert mønster. En sammenlikning av trafikkprognoser for danske vegprosjekter påbegynt før 1980 med tilsvarende prosjekter fra perioden 1980-90 kan imidlertid tyde på at det forekom strategisk undervurdering av

trafikprognosene i 1980-årene, da Danmark la større vekt enn de fleste andre europeiske land på å begrense forbruket av bensin (Næss, Flyvbjerg & Buhl, 2006).

Den såkalte referanseprognosemetoden kan være et middel til å unngå overoptimistiske eller fordreide anslag over utbyggingskostnader eller den framtidige bruken av et investeringsprosjekt. Metoden krever at det foreligger et tilstrekkelig stort antall sammenliknbare prosjekter til at det lar seg gjøre å utføre statistiske analyser av hva som er ”normale” utfall mht. kostnader eller antall brukere. Metoden innebærer at man i stedet for å utlede den framtidige etterspørselen eller utbyggingskostnadene ”innenfra”, dvs. ut fra en betraktning av prosjektets tekniske egenskaper og tenkelige forhold som kunne påvirke behovet, anlegger et ”utenfraperspektiv”. Et ”utenfraperspektiv” innebærer å sammenlikne det aktuelle prosjektet med en gruppe av tilsvarende prosjekter og utnytte foreliggende informasjon om disse. Ifølge Flyvbjerg et al. (2005) har forskning (Gilovich, Griffin & Kahneman, 2002) vist at referanseprognosemetoden gir mer treffsikre og pålitelige resultater enn prognoser basert på tradisjonelle metoder. Dette fordi referanseprognosemetoden unngår kognitive og organisatoriske fordreininger som f.eks. ”prosjektbedømmelsesoptimisme” (appraisal optimism) og strategisk skjev framstilling, men i stedet fokuserer direkte på utfallet.

For å motvirke at planleggere ”lyver med tall” er det nødvendig å unngå at ”bukken passer havresekken”. Analysene bør utføres av et mest mulig nøytralt organ, med bred, tverrfaglig bemanning. Folkemøter, borgerjuryer og liknende bør organiseres for å gjøre det mulig for interessegrupper og sivilsamfunnet å komme med kritikk av eller støtte til analysene. Analysene bør bli gjenstand for uavhengig faglig vurdering (peer review). Faglige og eventuelt også strafferettslige sanksjoner bør dessuten rettes mot planleggere som gjentatte ganger produserer grovt misvisende prognoser (Flyvbjerg et al., 2003). Større offentlighet og uavhengig fagkritikk av analysene vil også kunne bidra til å motvirke mangelfull kartlegging av hvilke relevante behov og effekter som forekommer.

## 2.7 Referanser

- Ackerman, F. & Heinzerling, L. (2004): *Priceless: On Knowing the Price of Everything and the Value of Nothing*. New York: New Press.
- Aftenposten (2002): ”Nytt skytefelt i skuddlinjen.” *Aftenposten*, 24.5.02. <http://www.aftenposten.no/nyheter/iriks/politikk/article.jhtml?articleID=337149>
- Arge, N.; Homleid, T & Stølan, A. (2000): *Modeller på randen... Bruk av transportmodeller i norske byområder. En evaluering*. Oslo: LOKTRA-prosjekt.
- Banfield, E. C. (1959): “Ends and means in planning.” *International Social Science Journal*, Vol. XI, no. 3, 1959.
- Berntsen, B. (1994): *Grønne linjer. Natur- og miljøvernets historie i Norge*. Oslo: Grøndahl Dreyer.
- Brekke, K. A. (2004): *Realopsjoner og fleksibilitet i store offentlige investeringsprosjekter*. Concept-rapport 1050-4. Oslo: Frischsenteret for samfunnsøkonomisk analyse
- Dreborg, K. H. (1996): “Essence of backcasting.” *Futures*, 28(9), 813-828.

- Engebretsen, Ø.; Lian, O. & Strand, S. (1998): *Samferdsel og robuste bo- og arbeidsmarkedsregioner*. TØI arbeidsdokument TR/0802/1998. Oslo: Transportøkonomisk institutt.
- European Commission – DG Energy and Transport (2001): *Manual on Strategic Environmental Assessment of Transport Infrastructure Plans*. Amersfoort: DHV Environment and Infrastructure BV.
- Finansdepartementet (2000): *Veiledning i samfunnsøkonomiske analyser*. Oslo: Finansdepartementet/Statens forvaltningstjeneste.  
[http://odin.dep.no/filarkiv/116093/Samf\\_oe.pdf](http://odin.dep.no/filarkiv/116093/Samf_oe.pdf)
- Flyvbjerg, B., Bruzelius, N. and Rothengatter, W. (2003): *Megaprojects and Risk. An Anatomy of Ambition*. Cambridge: Cambridge University Press.
- Flyvbjerg, B., Holm, M. S. & Buhl, S. (2005) “How (In)accurate Are Demand Forecasts in Public Works Projects? The Case of Transportation.” *Journal of the American Planning Association*, 71, pp. 131-146.
- Gilovich, T.; Griffin, D. & Kahneman, D. (2002) (eds.): *Heuristics and Biases: The Psychology of Intuitive Judgement*. Cambridge: Cambridge University Press.
- Gaardmand, A. (1996): *Magt og medløb. Om mahognibordsmetoden og den korporative planlægning*. København: Arkitektens Forlag.
- Hiebert, M. B. & Smallwood, W. M. (1987): “Now for a completely different look at needs analysis.” *Training and Development Journal*, Vol. 41, pp. 75-79
- Jonsson, D. (2004): *Granskning av regionala transportinfrastrukturplaner med avseende på indirekta miljöeffekter*. Rapport från Naturvårdsverket. Stockholm.
- Klakegg, O. J. (2004): *Målformulering i store statlige investeringsprosjekt*. Concept-rapport nr. 1050-2. Trondheim: NTNU
- Kenworthy, J. (1990): “Don’t Shoot Me – I’m Only the Transport Planner”, in Newman, P.; Kenworthy, J. & Lyons, T. *Transport Energy Conservation*. Perth: Murdoch University.
- Kunnskapsforlaget (1992): *Aschehoug og Gyldendals store norske ordbok*. Oslo: Kunnskapsforlaget.
- Kunnskapsforlaget (1996-98): *Kunnskapsforlagets store norske leksikon*. Oslo: Kunnskapsforlaget
- McKillip, J. (1987): *Need Analysis. Tools for the Human Services and Education*. Series: Applied Social Research Methods, Vol. 10. Newbury Park/Beverly Hills/London/New Delhi: Sage Publications.

- Medalen, T. (1987): *Konflikter i vegplanlegging: en beskrivelse og analyse av konflikter i hovedplanprosesser*. Dr. ing.-avhandling. Trondheim: Norges tekniske høgskole, Institutt for by- og regionplanlegging.
- Moen, B. & Strand, A. (2000): "Når kapasitetsproblemer i vegnettet oppstår, skal andre...". Prosjektrapport 2000:1. Oslo: Norsk institutt for by- og regionforskning.
- Nielsen, K. L. (2000): *Environmental Appraisal of Large Scale Transport Infrastructure Projects.* Ph.D.-avhandling. Aalborg: Aalborg Universitet, Istitut for Samfundsudvikling og Planlægning.
- NOU 1999:28 Gardermoprojektet. *Evaluering av planlegging og gjennomføring*. Utredning fra en gruppe oppnevnt ved kongelig resolusjon av 15. mai 1998 <http://odin.dep.no/sd/norsk/publ/utredninger/NOU/028005-020002/hov001-nu.html>
- NRK Hordaland (2003): *Ber Kristin stanse bombeplaner. Voss kommunestyre vedtok i dag enstemmig å be forsvarsministeren stanse bombeplanene på Mjølfjell*. NRK Hordaland, 20.6.03.
- Næss, P. (2006) "Cost-benefit analyses of transportation investments: neither critical nor realistic." *Journal of Critical Realism*, Vol. 5, No. 1, 2006.
- Næss, P.; Flyvbjerg, B. & Buhl, S. (2006): "Do road planners produce more 'honest numbers' than rail planners? An analysis of accuracy in road-traffic forecasts in cities vs. peripheral regions." Forthcoming in *Transport Reviews*, Vol. 26, 2006.
- Næss, P. (2004a): *Bedre behovsanalyser. Erfaringer og anbefalinger om behovsanalyser i store offentlige investeringsprosjekter*. NIBR-rapport 2004:15. Oslo: Norsk institutt for by- og regionforskning.
- Næss, P. (2004b): *Tidlig fase i store offentlige investeringsprosjekter. Vurdering av behov, mål, fleksibilitet og effekt*. Med bidrag fra O. J. Klakegg, N. Olsson og K. A. Brekke. NIBR-rapport 2004:22.
- Riksrevisjonen (1998): «Riksrevisjonens undersøkelse vedrørende hovedflyplassens og Gardermobanens bedriftsøkonomiske lønnsomhet», *Dokument nr. 3:10*, 1997-1998.
- Røpke, I. (1993): "Hvor mange "lag"? –miljøkonsekvenser i strukturelt perspektiv." I Kullinger, B. og Strömberg, U.-B. (red.): *Planera för en bärkraftig utveckling*, s. 133-144. Stockholm: Byggeforskningsrådet.
- SACTRA, 1999: *Transport and the Economy*, London: Standing Advisory Committee on Trunk Road Appraisal.
- Sager, T. (1991): *Planlegging med samfunnsperspektiv*. Analysemetode. Trondheim: Tapir forlag.
- Samset, K. (2003): *Project Evaluation. Making investments succeed*, Trondheim, Tapir Akademisk Forlag

Simon, H. (1959): "Administrative behavior". *Public Administration Review*, Vol. XI

St.meld. nr. 23 (2001-2002) *Bedre miljø i byer og tettsteder*. Oslo: Miljøverndepartementet.  
<http://odin.dep.no/md/norsk/publ/stmeld/022001-040015/dok-bn.html>

St.meld. nr. 26 (2001-2002) *Bedre kollektivtransport*. Oslo: Samferdselsdepartementet.  
<http://odin.dep.no/sd/norsk/publ/stmeld/028001-040007/dok-bn.html>

St.prp. nr. 90 (1991-92) *Utbygging og finansiering av hovedflyplass for Oslo-området på Gardermoen med tilhørende tilbringersystem og konsekvenser for Forsvaret*.

Tennøy, A. (2003): *Bidrar bruk av transportanalyser i byplanleggingen til vekst i biltrafikken?* Paper til Trafikdage på Aalborg universitet, 25. – 26. august 2003.

Wachs. M. (1989): "When Planners Lie With Numbers." *Journal of American Planning Association*, Vol. 55, no. 4, pp. 476-479.

# 3 HOW OPTIMISM BIAS AND STRATEGIC MISREPRESENTATION IN EARLY PROJECT DEVELOPMENT UNDERMINE IMPLEMENTATION

BENT FLYVBJERG\*

## 3.1 Introduction

For a number of years the research program on large infrastructure at Aalborg University, Denmark, has explored different aspects of the planning and implementation of large infrastructure projects (Flyvbjerg, Bruzelius, and Rothengatter, 2003; Flyvbjerg, Holm, and Buhl, 2002, 2004, 2005; Flyvbjerg and Cowi, 2004; Flyvbjerg, 2005a, 2005b).<sup>2</sup> This paper takes stock of what may be learned from the research so far, especially as regards optimism bias and strategic misrepresentation in the early phases of project development.

First I will argue that a major problem in the planning of large infrastructure projects is the high level of misinformation about costs and benefits that is produced during the early stages of project development and the high risks such misinformation generates. Second I will explore the causes of misinformation and risk, mainly in the guise of optimism bias and strategic misrepresentation. In a subsequent paper, I will present a number of measures aimed at improved planning and decision making, including changed incentive structures and better planning methods (Flyvbjerg, 2007).

The emphasis will be on transportation infrastructure projects. It should be mentioned at the outset, however, that comparative research shows that the problems and causes identified for transportation apply to a wide range of other project types including power plants, dams, water projects, concert halls, museums, sports arenas, convention centers, IT systems, oil and gas extraction projects, aerospace projects, and weapons systems (Flyvbjerg, Bruzelius, and Rothengatter, 2003: 18-19; Flyvbjerg, Holm, and Buhl, 2002: 286; Flyvbjerg, 2005a; Altshuler and Luberoff, 2003).

---

\* Bent Flyvbjerg er professor ved Aalborg Universitet, Department of Development and Planning ([bentflyvbjerg@stofanet.dk](mailto:bentflyvbjerg@stofanet.dk)), og ved Delft University of Technology, Nederland.

<sup>2</sup> "Large infrastructure projects" are here defined as the most expensive infrastructure projects that are built in the world today, typically at costs per project from around a hundred million to several billion dollars.

## 3.2 Characteristics of Large Infrastructure Projects

Large infrastructure projects, and planning for such projects, generally have the following characteristics (Flyvbjerg and Cowi, 2004):

- Such projects are inherently risky due to long planning horizons and complex interfaces.
- Technology is often not standard.
- Decision making and planning is often multi-actor processes with conflicting interests.
- Often the project scope or ambition level change significantly over time.
- Statistical evidence shows that such unplanned events are often unaccounted for, leaving budget contingencies inadequate.
- As a consequence, misinformation about costs, benefits, and risks is the norm.
- The result is cost overruns and/or benefit shortfalls with a majority of projects.

## 3.3 Size and Frequency of Cost Overruns and Benefit Shortfalls

For transportation infrastructure projects, Table 1 shows the inaccuracy of construction cost estimates measured as the size of cost overrun. The cost study covers 258 projects in 20 nations on five continents. All projects for which data was obtainable were included in the study.<sup>3</sup> For rail, average cost overrun is 44.7 percent measured in constant prices. For bridges and tunnels, the equivalent figure is 33.8 percent, and for roads 20.4 percent. The difference in cost overrun between the three project types is statistically significant, indicating that each type should be treated separately (Flyvbjerg, Holm, and Buhl, 2002).

The large standard deviations shown in Table 3-1 are as interesting as the large average cost overruns. The size of the standard deviations demonstrate that uncertainty and risk regarding cost overruns are large, indeed.

---

<sup>3</sup> The data are from the largest database of its kind. All costs are construction costs measured in constant prices. Cost overrun, also sometimes called "cost increase" or "cost escalation," is measured according to international convention as actual out-turn costs minus estimated costs in percent of estimated costs. Actual costs are defined as real, accounted construction costs determined at the time of project completion. Estimated costs are defined as budgeted, or forecasted, construction costs at the time of decision to build. For reasons explained in Flyvbjerg, Holm, and Buhl (2002) the figures for cost overrun presented here must be considered conservative. Ideally financing costs, operating costs, and maintenance costs would also be included in a study of costs. It is difficult, however, to find valid, reliable, and comparable data on these types of costs across large numbers of projects. For details on methodology, see Flyvbjerg, Holm, and Buhl (2002).



The following key observations pertain to cost overruns in transportation infrastructure projects:

- 9 out of 10 projects have cost overrun.
- Overrun is found across the 20 nations and 5 continents covered by the study.
- Overrun is constant for the 70-year period covered by the study, estimates have not improved over time.

*Table 3-1 Inaccuracy of transportation project cost estimates by type of project, in constant prices.*

Type of project	No. of cases (N)	Avg. cost overrun %	Standard deviation
Rail	58	44.7	38.4
Bridges and tunnels	33	33.8	62.4
Road	167	20.4	29.9

Table 3-2 shows the inaccuracy of travel demand forecasts for rail and road projects. The demand study covers 208 projects in 14 nations on five continents. All projects for which data was obtainable were included in the study.<sup>4</sup> For rail, actual passenger traffic is 51.4 percent lower than estimated traffic on average. This is equivalent to an average overestimate in rail passenger forecasts of no less than 105.6 percent. The result is large benefit shortfalls for rail. For roads, actual vehicle traffic is on average 9.5 percent higher than forecasted traffic. We see that rail passenger forecasts are biased, whereas this is not the case for road traffic forecasts. The difference between rail and road is statistically significant at a high level. Again the standard deviations are large, indicating that forecasting errors vary widely across projects (Flyvbjerg, Holm, and Buhl, 2005; Flyvbjerg, 2005b).

The following observations hold for traffic demand forecasts:

- 84 percent of rail passenger forecasts are wrong by more than  $\pm 20$  percent.
- 9 out of 10 rail projects have overestimated traffic.
- 50 percent of road traffic forecasts are wrong by more than  $\pm 20$  percent.
- The number of roads with overestimated and underestimated traffic, respectively, is about the same.
- Inaccuracy in traffic forecasts are found in the 14 nations and 5 continents covered by the study.
- Inaccuracy is constant for the 30-year period covered by the study, forecasts have not improved over time.

---

<sup>4</sup> Following international convention, inaccuracy is measured as actual traffic minus estimated traffic in percent of estimated traffic. Rail traffic is measured as number of passengers; road traffic as number of vehicles. The base year for estimated traffic is the year of decision to build. The forecasting year is the first full year of operations. Two statistical outliers are not included here. For details on methodology, see Flyvbjerg (2005b).

It must be concluded that if techniques and skills for arriving at accurate cost and traffic forecasts have improved over time, these improvements have not resulted in an increase in the accuracy of forecasts.

*Table 3-2 Inaccuracy in forecasts of rail passenger and road vehicle traffic.*

Type of project	No. of cases (N)	Avg. inaccuracy %	Standard deviation
Rail	25	-51.4	28.1
Road	183	9.5	44.3

If we combine the data in Table 3-1 and Table 3-2, we see that for rail an average cost overrun of 44.7 percent combines with an average traffic shortfall of 51.4 percent.<sup>5</sup> For roads, an average cost overrun of 20.4 percent combines with a fifty-fifty chance that traffic is also wrong by more than 20 percent. As a consequence, cost benefit analyses and social and environmental impact assessments based on cost and traffic forecasts like those described above will typically be highly misleading.

Cost benefit analyses and social and environmental impact assessments are typically central in the early phases of project development, for instance in pre-feasibility and feasibility studies. Ideally, decisions in early-phase development should be informed. Given the data presented above, we see that in reality decisions are often misinformed. This presents a major risk to those paying for projects and hoping to benefit from their impacts.

### **3.4 Projects With Cost Overruns and Benefit Shortfalls**

The list of examples of projects with cost overruns and/or benefit shortfalls is seemingly endless (Flyvbjerg, 2005a). Boston's Big Dig, otherwise known as the Central Artery/Tunnel Project, were 275 percent or US\$11 billion over budget in constant dollars when it opened, and further overruns are accruing due to faulty construction. Actual costs for Denver's \$5 billion International Airport were close to 200 percent higher than estimated costs. The overrun on the San Francisco-Oakland Bay Bridge retrofit was \$2.5 billion, or more than 100 percent, even before construction started. The Copenhagen metro and many other urban rail projects worldwide have had similar overruns. The Channel tunnel between the UK and France came in 80 percent over budget for construction and 140 percent over for financing. At the initial public offering, Eurotunnel, the private owner of the tunnel, lured investors by telling them that 10 percent "would be a reasonable allowance for the possible impact of unforeseen circumstances on construction costs."<sup>6</sup> Outside of transportation, the \$4 billion cost overrun for the Pentagon spy satellite program and the over \$5 billion overrun on the

<sup>5</sup> For each of twelve urban rail projects, data exist for both cost overrun and traffic shortfall. For these projects average cost overrun is 40.3 percent; average traffic shortfall is 47.8 percent.

<sup>6</sup> Quoted from "Under Water Over Budget," *The Economist*, October 7, 1989, 37-38.

International Space Station are typical of defense and aerospace projects. Our studies show that large infrastructure and technology projects tend statistically to follow a pattern of cost underestimation and overrun. Many such projects end up financial disasters. Unfortunately, the consequences are not always only financial, as is illustrated by the NASA space shuttle. Here, the cooking of budgets to make this under-performing project look good on paper has been linked with shortchanged safety upgrades related to the deaths of seven astronauts aboard the Columbia shuttle in 2003 (Flyvbjerg 2004).

As for benefit shortfalls, consider Bangkok's US\$2 billion Skytrain, a two-track elevated urban rail system designed to service some of the most densely populated areas from the air. The system is greatly oversized, with station platforms too long for its shortened trains. Many trains and cars sit in the garage, because there is no need for them. Terminals are too large, etc. The reason is that actual traffic turned out to be less than half that forecast (Flyvbjerg, Holm, and Buhl, 2005: 132). Every effort has been made to market and promote the train, but the project company has ended up in financial trouble. Even though urban rail is probably a good idea for a dense, congested, and air-polluted city like Bangkok, overinvesting in idle capacity is hardly the best way to use resources, especially in a developing nation in which capital for investment is particularly scarce. Such benefit shortfalls are common and have also haunted the Channel tunnel, the Los Angeles and Copenhagen metros, and Denver's International Airport.

Other projects with cost overruns and/or benefit shortfalls are, in North America: the F/A-22 fighter aircraft; FBI's Trilogy information system; Ontario's Pickering nuclear plant; subways in numerous cities, including Miami and Mexico City; convention centers in Houston, Los Angeles, and other cities; the Animas-La Plata water project; the Sacramento regional sewer system renewal; the Quebec Olympic stadium; Toronto's Sky Dome; the Washington Public Power Supply System; and the Iraq reconstruction effort. In Europe: the Eurofighter military jet, the new British Library, the Millennium Dome, the Nimrod maritime patrol plane, the UK West Coast rail upgrade and the related Railtrack fiscal collapse, the Astute attack submarine, the Humber Bridge, the Tyne metro system, the Scottish parliament building, the French Paris Nord TGV, the Berlin-Hamburg maglev train, Hanover's Expo 2000, Athens' 2004 Olympics, Russia's Sakhalin-1 oil and gas project, Norway's Gardermø airport train, the Øresund Bridge between Sweden and Denmark, and the Great Belt rail tunnel linking Scandinavia with continental Europe. In Australasia: Sydney's Olympic stadiums, Japan's Joetsu Shinkansen high-speed rail line, India's Sardar Sarovar dams, the Surat-Manor toll way project, Calcutta's metro, and Malaysia's Pergau dam. The list ends here only for reasons of space.

This is not to say that projects do not exist for which costs and/or benefits were on or better than the budget, even if they are harder to find. For instance, costs for the Paris Southeast and Atlantique TGV lines were on budget, as was the Brooklyn Battery tunnel. The Third Dartford Crossing in the UK, the Pont de Normandie in France, and the Great Belt road bridge in Denmark all had higher traffic and revenues than projected. Finally, the Bilbao Guggenheim Museum is an example of that rare breed of projects, the cash cow, with costs on budget and revenues much higher than expected.<sup>7</sup>

---

<sup>7</sup> For an explanation of the success of the Bilbao Guggenheim Museum, see Flyvbjerg (2005a).

### 3.5 Why Cost Overruns and Benefit Shortfalls Are a Problem

Cost overruns and benefit shortfalls of the frequency and size described above are a problem for the following reasons:

- They lead to a Pareto-inefficient allocation of resources, i.e., waste.
- They lead to delays and further cost overruns and benefit shortfalls.
- They destabilize policy, planning, implementation, and operations of projects.
- The problem is getting bigger, because projects get bigger.

Let's consider each point in turn. First, an argument often heard in the planning of large infrastructure projects is that cost and benefit forecasts at the planning stage may be wrong, but if one assumes that forecasts are wrong by the same margin across projects, cost-benefit analysis would still identify the best projects for implementation. The ranking of projects would not be affected by the forecasting errors, according to this argument. However, the large standard deviations shown in Table 3-1 and Table 3-2 falsify this argument. The standard deviations shows that cost and benefit estimates are not wrong by the same margin across projects; errors vary extensively and this will affect the ranking of projects. Thus we see that misinformation about costs and benefits at the planning stage is likely to lead to Pareto-inefficiency, because in terms of standard cost-benefit analysis decision makers are likely to implement inferior projects.

Second, cost overruns of the size described above typically lead to delays, because securing additional funding to cover overruns often takes time. In addition, projects may need to be re-negotiated or re-approved when overruns are large, as the data show they often are (Flyvbjerg, 2005a). In a separate study, it was demonstrated that delays in transportation infrastructure implementation are very costly, increasing the percentage construction cost overrun measured in constant prices by 4.64 percentage points per year of delay incurred after the time of decision to build (Flyvbjerg, Holm, and Buhl, 2004). For a project of, say, US\$8 billion--that is the size range of the Channel tunnel and about half the size of Boston's Big Dig--the expected average cost of delay would be approximately \$370 million/year, or about \$1 million/day.--Benefit shortfalls are an additional consequence of delays, because delays result in later opening dates and thus extra months or years without revenues. Because many large infrastructure projects are loan-financed and have long construction periods, they are particularly sensitive to delays, as delays result in increased debt, increased interest payments, and longer payback periods.

Third, large cost overruns and benefit shortfalls tend to destabilize policy, planning, implementation, and operations. For example, after several overruns in the initial phase of the Sydney Opera House, the Parliament of New South Wales decided that every further 10 percent increase in the budget would need their approval. After this decision, the Opera House became a political football needing constant re-approval. Every overrun set off an increasingly menacing debate about the project, in Parliament and outside, with total cost overruns ending at 1,400 percent. The unrest drove the architect off the project, destroyed his career and oeuvre, and produced an Opera House unsuited for opera. Many other projects have experienced similar, if less spectacular, unrest, including the Channel Tunnel, Boston's Big Dig, and Copenhagen's metro.

Finally, as projects grow bigger, the problems with cost overruns and benefit shortfalls also grow bigger and more consequential (Flyvbjerg, Holm, and Buhl, 2004: 12). Some megaprojects are becoming so large in relation to national economies that cost overruns and benefit shortfalls from even a single project may destabilize the finances of a whole country or region. This occurred when the billion-dollar cost overrun on the 2004 Athens Olympics affected the credit rating of Greece and when benefit shortfalls hit Hong Kong's new \$20 billion Chek Lap Kok airport after it opened in 1998. The desire to avoid national fiscal distress has recently become an important driver in attempts at reforming the planning of large infrastructure projects, as we will see later.

### 3.6 Policy Implications

The policy implications of the results presented above are as follows:

- Lawmakers, investors, and the public cannot trust information about costs, benefits, and risks of large infrastructure projects produced by promoters and planners of such projects.
- The current way of planning large infrastructure projects is ineffective in conventional economic terms, i.e., it leads to Pareto-inefficient investments.
- There is a strong need for reform in policy and planning for large infrastructure projects.

Next, the paper examines the causes of cost overruns and benefit shortfalls.

### 3.7 Causes of Cost Overruns and Benefit Shortfalls

Three main types of explanation exist that claim to account for inaccuracy in forecasts of costs and benefits: technical, psychological, and political-economic explanations.

Technical explanations account for cost overruns and benefit shortfalls in terms of imperfect forecasting techniques, inadequate data, honest mistakes, inherent problems in predicting the future, lack of experience on the part of forecasters, etc. This is the most common type of explanation of inaccuracy in forecasts (Ascher, 1978; Flyvbjerg, Holm, and Buhl, 2002, 2005; Morris and Hough, 1987; Vanston and Vanston, 2004; Wachs, 1990). Technical error may be reduced or eliminated by developing better forecasting models, better data, and more experienced forecasters, according to this explanation.

Psychological explanations account for cost overruns and benefit shortfalls in terms of what psychologists call the planning fallacy and optimism bias. Such explanations have been developed by Kahneman and Tversky (1979), Kahneman and Lovallo (1993), and Lovallo and Kahneman (2003). In the grip of the planning fallacy, planners and project promoters make decisions based on delusional optimism rather than on a rational weighting of gains, losses, and probabilities. They overestimate benefits and underestimate costs. They involuntarily spin scenarios of success and overlook the potential for mistakes and miscalculations. As a result, planners and promoters pursue initiatives that are unlikely to come in on budget or on time, or to ever deliver the expected returns. Overoptimism can be traced to cognitive biases, that is, errors in the way the mind processes information. These biases are thought to be ubiquitous, but their effects can be tempered by simple reality checks, thus reducing the odds that people and organizations will rush blindly into unprofitable investments of money and time.

Political-economic explanations see planners and promoters as deliberately and strategically overestimating benefits and underestimating costs when forecasting the outcomes of projects. They do this in order to increase the likelihood that it is their projects, and not the competition's, that gain approval and funding. Political-economic explanations have been set forth by Flyvbjerg, Holm, and Buhl (2002, 2005) and Wachs (1989,1990). According to such explanations planners and promoters purposely spin scenarios of success and gloss over the potential for failure. Again, this results in the pursuit of ventures that are unlikely to come in on budget or on time, or to deliver the promised benefits. Strategic misrepresentation can be traced to political and organizational pressures, for instance competition for scarce funds or jockeying for position, and it is rational in this sense. If we now define a lie in the conventional fashion as making a statement intended to deceive others (Bok, 1979: 14; Cliffe et al., 2000: 3), we see that deliberate misrepresentation of costs and benefits is lying, and we arrive at one of the most basic explanations of lying that exists: Lying pays off, or at least political and economic agents believe it does. Where there is political pressure there is misrepresentation and lying, according to this explanation, but misrepresentation and lying can be moderated by measures of accountability.

### 3.8 Testing the Validity of Explanations

How well does each of the three explanations of forecasting inaccuracy--technical, psychological, and political-economic--account for the data on cost overruns and benefit shortfalls presented earlier? This is the question to be answered in this section.

Technical explanations have, as mentioned, gained widespread credence among forecasters and planners (Ascher, 1978; Flyvbjerg, Holm, and Buhl, 2002, 2005). It turns out, however, that such credence could mainly be upheld because until now samples have been too small to allow tests by statistical methods. The data presented above, which come from the first large-sample study in the field, lead us to reject technical explanations of forecasting inaccuracy. Such explanations do not fit the data well.

First, if misleading forecasts were truly caused by technical inadequacies, simple mistakes, and inherent problems with predicting the future, we would expect a less biased distribution of errors in forecasts around zero. In fact, it was found with high statistical significance that for four out of five distributions of forecasting errors, the distributions have a mean statistically different from zero. Only the data for inaccuracy in road traffic forecasts have a statistical distribution that seem to fit with explanations in terms of technical forecasting error.

Second, if imperfect techniques, inadequate data, and lack of experience were main explanations of inaccuracies, we would expect an improvement in accuracy over time, since in a professional setting errors and their sources would be recognized and addressed through the refinement of data collection, forecasting methods, etc. Substantial resources have in fact been spent over several decades on improving data and methods. Still our data show that this has had no effect on the accuracy of forecasts. Technical factors, therefore, do not appear to explain the data.

It is not so-called forecasting "errors" or their causes that need explaining. It is the fact that in a large majority of cases, costs are underestimated and benefits overestimated. One may agree with proponents of technical explanations that it is, for example, impossible to predict for the individual project exactly *which* geological, environmental, or safety problems will appear and

make costs soar. But one must also maintain that it is possible to predict the risk, based on experience from other projects, *that* some such problems will haunt a project and how this will affect costs. Such risk can and should be accounted for in forecasts of costs, but typically is not. For technical explanations to be valid, they would have to explain why forecasts are so consistent in ignoring cost and benefit risks over time, location, and project type.

Psychological explanations better fit the data. The existence of optimism bias in planners and promoters would result in actual costs being higher and actual benefits being lower than those forecasted. Consequently, the existence of optimism bias would be able to account, in whole or in part, for the peculiar bias found in most of our data. Interestingly, however, when you ask forecasters about causes for forecasting inaccuracies in actual forecasts, they do not mention optimism bias as a main cause of inaccuracy (Flyvbjerg, Holm, and Buhl, 2005: 138-140). This could of course be because optimism bias is unconscious and thus not reflected by forecasters. After all, there is a large body of experimental evidence for the existence of optimism bias (Buehler et al., 1994; Buehler, Griffin, and MacDonald, 1997; Newby-Clark et al. 2000). However, the experimental data are mainly from simple, non-professional settings. This is a problem for psychological explanations, because it remains an open question whether they are general and apply beyond such simple settings.

Optimism bias would be an important and credible explanation of underestimated costs and overestimated benefits in infrastructure forecasting if estimates were produced by inexperienced forecasters, i.e., persons who were estimating costs and benefits for the first or second time and who were thus unknowing about the realities of infrastructure building and were not drawing on the knowledge and skills of more experienced colleagues. Such situations may exist and may explain individual cases of inaccuracy. But given the fact that in modern society it is a defining characteristic of professional expertise that it is constantly tested--through scientific analysis, critical assessment, and peer review--in order to root out bias and error, it seems unlikely that a whole profession of forecasting experts would continue to make the same mistakes decade after decade instead of learning from their actions. Learning would result in the reduction, if not elimination, of optimism bias, which would then result in estimates becoming more accurate over time. But our data clearly shows that this has not happened.

The profession of forecasters would indeed have to be an optimistic--and non-professional--group to keep their optimism bias throughout the 70-year period our study covers for costs, and the 30-year period covered for patronage, and not learn that they were deceiving themselves and others by underestimating costs and overestimating benefits. This would account for the data, but is not a credible explanation. Therefore, on the basis of our data, we are led to reject optimism bias as a primary cause of cost underestimation and benefit overestimation.

Political-economic explanations and strategic misrepresentation account well for the systematic underestimation of costs and overestimation of benefits found in the data. A strategic estimate of costs would be low, resulting in cost overrun, whereas a strategic estimate of benefits would be high, resulting in benefit shortfalls. A key question for explanations in terms of strategic misrepresentation is whether estimates of costs and benefits are intentionally biased to serve the interests of promoters in getting projects started. This question raises the difficult issue of lying. Questions of lying are notoriously hard to answer, because a lie is making a statement intended to deceive others, and in order to establish

whether lying has taken place, one must therefore know the intentions of actors. For legal, economic, moral, and other reasons, if promoters and planners have intentionally cooked estimates of costs and benefits to get a project started, they are unlikely to formally tell researchers or others that this is the case. Despite such problems, two studies exist that succeeded in getting forecasters to talk about strategic misrepresentation (Flyvbjerg and Cowi, 2004; Wachs 1990).

Flyvbjerg and Cowi (2004) interviewed public officials, planners, and consultants who had been involved in the development of large UK transportation infrastructure projects. A planner with a local transportation authority is typical of how respondents explained the basic mechanism of cost underestimation:

“You will often as a planner know the real costs. You know that the budget is too low but it is difficult to pass such a message to the counsellors [politicians] and the private actors. They know that high costs reduce the chances of national funding.”

Experienced professionals like the interviewee know that outturn costs will be higher than estimated costs, but because of political pressure to secure funding for projects they hold back this knowledge, which is seen as detrimental to the objective of obtaining funding.

Similarly, an interviewee explained the basic mechanism of benefit overestimation:

“The system encourages people to focus on the benefits--because until now there has not been much focus on the quality of risk analysis and the robustness [of projects]. It is therefore important for project promoters to demonstrate all the benefits, also because the project promoters know that their project is up against other projects and competing for scarce resources.”

Such a focus on benefits and disregard of risks and robustness may consist, for instance, in the discounting of spatial assimilation problems described by Priemus (forthcoming) elsewhere in this issue. Competition between projects and authorities creates political and organizational pressures that in turn create an incentive structure that makes it rational for project promoters to emphasize benefits and deemphasize costs and risks. A project that looks highly beneficial on paper is more likely to get funded than one that does not.

Specialized private consultancy companies are typically engaged to help develop project proposals. In general, the interviewees found that consultants showed high professional standard and integrity. But interviewees also found that consultants appeared to focus on justifying projects rather than critically scrutinizing them. A project manager explained:

"Most decent consultants will write off obviously bad projects but there is a grey zone and I think many consultants in reality have an incentive to try to prolong the life of projects which means to get them through the business case. It is in line with their need to make a profit."

The consultants interviewed confirmed that appraisals often focused more on benefits than on costs. But they said this was at the request of clients and that for specific projects discussed "there was an incredible rush to see projects realized."



One typical interviewee saw project approval as "passing the test" and precisely summed up the rules of the game like this:

"It's all about passing the test [of project approval]. You are in, when you are in. It means that there is so much focus on showing the project at its best at this stage."

In sum, the UK study shows that strong interests and strong incentives exist at the project approval stage to present projects as favorably as possible, that is, with benefits emphasized and costs and risks deemphasized. Local authorities, local developers and land owners, local labor unions, local politicians, local officials, local MPs, and consultants all stand to benefit from a project that looks favorable on paper and they have little incentive to actively avoid bias in estimates of benefits, costs, and risks. National bodies, like certain parts of the Department for Transport and the Ministry of Finance who fund and oversee projects, may have an interest in more realistic appraisals, but so far they have had little success in achieving such realism, although the situation may be changing with the initiatives to curb bias set out in HM Treasury (2003) and Flyvbjerg and Cowi (2004).

The second study was carried out by Wachs (1990, 1986). Wachs interviewed public officials, consultants, and planners who had been involved in transit planning cases in the US. He found that a pattern of highly misleading forecasts of costs and patronage could not be explained by technical errors, honest mistakes, or inadequate methods. In case after case, planners, engineers, and economists told Wachs that they had had to "revise" their forecasts many times because they failed to satisfy their superiors. The forecasts had to be cooked in order to produce numbers that were dramatic enough to gain federal support for the projects whether or not they could be fully justified on technical grounds. Wachs (1990: 144) recounts from his interviews:

"One young planner, tearfully explained to me that an elected county supervisor had asked her to estimate the patronage of a possible extension of a light-rail (streetcar) line to the downtown Amtrak station. When she carefully estimated that the route might carry two to three thousand passengers per day, the supervisor directed her to redo her calculations in order to show that the route would carry twelve to fifteen thousand riders per day because he thought that number necessary to justify a federal grant for system construction. When she refused, he asked her superior to remove her from the project, and to get someone else to 'revise' her estimates."

In another typical case of cost underestimation and benefit overestimation, Wachs (1990: 144-145) gives the following account:

"a planner admitted to me that he had reluctantly but repeatedly adjusted the patronage figures upward, and the cost figures downward to satisfy a local elected official who wanted to compete successfully for a federal grant. Ironically, and to the chagrin of that planner, when the project was later built, and the patronage proved lower and the costs higher than the published estimates, the same local politician was asked by the press to explain the outcome. The official's response was to say, 'It's not my fault; I had to rely on the forecasts made by our staff, and they seem to have made a big mistake here'."

Like in the UK study above, Wachs specifically interviewed consultants. He found, as one consultant put it, that "success in the consulting business requires the forecaster to adjust results to conform with the wishes of the client," and clients typically wish to see costs underestimated and benefits overestimated (1990: 151-152).

On the basis of his pioneering study, Wachs (1990: 145) concludes that forecasts of costs and benefits are presented to the public as instruments for deciding whether or not a project is to be undertaken, but they are actually instruments for getting public funds committed to a favored project. Wachs (1990: 146, 1986: 28) talks of "nearly universal abuse" of forecasting in this context, and he finds no indication that it takes place only in transit planning; it is common in all sectors of the economy where forecasting routinely plays an important role in policy debates, according to Wachs.

### 3.9 Survival of the Unfittest

In conclusion, the UK and US studies arrive at results that are basically similar. Both studies account well for existing data on cost underestimation and benefit overestimation. Both studies falsify the notion that in situations with high political and organizational pressure the lowballing of costs and highballing of benefits is caused by non-intentional technical error or optimism bias. Both studies support the view that in such situations promoters and forecasters intentionally use the following formula in order to secure approval and funding for their projects:

$$\begin{aligned} & \text{Underestimated costs} \\ & + \text{Overestimated benefits} \\ & = \text{Project approval} \end{aligned}$$

Using this formula, and thus "showing the project at its best" as one interviewee said above, results in an inverted Darwinism, i.e., the "survival of the unfittest." It is not the best projects that get implemented, but the projects that look best on paper. And the projects that look best on paper are the projects with the largest cost underestimates and benefit overestimates, other things being equal. But these are the worst, or "unfittest," projects in the sense that they are the very projects that will encounter most problems during construction and operations in terms of the largest cost overruns, benefit shortfalls, and risks of non-viability. They have been designed like that.

A later paper will present ideas of how this situation may be improved (Flyvbjerg, 2007).

### 3.10 Summary

This paper identifies as the main problem in major infrastructure developments pervasive misinformation in the early phases of the project cycle about the costs, benefits, and risks involved. A consequence of misinformation is massive cost overruns, benefit shortfalls, and waste. The paper then explores the causes of misinformation and finds that optimism bias and strategic misinformation best account for the available evidence. In the case of strategic misrepresentation, planners and promoters deliberately misrepresent costs, benefits, and risks in order to increase the likelihood that it is their projects, and not the competition's, that gain

approval and funding. This results in a negated Darwinism with "survival of the unfittest," where often it is not the best projects that are built, but the most misrepresented ones.

### 3.11 References

- Altshuler, A. and D. Luberoff, 2003, *Mega-Projects: The Changing Politics of Urban Public Investment* (Washington, DC: Brookings Institution).
- Ascher, W., 1978, *Forecasting: An appraisal for policy-makers and planners* (Baltimore: The Johns Hopkins University Press)
- Bok, S., 1979, *Lying: moral choice in public and private life* (New York: Vintage)
- Buehler, R., Griffin, D., and MacDonald, H., 1997, The role of motivated reasoning in optimistic time predictions. *Personality and Social Psychology Bulletin*, 23, 3, pp 238-247
- Buehler, R., Griffin, D., and Ross, M., 1994, Exploring the "planning fallacy": Why people underestimate their task completion times. *Journal of Personality and Social Psychology*, 67, pp 366-381
- Cliffe, L., Ramsey, M., & Bartlett, D., 2000, *The politics of lying: Implications for democracy* (London: Macmillan)
- Flyvbjerg, Bent, 2005a, "Design by Deception: The Politics of Megaproject Approval." *Harvard Design Magazine*, no. 22, Spring/Summer, pp 50-59
- Flyvbjerg, Bent, 2005b, "Measuring Inaccuracy in Travel Demand Forecasting: Methodological Considerations Regarding Ramp Up and Sampling." *Transportation Research A*, vol. 39, no. 6, pp 522-530
- Flyvbjerg, Bent, Nils Bruzelius, and Werner Rothengatter, 2003, *Megaprojects and Risk: An Anatomy of Ambition* (Cambridge University Press)
- Flyvbjerg, Bent, 2004, "Megaprojects and Risk: A Conversation With Bent Flyvbjerg." Interview conducted by Renia Ehrenfeucht. *Critical Planning*, vol. 11, 2004, pp. 51-63.
- Flyvbjerg, 2007, other article in this volume.
- Flyvbjerg, Bent and Cowi, 2004, *Procedures for Dealing with Optimism Bias in Transport Planning: Guidance Document* (London: UK Department for Transport)
- Flyvbjerg, Bent, Mette K. Skamris Holm, and Søren L. Buhl, 2002, "Underestimating Costs in Public Works Projects: Error or Lie?" *Journal of the American Planning Association*, vol. 68, no. 3, Summer, pp 279-295
- Flyvbjerg, Bent, Mette K. Skamris Holm, and Søren L. Buhl, 2004, "What Causes Cost Overrun in Transport Infrastructure Projects?" *Transport Reviews*, vol. 24, no. 1, pp 3-18

- Flyvbjerg, Bent, Mette K. Skamris Holm, and Søren L. Buhl, 2005, "How (In)accurate Are Demand Forecasts in Public Works Projects? The Case of Transportation." *Journal of the American Planning Association*, vol. 71, no. 2, Spring, pp 131-146
- HM Treasury, 2003, *The Green Book: Appraisal and Evaluation in Central Government, Treasury Guidance* (London: TSO)
- Kahneman, D. and Lovallo, D., 1993, Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, 39, pp 17-31
- Kahneman, D. and Tversky, A., 1979, "Prospect theory: An analysis of decisions under risk." *Econometrica*, 47, pp 313-327
- Lovallo, Dan and Daniel Kahneman, 2003, "Delusions of Success: How Optimism Undermines Executives' Decisions," *Harvard Business Review*, July, pp 56-63
- Morris, Peter W. G. and George H. Hough, 1987, *The Anatomy of Major Projects: A Study of the Reality of Project Management* (New York: John Wiley and Sons)
- Newby-Clark, I. R., McGregor, I., and Zanna, M. P., 2002, Thinking and caring about cognitive inconsistency: when and for whom does attitudinal ambivalence feel uncomfortable? *Journal of Personality and Social Psychology*, 82, pp 157-166
- Priemus, Hugo, forthcoming, "Design of Large Infrastructure Projects: Disregarded Alternatives and Issues of Spatial Planning." *Environment and Planning B*
- Vanston, John H. and Lawrence K. Vanston, 2004, "Testing the Tea Leaves: Evaluating the alidity of Forecasts," *Research-Technology Management* 47(5), pp 33-39
- Wachs, M., 1986, Technique vs. advocacy in forecasting: A study of rail rapid transit. *Urban Resources*, 4(1), pp 23-30
- Wachs, M., 1989, When Planners Lie with Numbers. *Journal of the American Planning Association*, 55(4), pp 476-479
- Wachs, M., 1990, Ethics and advocacy in forecasting for public policy. *Business and Professional Ethics Journal*, 9(1-2), pp 141-157

## 4 HVLKE MULIGHETER HAR VI TIL Å FORUTSI I EN TIDLIG FASE?

KNUT SAMSET\*

### 4.1 Generelt

Ønsket om å kunne forutsi fremtidige hendelser er dypt forankret hos mennesker, og gjenspeiles på mange områder, ikke minst religion, filosofi, litteratur, psykologi, matematikk, økonomi og meteorologi. Det faglige spennet er enormt fra ikke-vitenskapelige spekulasjoner innen det som betegnes som astrologi, til presis matematisk modellering innen kvantefysikk og astronomi. Tidsspennet er også enormt, fra oldtidens orakel i byen Delphi, til nåtidens prosedyrer for gruppebasert systematisk gjetting, kalt Delphi- metoden.

Forutsigelser gjøres på mer eller mindre obskurt grunnlag, uttrykt ved forholdet mellom gjetning og fakta som inngår. I de fleste tilfeller vil det være et betydelig element av gjetning, også i en enkel trendfremskrivning, fordi vi vet at trender kan ta nye veier i tiden som kommer selv om historiske fakta gir en entydig indikasjon på hva som kommer til å skje.

Forskning innen kognitiv psykologi har vist at vi som individer har en tendens til å overvurdere vår egen evne til å forutsi. Samtidig nedvurderer vi andres evne til det samme og har lett for å avvise systematisk virksomhet på dette området, for eksempel fremtidsforskning, som useriøs. Vi ser da bort fra at det finnes meget vellykkete forsøk på å forutsi fremtidige hendelser, ikke bare på veldokumentert naturvitenskapelig grunnlag, men også på grunnlag av meget svak informasjon. En rekke Delphi-undersøkelser, som ofte har vært brukt til å vurdere fremtidig teknologisk utvikling er eksempel på dette. Et av de mer eksotiske eksemplene er hentet fra University of Washington der forskere studerte følelsesmessige responser mellom ektefeller. De videofilmte ektepar i dagligdags samtale i laboratoriet og analyserte deretter ansiktsuttrykket hos begge parter sekund for sekund. Disse ble gradert på en skala med et stort følelsesmessig register fra negativt til positivt, det vil si fra forakt til beundring. Ved å analysere dataene har det senere vist seg at forskerne bare på dette grunnlaget var i stand til å forutsi hvilke av parene som fortsatt ville være gift etter 15 år og hvilke som ville være skilt, med en nøyaktighet på 95 prosent, (Gladwell, 2005).

### 4.2 Forutsigelser i prosjekter

Vurderingene i den tidligste fasen av et prosjekt vil i stor grad baseres på antakelser. Evnen og muligheten for å forutsi vil være avgjørende for beslutningenes godhet. Muligheten for å forutsi er til dels avhengig av den typen informasjon som er tilgjengelig, dels lengden av tidsspennet som antakelsen gjelder, dels av dynamikken og kompleksiteten i de prosessene som inngår, dels av begrensningene i menneskers fantasi og vurderingsevne generelt, og dels av modellene og verktøyene som brukes for å bearbeide og analysere informasjon. I tillegg kommer at enkelte forhold ganske enkelt ikke kan forutsies. Dette indikerer at en står overfor

---

\* Knut Samset er professor ved NTNU, Institutt for bygg, anlegg og transport.

formidable utfordringer innledningsvis som blir større desto lenger tidligfasen varer. I de tilfellene der den strekker seg over tiår er kanskje utfordringen uoverkommelig.

Samtidig ser vi eksempler på at mange forhold er mer forutsigbare enn en forventer. En studie av 250 utviklingsprosjekter så nærmere på dette (Samset 1998). Studien tok utgangspunkt i de problemene som ble rapportert ex post i evalueringsrapporter å ha vært avgjørende for hvorvidt prosjektet lykkes eller ikke. En forsøkte deretter å spore tilbake hva som var årsaken til disse problemene, i hvilken grad årsakene kunne vært unngått og i hvilken grad de kunne vært forutsett. En fant at om lag to tredjedeler av problemene var forårsaket av forhold som prosjektet selv hadde hånd om, som for eksempel *prosjektledelse* og *prosjektutforming*. Bare en tredjedel av problemene hadde med utenforliggende forhold å gjøre, som for eksempel *markedets respons*, *miljømessige forhold* etc.

En fant også at i om lag 70 prosent av tilfellene kunne årsaken til problemene med stor sannsynlighet vært forutsett. Eksempler på slike forhold var ”uklar arbeidsdeling mellom institusjoner”, ”lav prioritering hos brukerne”, og ”dårlig vannkvalitet”. I de resterende 30 prosent ble årsaken til problemene vurdert som delvis forutsigbar og bare i noen tilfeller som ikke forutsigbar. Dette gjaldt forhold som for eksempel ”sene beslutninger”, ”dårlig vedlikehold” ”arbeidskonflikt” og ”sabotasje”.

Dersom denne undersøkelsen gjenspeiler forholdene mer generelt, så er det oppsiktsvekkende fordi folks oppfatning heller nærmest i motsatt retning: En antar at det først og fremst er utenforliggende årsaker som forårsaker de største problemene i prosjekter, og at disse er vanskelig forutsigbare.

### **Forutsigelse av enkle hendelser**

Ikke minst er det muligheten for å skaffe seg informasjon om tilsvarende og liknende prosjekter og erfaringene med disse som bedrer muligheten for å forutsi på et tidlig tidspunkt. Problemet i praksis ligger ikke så mye i at slik informasjon ikke er lett tilgjengelig. I de fleste tilfellene er det mulig å finne mange kilder til slik informasjon. Bøygen er heller ganske enkelt at en ikke har en tradisjon for å bruke ressurser på dette slik at det ofte ikke blir gjort i tilstrekkelig grad.

En måte å betrakte usikkerhet i beslutning på er ved å se på avviket mellom det faktiske utfallet og det som var forutsigelsen. Avviket mellom disse er et direkte uttrykk for forutsigelsens presisjon. Imidlertid kan dette bare verifiseres i etterkant. Vurderinger i forkant må baseres på skjønn og uttrykkes gjerne i sannsynlighet for at forutsigelsen er korrekt. Det er gjort mange eksperimenter innen kognitiv psykologi på dette området der respondentene skal foreta en gjetting men også selv vurdere sannsynligheten for at de gjetter riktig. Noen av konklusjonene fra slike eksperimenter (Wright and Ayton, 1987) peker i retning av at:

- folks evne til å gjette på tidsrelaterte hendelser er vanligvis bedre enn på kunnskapstemaer
- Folk med spesiell kunnskap og erfaring er bedre til å vurdere godheten av sine egne anslag
- Når treffprosenten er høy har folk en tilbøyelighet til å undervurdere vurderingsevnen og når den er lav overvurderer de godheten av sine anslag
- Grupper er bedre til å forutsi og vurdere godheten av forutsigelsene

- Evnen til å forutsi øker med trening – og særlig i den tidlige delen av treningsforløpet
- Evnen til å forutsi bedres vesentlig når komplekse hendelser brytes ned i flere enkle hendelser

Slike funn gir grunnlag for optimisme og kan ha overføringsverdi til det praktiske liv i å tilrettelegge for fornuftige beslutninger med best mulig utnyttelse av ressursene som inngår.

### **Forutsigelse av komplekse hendelser**

Det som er nevnt ovenfor gjelder forutsigelser av enkle hendelser i en eksperimentsituasjon. Dette avviker fra den situasjonen en står overfor i virkeligheten, dels ved at problemstillingene er mer komplekse og dels ved at vurderingsprosessen er generativ i den forstand at folk ofte bearbeider og endrer sitt skjønn i løpet av vurderingsprosessen. I virkelighetens verden er forutsigelse en dynamisk og sosial prosess som ikke bare bygger på tilgjengelig informasjon og erfaring, men også på interaksjon med andre mennesker. Situasjonen kompliseres ytterligere ved at hendelser sjelden kan betraktes som enkle årsak-virknings-forhold, men som dynamiske prosesser som påvirkes i ulik grad av forskjellige utenforliggende faktorer.

De siste 30-40 årene har det vært gjort mange forsøk på å anvende dynamiske simuleringmodeller for å beskrive komplekse systemer og forutsi hendelser. Dette har vært meget vellykket innenfor naturvitenskapen, mens det i samfunnsrelaterte problemstillinger har vist seg lite egnet. Mange tvilsomme analyser er blitt presentert som troverdige og har møtt voldsom faglig kritikk.<sup>8</sup> Det har derfor vært en økende erkjennelse om at matematisk simulering ikke er egnet i analyse av selvjusterende, samfunnsbaserte prosesser. Problemet bunner i at elementene i et sosialt system i motsetning til et fysisk system kan ta egne beslutninger. Et fysisk system er i prinsippet forutsigbart både på elementært nivå og på et aggregert nivå. Et sosialt system kan til en viss grad være forutsigbart på elementært nivå men i liten grad på aggregert nivå.

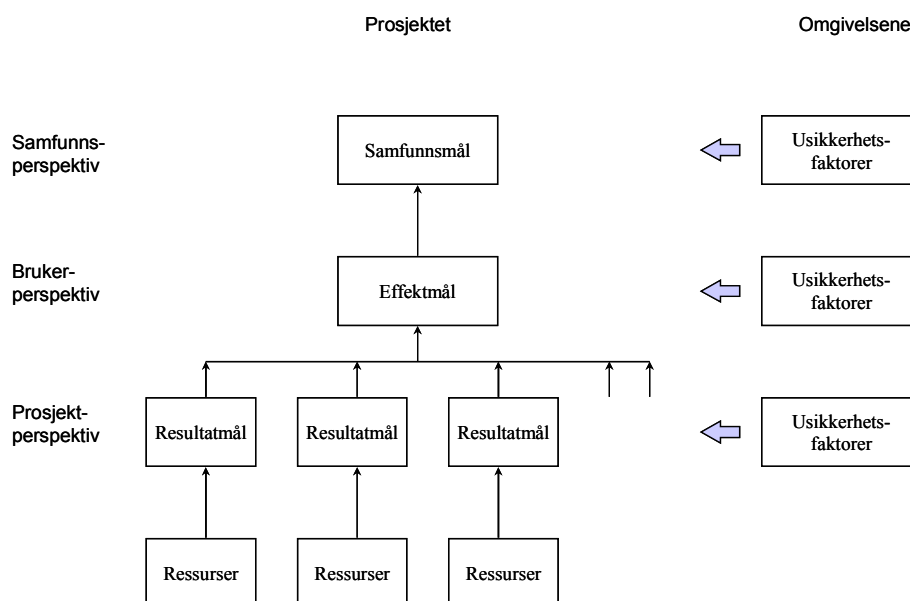
Samfunnsbaserte systemer – og et prosjekt er et slikt system – kan beskrives som et komplekst samspill av hendelser. Over tid kan en observere mønstre i det som skjer. Tiltak kan iverksettes på forskjellig tidspunkt som viser seg å påvirke hendelser i ønsket eller uønsket retning. Hendelser kan være tilsiktet og derfor mulige å forutsi - eller utilsiktet og i så fall vanskelige å forutsi. Det er vanlig å skille mellom *trender* og *overskridelser*. Trender kan identifiseres retrospektivt på empirisk grunnlag og fremskrives for å forsøke å forutsi fremtidig utvikling, for eksempel trafikkprognoser eller kriminalitet. Prognosen blir mer usikker desto lengre en trekker den inn i fremtiden. Når det gjelder overskridelser, forårsakes disse av en type hendelser som fører til kvalitative sprang eller vesentlige diskontinuerlige endringer i eksisterende trender. Det kan for eksempel være en naturkatastrofe, eller en ny og banebrytende teknologi, eller et regimeskifte – hendelser som får uventete konsekvenser og snur handlingsmønstre og trender i samfunnet. Slike overskridende hendelser er i prinsippet uforutsigbare og kaster i tillegg om på det forutsagte.

---

<sup>8</sup> En av de første og mest kjente var "Limits to Growth" utført ved Massachusetts Institute of Technology på begynnelsen av 1970-tallet, der en simulerte global ressursbruk og konsekvensene av dette i et 100-års perspektiv.

### 4.3 Behovet for å forutsi i prosjekters tidligste fase

I en situasjon der realiseringen av prosjektet ligger ett eller flere tiår frem i tid er muligheten for å forutsi med noen grad av presisjon svært begrenset. Det er nok heller ikke nødvendig. På dette tidspunktet handler det om å forstå problemet og kunne beskrive realistiske strategier for å løse dette – og identifisere vesentlige faktorer som kan tenkes å påvirke realiseringen av disse strategiene. Problemet fremstår som nåtidige men vil selvsagt ofte være projisert frem i tid i forsterkende retning. Det kan være behov for å forutsi omfang og endringer i problemene over tid. Dette kan gjøres kvalitativt eller kvantifiseres med lavt presisjonsnivå.



Figur 4-1 Prosjektstrategi fremstilt som et hierarki av hendelser

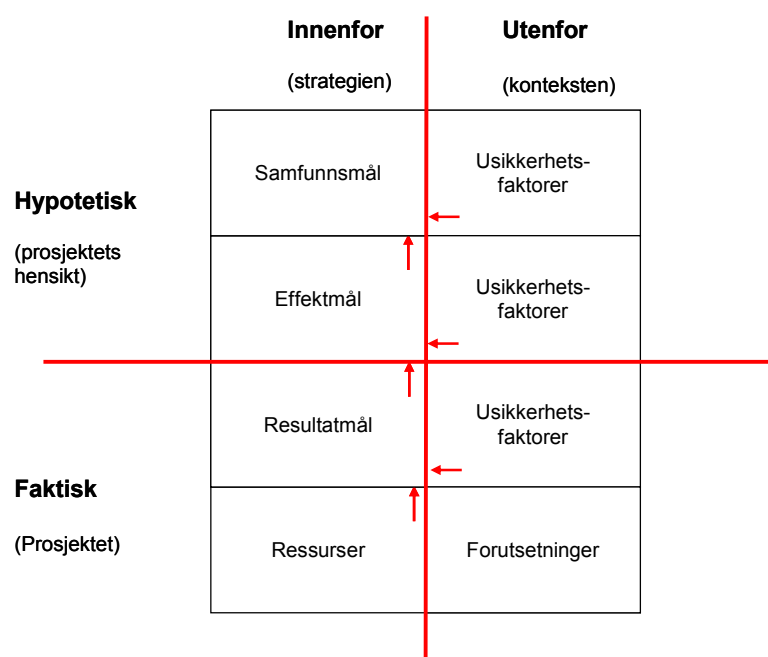
Prosjektstrategien beskrives i prinsippet slik det er gjort i figur 4.1, i form av en årsak-virkningskjede av hendelser som skal realiseres - formulert som *mål*. Elementene i denne årsak-virkningskjeden er størrelser som i prinsippet bestemmes på et normativt grunnlag og i mindre grad basert på forutsigelser. Målene skaleres gjerne til omfanget av det problemet som skal løses slik det er forutsett. Samtidig må det skaleres i forhold til ressursinnsatsen og kapasiteten til det som prosjektet skal levere. Når det gjelder leveransen, som her er betegnet med resultatmål, så er denne i prinsippet forutsigbar, en antar at det er 100 prosent sannsynlig at prosjektet leverer det som er avtalt.

De eksterne forhold som kan tenkes å påvirke strategien fremstår i form av hendelser og er her betegnet som *usikkerhetsfaktorer*. Disse må forutsees. På et tidlig tidspunkt vil det normalt være tilstrekkelig å uttrykke usikkerhetsfaktorene i kvalitative termer, men det kan være nyttig å forsøke å forutsi når disse kan tenkes å inntreffe, og også hvor stor effekt de kan få. Dette innebærer en risikovurdering på et forholdsvis lavt presisjonsnivå, ettersom risiko defineres som sannsynlighet multiplisert med konsekvens.



## 4.4 Et rammeverk for tidlig vurdering av prosjektstrategi

Det som er nevnt ovenfor er sammenfattet i modellen beskrevet i figur 4.2 som har fått stor utbredelse innenfor FN, EU og OECD som en modell og metode for konseptvurdering. Modellen er kjent som "Log-frame-modellen", det logiske rammeverket, og brukes som en første tilnærming til å beskrive og risikovurdere en strategi. Anvendelsen omtales derfor også som strategianalyse (Samset 1999). Metoden er metodisk *konsistent* i den forstand at matrisen beskriver et globalt utfallsrom som kan romme alle tenkelige hendelser innenfor de dimensjonene som er gitt (en hendelse er enten faktisk realiserbar - eller hypotetisk, dvs. at sannsynligheten for realisering er mindre enn 100 prosent. Hendelsen hører enten til innenfor strategien, dvs. prosjektets ansvarsområde - eller utenfor). Samtidig er utfallsrommet delt i gjensidig utelukkende kategorier (det som er innenfor strategien - versus det som er utenfor, og det som er faktisk - versus det som er hypotetisk). Ved å dele inn denne firefeltsmatrisen som vist i figuren slik at en får fire nivåer vertikalt, fremstår prosjektstrategien på venstre side i matrisen som en årsak-virkningskjede, og usikkerhetsfaktorene knyttet til hvert nivå i strategien på høyre side.



Figur 4-2 Rammeverk for tidlig, kvalitativ vurdering av prosjektkonsepter

Den utfordringen en står overfor i den tidligste fasen er som nevnt å beskrive helheten, det vil si at strategien skal være konsistent og realistisk, og at usikkerhetsfaktorene skal være slik at de ikke truer realiseringen av strategien. Mer konkret innebærer dette følgende krav og tilhørende utfordringer når det gjelder informasjon og vurdering (i kursiv):

1. Strategien skal utgjøre en konsistent årsak-virkningskjede
  - *dette krever vurdering av logikken*
2. Samfunnsmål og effektmål skal være realiserbare innenfor de tidsrammene en har satt
  - *dette krever vurdering av sannsynlighet for realisering*
3. Resultatmålet skal være realiserbart i forhold til ressursinnsatsen

- *dette krever en økonomisk vurdering*
4. De vesentligste usikkerhetsfaktorene skal identifiseres
    - *dette krever erfaring og evne til å forutsi*
  5. Usikkerhetsfaktorene skal tilordnes riktig nivå i strategien
    - *dette krever vurdering av logikken*
  6. Usikkerhetsfaktorene skal vurderes med hensyn til sannsynlighet og konsekvens og kritiske elementer identifiseres
    - *det krever sannsynlighetsvurdering*
  7. Samlet effekt av hendelsesusikkerheten skal vurderes
    - *det krever skjønnsmessig vurdering av sannsynlighet*

Den modellen som er beskrevet ovenfor har stor praktisk utbredelse som en første tilnærming til utviklingen og vurderingen av prosjektkonsepter. En kan si at den innebærer en første grov risikovurdering av en prosjekt-ide. Det fremgår av det som er nevnt ovenfor - og dette bekreftes av den erfaringen en har med metoden, at det ikke bare er meningsfullt, men også spesielt effektivt å basere en slik analyse stort sett på kvalitativ informasjon. Med det menes ikke det samme som ukvalifiserte antakelser. En har selvsagt et stort behov for relevant informasjon for å gjennomføre slike analyser. Grundig beskrivelse og vurdering av problemer og behov, samt interesser og prioriteringer hos berørte parter er viktig. Det samme gjelder forhold som kan tenkes å påvirke gjennomføringen av en strategi. I den grad informasjonen er kvantifisert er det naturlig å sammenfatte den i kvalitative termer i analysen for å kunne konsentrere denne om det prinsipielle og ikke om størrelser.

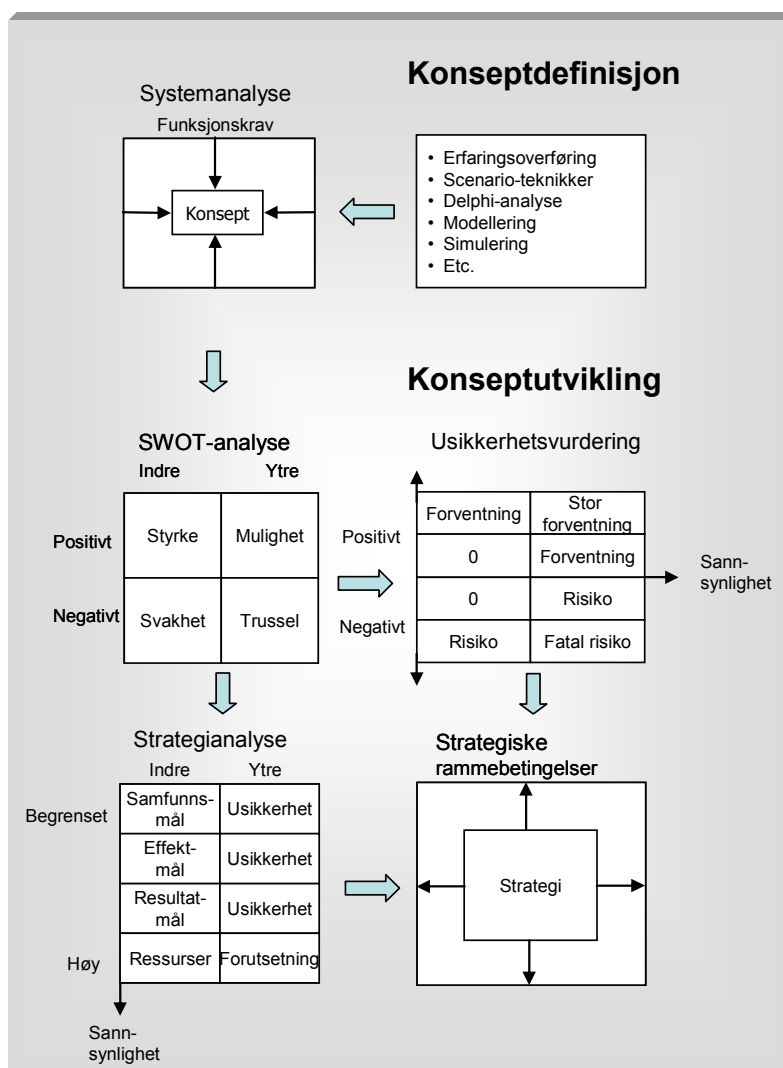
En vet at individuell erfaring alene ikke er det beste grunnlag for å gjøre gode forutsigelser. Videre, at skjønnsmessige vurderinger ofte påvirkes av individers holdninger og preferanser, og at slike skjevheter ofte er systematiske og robuste. En vet også at forutsigelser basert på gruppebasert skjønn generelt faller bedre ut enn individers skjønn, også bedre enn eksperters skjønn (Surowiecki, 2004). Ennå viktigere er det kanskje at systematisk bruk av relevant erfaringsmateriale og modeller ofte gir bedre forutsigelser enn eksperters og gruppers skjønnsmessige vurderinger. Dette til tross, er det overraskende hvor lite ressurser som i dag brukes til å trekke lærdom fra prosjekter som har vært gjennomført, systematisere dette, gjøre det tilgjengelig og bruke lærdommen i forbindelse med nye prosjekter. På dette området ligger det et stort og til dels utforsket utviklingspotensial.

Dette betyr at det normalt er behov for betydelige forundersøkelser supplert med kunnskap fra tilsvarende prosjekter. Om en i tillegg gjennomfører analysen i et samarbeid med godt informerte personer i en konsensusbasert prosess, kan en oppnå meget gode resultater med slike enkle hjelpemidler. Uansett vil det i mange tilfeller være vesentlig bedre enn hva som ofte er praksis i dag – nemlig å ikke gjøre en grunnleggende systematisk innledende vurdering.

## 4.5 Enkelte verktøy og teknikker

Denne studien fokuserer utelukkende på konseptfasen, som i figur 4.3 er delt inn i to delfaser: Konseptdefinisjon og konseptutvikling. Skissen viser hvordan vanlige teknikker er knyttet til hverandre og representerer et helhetlig og systematisk sett av analytiske steg der valg av parametere og bearbeiding av informasjon utgjør en helhet. Elementene som er illustrert i figuren er summarisk beskrevet nedenfor.

Tilnærmingen som beskrives her er selvsagt bare en av mange mulige måter å foreta systematisk problemløsning på. Imidlertid bygger den på kjente og mye anvendte metoder – og viser hvordan disse bygger på hverandre - og bringer inn nye parametere etter hvert - som gir grunnlag for videre bearbeiding. Et viktig kjennetegn ved de tre sentrale elementene i figuren, det vil si SWOT-analysen, strategianalysen og usikkerhetsanalysen, er at de er representert som fullstendige, globale utfallsrom som er delt opp i gjensidig utelukkende kategorier, slik at de gir konsistente rammer for å analysere konsepter.



Figur 4-3 Systematisk informasjonsbearbeiding på veien fra problem til konsept

## Konseptdefinisjon

Konseptet skal i prinsippet være det mest hensiktsmessige svaret på ett eller flere mer eller mindre konkrete uttalte behov. En helhetlig tilnærming tilsier at en først må klargjøre hva disse behovene er og de funksjonskravene som behovene medfører i den konteksten<sup>9</sup> som konseptet vil være en del av. Kravet til den prosessen som fører fram til identifisering av et prosjekt konsept er først og fremst at den skal være åpen. Det betyr at den ikke er bundet av på forhånd bestemte valg, men legger til rette for å identifisere ulike alternativer på en fordomsfri måte.

Systemanalyse representerer en systematisk tilnærming til problemløsning ved at en beskriver problemstillingen som et system samt systemets ytre betingelser for å lykkes, og deretter identifiserer konkrete løsninger og tester dem ut mot systemets ytre betingelser. Dette gir grunnlag for å foreta et valg. Systemanalysen er altså en utenfra-og-inn tilnærming som antydnet i figur 4.3. En bryter derved med tradisjonell problemløsning som tar utgangspunkt i et grunnleggende valg og deretter drøfter effekten av dette.

Metodegrunlaget som benyttes kan være alt fra enkle konseptuelle modeller til systemdynamisk simulering. Som nevnt ovenfor er det imidlertid helt grunnleggende at en trekker inn relevant erfaring fra liknende situasjoner eller prosjekter. Tid og systemets livsløp vil være en sentral parameter i systemanalysen. En generell prosedyre kan være som følger:

1. **Systemdefinisjon:** Fokuserer på et problemkompleks og avgrensning av systemet
2. **Behovsdefinisjon:** Identifisere sentrale parter som inngår eller påvirkes, deres behov og prioriteringer, og eventuelle endringer over tid i hele livsløpet.
3. **Kravspesifikasjon:** Ut fra de behov som foreligger identifisere funksjonelle krav (hvordan systemet skal virke), fysiske krav (hvordan det skal bygges og hva det skal tåle), operasjonelle krav (hvordan systemet skal drives), og økonomiske krav (hva systemet skal koste i utvikling og drift)
4. **Alternative delsystemer:** Identifisere teknologier eller konsepter som kan bidra til å tilfredsstille systemets krav
5. **Ytelsesvurdering:** Teste delsystemer mot kravspesifikasjonen
6. **Konseptvalg:** Evaluere alternativer for å komme frem til det som er mest egnet
7. **Sårbarhetstesting:** Simulere systemet under forskjellige scenarier med endringer, feil og forstyrrelser, for eksempel ved hjelp av modeller
8. **Beskrive systemet**

For å gjennomføre en slik omfattende prosedyre vil en ha behov for konkrete verktøy og teknikker som antydnet i figuren. Det kan være databaser og nettverk for å sikre tilgang til relevant erfaring fra liknende prosjekter, Delphi-teknikk eller Monte Carlo simulering for å forutsi eller estimere, scenario-verktøy og mind-mapping teknikker for å beskrive/visualisere, flermålsanalyse for å rangere alternativer, og systemdynamiske modeller for å vurdere effekt og foreta sensitivitetstesting.

---

<sup>9</sup> sosialt, økonomisk, institusjonelt, teknisk, miljømessig og politisk

## Konseptutvikling

Gitt at en har funnet fram til ett eller flere lovende konsepter vil det være behov for å konkretisere og utvikle disse fra idé til det som vil være det strategiske valget. For å kunne gjøre vil det være behov for ytterligere informasjon og konkretisering av hva en ønsker å oppnå, hva prosjektet skal produsere, hvilke ressurser som står til rådighet, og hvilke muligheter og begrensninger som foreligger for å gjennomføre prosjektet.

Et første steg kan være å foreta en *SWOT-analyse*.<sup>10</sup> Denne teknikken trekker opp et utfallsrom delt i fire kategorier for å beskrive kjennetegn ved det påtenkte prosjektet. Den skiller altså mellom det interne og det eksterne på den ene siden og det positive og det negative på den andre siden.

SWOT-metoden brukes for å hente fram informasjon som grunnlag for strategidannelse. Strategien blir til i grensesnittet mellom en kartlegging og vurdering av de muligheter og trusler som prosjektet står overfor i den konteksten det opererer innenfor – der en også vurderer prosjektets egen styrke og svakheter. Strategiens mål er å utnytte de mulighetene som finnes på bakgrunn av prosjektets styrke, samtidig som en unngår ytre trusler og tar høyde for indre svakheter. Målet er altså ambisiøst. Metoden har imidlertid et begrenset siktemål. Den brukes først og fremst prospektivt som et hjelpemiddel til å kartlegge viktige forhold som er relevant for strategisk planlegging i tidligfasen. SWOT analysen forutsetter imidlertid at en går nærmere inn på prosjektet for å tenke det inn i sin kontekst institusjonelt, teknisk, miljømessig, samfunnsmessig, etc.

Svakheter ved metoden er at en slik enkel beskrivelse av virkeligheten i kategorier ikke bidrar til å illustrere samspillet og dynamikken i de prosessene en studerer, og derfor bare kan brukes i en første tilnærming til en analyse. Resultatet avhenger av innsikten og forståelsen hos de som medvirker i analysen, og ikke minst sammensetningen av disse. Dersom ikke de viktigste interessegruppene er representert vil resultatet ha begrenset verdi og gyldighet. Dersom resultatet bygger på en konsensusbasert dialog mellom informerte parter kan resultatet av SWOT-analyser være meget verdifullt.

Et neste logisk steg vil være å foreta en *strategianalyse*, som beskrevet i figur 4.2. Den viderefører SWOT-analysen og beskriver igjen et metodisk konsistent utfallsrom der en skiller mellom indre og ytre forhold langs den ene aksene, og innfører sannsynlighet for realisering som en ny parameter langs den andre aksene. Et sentralt poeng her er skillet mellom hva som er *faktisk* (realiserbart), og hva som er *hypotetisk* (ønskelig). Når utfallsrommet deles inn i åtte deler som vist i figur 4.3 kalles det et logisk rammeverk. Dette beskriver prosjektstrategien på venstre side, og usikkerhetsfaktorene som påvirker strategien på høyre side. Rammeverket brukes for å drøfte og analysere sannsynlighet for realisering med sikte på å få fram et mest mulig realistisk gjennomførbart konsept på et tidlig tidspunkt.

Metoden er en enkel målstyringsteknikk som først og fremst egner seg for kvalitative vurderinger. Den anvendes ofte i arbeidsgrupper der nøkkelpersoner og ressurspersoner analyserer alternative prosjektstrategier på grunnlag av egen erfaring, opplysninger fra forundersøkelser osv. I slike sammenhenger vil mange av innspillene være basert på

---

<sup>10</sup> SWOT står for Strength, Weaknesses, Opportunities and Threats

subjektive vurderinger og antagelser. Et premiss i slike arbeidsgrupper er normalt at alle konklusjoner skal være basert på konsensus.

Teknikken er også godt egnet til å foreta en enkel vurdering av systematikken og realismen i eksisterende konsepter eller prosjektutforminger. Når en plotter inn prosjektets mål i matrisens venstre side og usikkerhetselementene på høyre side vil en ofte finne at prosjekter har uklare, ambisiøse, komplekse, eller sammensatte mål som ikke gir en klar retning for hvor en vil hen, og at det kanskje er interne målkonflikter. En finner av og til at det som er angitt som resultatmål i virkeligheten er effektmål, og omvendt – og at det som er angitt som mål ofte er usikkerhetsfaktorer utenfor prosjektets mandat, og omvendt. En slik gjennomgang av et eksisterende prosjekt vil gjerne gi en avklaring av konsistens og realisme og føre til vesentlige forbedringer i prosjektstrategien.

I den innledende fasen av konseptutviklingen er det vanlig å begrense ambisjonsnivået til å foreta rent kvalitative vurderinger og ikke tallfeste omfang, kostnad eller tid. Etter hvert som vurderingen av konseptet skrider frem, vil en i økende grad kvantifisere enkeltkomponentene i strategien og de eksterne usikkerhetsfaktorene, for å øke presisjonen i vurderingene og tilpasse konseptet til de funksjonskravene og den ressurstilgangen som er aktuell. Teknikken bidrar ikke minst til å identifisere omgivelsesbasert usikkerhet, og vurdere de muligheter og risiki som usikkerheten innebærer.

Denne metoden er primitiv i den forstand at den bare benytter enkle, ofte intuitive sannsynlighetsvurderinger. Erfaringen viser imidlertid at når informasjonsunderlaget er godt og metoden brukes av fornuftig sammensatte arbeidsgrupper, kan en oppnå gode resultater på en enkel og effektiv måte. Ikke minst bidrar den til kommunikasjon mellom dem som deltar i prosessen og fremmer en felles forståelse av prosjektet som kommer til nytte under den påfølgende detaljutformingen og gjennomføringen.

Gitt at en har definert en strategi vil det neste opplagte steget være å foreta en *vurdering av usikkerhet* knyttet til realisering av strategien. Figur 4.3 skisserer et rammeverk for grovanalyse av usikkerhetsfaktorer. Her viderefører en skillet mellom det som antas positivt og negativt fra SWOT-analysen for å skille mellom muligheter og risiki. Igjen innføres sannsynlighet som en ny parameter. Ved å dele inn utfallsrommet i underkategorier vil en identifisere lavrisiko- og lavmulighets- hendelser der både konsekvens og sannsynlighet er lav. Dette er hendelser en kan se bort fra. På den negative siden kan en identifisere hendelser som har stor konsekvens og stor sannsynlighet for å inntreffe og som derfor krever at en redefinerer konseptet, såkalt fatal risiko. Grovanalysen gir også oversikt over de gjenstående usikkerhetsfaktorene som det må tas hensyn til i den videre vurderingen av strategien.

Resultatet av konseptutviklingen vil i første instans være at det etableres et *strategisk rammeverk* for utforming og gjennomføring av prosjektet. Det logiske rammeverket i figur 4.2 beskriver prosjektkonseptet i et strategisk perspektiv. Usikkerheten knyttet til dette gjør at det vil være vanskelig å realisere prosjektet som forutsatt, og at en må tillate fleksibilitet i gjennomføringen. For å unngå at dette fører til for store avvik i den strategiske retningen etableres det strategiske rammebetingelser for den som gjennomfører prosjektet.

To hovedforutsetninger for at prosjekter skal lykkes er at en i tidligfasen kommer frem til en fornuftig *strategi* og at en i gjennomføringsfasen har den nødvendige fleksibilitet til å foreta nødvendige *taktiske* valg etter hvert. Som nevnt i kapittel 1 er skillet mellom det strategiske grepet og den taktiske fleksibiliteten viktig i prosjektstyring. I tilfeller der et prosjekt både

finansieres og gjennomføres av en og samme part unngår en at det oppstår prinsipielle interessekonflikter i grensedragningen mellom det strategiske og det taktiske. Ofte er dette ikke tilfellet, og det kan oppstå betydelige interessekonflikter mellom den finansierende partens ønske om strategisk retning og den gjennomførende partens behov for taktisk manøvreringsmulighet.

På grunnlag av de analysene som er beskrevet ovenfor vil en kunne trekke opp strategiske rammebetingelser for det tenkte prosjektet, det vil si at en spesifiserer de viktigste styringsparameterne sett fra prosjekteierens eller den som finansierer prosjektets side. Strategien gir de viktigste styringsparameterne for den som gjennomfører prosjektet. Skissen på figur 4.3 antyder at rammebetingelsene skal defineres så vidt at de gir den nødvendige taktiske fleksibilitet for den som gjennomfører prosjektet, dvs. at de gir rom for å håndtere usikkerhet og uforutsette hendelser i gjennomføringen. Strategien og rammebetingelsene vil utgjøre en fullstendig beskrivelse av det tenkte prosjektet på konseptstadiet.

## 4.6 Referanser

Bazerman, Max, 1994, *Judgment in Managerial Decision Making*, John Wiley & Sons, Inc., New York

Bennett, Deborah, 2004, *Logic made easy*, Penguin Books, London

Christensen Søren, Kreiner Kristian, *Prosjektledelse under usikkerhet*, Universitetsforlaget, Oslo 1991

Cooke Roger M., 1991, *Experts in Uncertainty. Opinion and Subjective Probability In Science*, Oxford University Press

de Jouvenel Bertrand, 1967, *The art of conjecture*

Gigerenzer Gerd, Todd Peter M., 1999, *Simple Heuristics That Make Us Smart*, Oxford University Press, Inc.

Gladwell Malcolm, 2005, *Blink. The Power of Thinking Without Thinking*, Penguin Books Ltd

Goodwin, Paul and Wright, George, 1996, *Decision Analysis for Management Judgment*, John Wiley & Sons, Inc., London

Gullvåg, Ingemund, 1990, *Rasjonalitet, forståelse og forklaring. Innføring i argumentasjonsteori, logikk og vitenskapsfilosofi*, Tapir Akademisk Forlag, Trondheim

Lewis H. W., 1997, *Why flip a coin? The art and science of good decisions*, John Wiley & Sons, Inc.

Samset, Knut, 1999, *The Logical Framework Approach (LFA), Handbook for objectives-oriented planning*, 4th edition, Norad, Oslo

Samset, Knut, 2001, *Prosjektvurdering i tidligfase, Fokus på konseptet*, Tapir Akademisk Forlag, Trondheim

Samset, Knut, 2006: Design of high-uncertainty projects in international aid, paper presentert på konferansen "*PROMAC 2006*", Sydney

Surowiecki James, 2004, *The Wisdom of Crowds. Why the Many Are Smarter Than the Few*, Little, Brown/ Doubleday

Teigen, Karl Halvor, 2006: *Skjønn og skivebom. Hvordan vi bedømmer usikkerhet*, Foredrag, Norsk Senter for Prosjektledelse, Oslo

Wright George and Ayton Peter (ed.), 1987, *Judgmental forecasting*, John Wiley & Sons Ltd.



## Del 2

# Generering av informasjon i en tidlig fase



## 5 EKSPERTVURDERINGER FOR Å AVDEKKE OG STRUKTURERE KVALITATIV INFORMASJON

KJELL J. SUNNEVÅG\*

### 5.1 Innledning<sup>11</sup>

En ekspertvurdering er rett og slett en vurdering gjort av en ekspert. Når det er knapt med tilgjengelige eller sammenlignbare kvantitative (statistiske) erfaringsdata, kan den best tilgjengelige informasjon være erfaringsbaserte og subjektive oppfatninger hos eksperter innenfor det aktuelle usikkerhetsområdet. Ekspertene kan også benyttes til å strukturere et problem, dvs. bestemme hvilke data og variable som er relevante for analysen, hvilke analytiske metoder som er hensiktsmessige, samt hva som er relevante forutsetninger.

Ekspertvurderinger er imidlertid ingen entydig definert tilnærming; det kan være alt fra ”gut feeling” til mer sofistikerte tilnærminger der ekspertens vurdering er understøttet av historiske data, en strukturert prosess for å avdekke ekspertinformasjon og sjekklister (”strukturert estimering”, jf. Jørgensen (2002)). Det er imidlertid utviklet en rekke ulike teknikker for å avdekke oppfatninger, estimater, anbefalinger eller beslutninger fra en enkelt eller grupper av eksperter, så som fremlokking av subjektive sannsynligheter, aggregering av oppfatninger, og konsensusbyggende teknikker så som Delfi-metoden.

Det ligger videre i sakens natur at ekspertinformasjon er subjektiv. Den er basert på oppfatninger og erfaringer hos en eller flere enkeltpersoner. For å avhjelpe dette kan hjelp hos flere eksperter søkes. Det kan også være nyttig å forstå den enkelte eksperts erfaringsbakgrunn slik at ekspertens oppfatning eventuelt kan vektas på denne bakgrunn. Vi skal se nærmere på årsaker til skjevheter i EV og hvordan de kan korrigeres.

### 5.2 Når benyttes eksperttilnærminger?

Innledningsvis ble det nevnt at ekspertbaserte tilnærminger først og fremst benyttes når det er knapt med tilgjengelige eller sammenlignbare kvantitative (statistiske) erfaringsdata. Likevel er det viktig å gjøre oppmerksom på at ekspertbasert informasjon ikke nødvendigvis fører til det mest korrekte estimatet, gitt tilgjengelig informasjon. Jørgensen (2002) gjennomgår femten ulike studier som sammenligner ekspertestimerer med estimater basert på mer formelle estimeringsmodeller, og finner ikke empirisk grunnlag som går i favør av verken ekspert- og modellbaserte tilnærminger. Et sentralt poeng er imidlertid at det er omstendigheter der den

---

\* Kjell J. Sunnevåg er forskningssjef Senter for Økonomisk Politikk ved Samfunns- og næringslivsforskning AS (SNF).

<sup>11</sup> Ole Jonny Klakegg har bidratt med mange nyttige kommentarer og innspill til dette kapitlet.

ene metoden er bedre enn den andre. Det blir dermed viktig å være klar over når den ene fremgangsmåten bør benyttes fremfor den andre.

Selv om hvilke omstendigheter som favoriserer ekspertbaserte tilnæringer ikke er et eksplisitt tema i de empiriske studiene som ble gjennomgått, prøver Jørgensen (*op. cit.*) å oppsummere noen hovedmomenter med et rimelig sterkt empirisk fundament. For det første vil ekspertestimerer gjerne være mer presise enn modellestimerer dersom eksperten besitter (og anvender) viktig kunnskap som ikke er inkludert i estimeringsmodellen. Det er videre slik at eksperter typisk benytter enkle estimeringsstrategier (heuristikk), og eksperter gjør det bedre enn modellbaserte tilnæringer når disse estimeringsstrategier er gyldige. I motsatt fall kan ekspertbaserte estimater føre til estimatskjevhet. Det er også slik at eksperter kan villedes av irrelevant informasjon, for eksempel være for optimistiske, eller påvirkes av nylige hendelser. Estimeringsmodeller er mindre utsatt for slike estimatskjevheter.

Når det gjelder spørsmålet om modellestimerer er bedre enn ekspertvurderinger når usikkerheten er relativt stor, er spørsmålet mer uavklart; her er det resultater som peker i ulike retninger.

### 5.3 Ulike tilnæringer for å avdekke ekspertinformasjon

Det finnes en rekke ulike teknikker for å fremskaffe ekspertinformasjon. Likevel kan disse inndeles i tre grunnleggende hovedkategorier:

- **Individuelle intervjuer**, der eksperten intervjues, vanligvis på tomannshånd. Dette er omgivelser der intervjueren gjennom et strukturert intervju kan få informasjon fra eksperten uten påvirkning fra andre eksperter. Situasjonen tillater også at intervjueren kan gå i dybden på enkelte temaer eller problemstillinger om nødvendig.
- **Interaktive grupper**, der en gruppe eksperter og en gruppeleder møtes ansikt til ansikt for å presentere, og diskutere data. Hvordan ekspertene interagerer med hverandre kan struktureres på ulike måter, fra den mer ustrukturerte fremgangsmåten til en tilnærming som er mer koreografert med hensyn til når den enkelte ekspert presenterer sine syn, og når det er åpen diskusjon. Hvilken praksis som velges kan være viktig i forhold til å avhjelpe ulike problematiske sider ved at eksperter interagerer i grupper, for eksempel påvirkning fra dominerende eksperter og at ulike eksperter i ulik grad er komfortable med å fremstå med synspunkter i slike fora.
- **Delfi-metoden** ble opprinnelig utviklet av Rand Corporation for å motvirke noen av de estimatskjevheter som kan oppstå når eksperter interagerer i grupper. Ved denne metoden gir eksperter, isolert fra hverandre, sine synspunkter til en undersøkelsesleder. Undersøkelseslederen anonymiserer vurderingene og omfordeler disse til ekspertene, og tillater dem å revidere sitt opprinnelige estimat eller sin opprinnelige vurdering. Disse iterasjonene kan i prinsippet gjentas inntil konsensus oppnås, om dette er en målsetting.

Det er mange fallgruver ved de ulike tilnærmingene. Hvilke, og hvordan de kan unngås avhenger av mange forhold. For eksempel kan det, dersom ekspertene er bedt om å oppgi sannsynligheter for en viss hendelse, være nødvendig å trene eksperten(e) opp i sannsynlighetsangivelse og muligheten for, eller ulike årsaker til estimatskjevhet på forhånd.

Det er også viktige spørsmål knyttet til hvem, og hvor mange ekspertene skal være. Det er innlysende at ekspertene forutsettes å inneha særskilt ekspertise på det aktuelle området. Det er likeså innlysende at eksperten ikke må ha særinteresser som kan påvirke svaret som gis. I et lite land som Norge kan det representere en særlig utfordring.

De ulike tilnærmingene kan også gi ulike svar. Morgan og Keith (1995), for eksempel, gjennomførte strukturerte intervjuer med 16 ledende klimaeksperter, og fant at oppfatningene om klimaendringer divergerte i mye større grad enn det man kan få inntrykk av gjennom vitenskapelige konsensusedokumenter.

### **Noen anbefalinger**

Uansett hvilken fremgangsmåte som benyttes for å avdekke ekspertinformasjon, er det en del momenter som er viktige i forhold til å skape en prosess som er effektiv i forhold til å fremskaffe best mulig beslutningsgrunnlag. Jørgensen (2002) refererer til ulike studier som peker på følgende elementer som viktige:

- **Evalueringens nøyaktighet bør evalueres, men et høyt evalueringstrykk bør unngås.** Ulike studier kan tyde på at dersom folk føler seg personlig ansvarlig for kvaliteten på ekspertinformasjonen, oppfatter estimeringsoppgaven som svært viktig eller dersom det er en pengemessig belønning knyttet til presise estimater, så kan det faktisk føre til redusert estimatpresisjon. På den annen side kan det virke som om denne reduksjonen i estimeringspresisjon først og fremst er et problem ved komplekse vurderinger. Ved enklere estimeringssituasjoner er det derimot ting som kan tyde på at presisjonen øker med motivasjonsgraden. Det er også interessant å merke seg resultater som tyder på at økt grad av evaluering av ekspertenes estimater kan føre til økt grad av svar basert på instinkt heller enn refleksjon. Med andre ord; dersom det er snakk om en estimeringsoppgave der gode estimater i større grad er avhengig av refleksjon og analyse, bør man unngå at estimeringsprosessen i for sterk grad er preget av at evaluering.
- **Unngå motstridende mål.** Dette dreier seg bl.a. om å unngå at eksperten settes i en situasjon der evalueringen kan påvirkes av konflikt mellom realisme og ønsketenkning. Et eksempel fra fotballens verden kan være at en ekspert fra Bergen bes om å vurdere Branns vinnere i serien. Lokalpatriotisme kan lett føre til at man tillegger det lokale fotballaget høyere vinnere enn det en uavhengig vurdering skulle tilsi. Dette er noe profesjonelle fotballgjembere ofte utnytter til sin fordel ved at oddsene i så fall blir skjeve. En implikasjon av dette er eksperten(e) som utarbeider estimatene ikke bør være den som er ansvarlig for, eller delaktig i gjennomføring av prosjektet.
- **Be ekspertene om å begrunne og kritisere egne estimater.** Ekspertens tiltro til eget estimat vil avhenge av hvor mye arbeid som legges ned i utarbeidelsen. Ulike studier tyder således på at å be ekspertene om å begrunne og kritisere egne estimater fører til en mer analytisk estimeringsprosess og reduserer risikoen for at for enkle estimeringsstrategier benyttes.<sup>12</sup> Det øker også graden av tiltro til estimatet. Det

---

<sup>12</sup> Dette diskuteres noe nærmere av Kirkebøen i kapittel 9.

samme gjelder graden av estimatnøyaktighet, særlig i situasjoner preget av høy usikkerhet.

- **Unngå irrelevant og upålitelig informasjon.** Folk har en implisitt tendens til å betrakte informasjon som relevant dersom den blir presentert i samme kontekst som estimeringsproblemet. Det er imidlertid en tendens til at mer erfarne eksperter i mindre grad påvirkes av irrelevant informasjon enn mer uerfarne. Implikasjonen av dette er at utvalget av eksperter er viktig. Videre vil det være viktig å *advare mot* eksistensen av upålitelig og irrelevant informasjon.
- **Bruk dokumenterte data fra tidligere estimeringsoppgaver.** Å gjøre dokumenterte data fra relevante gjennomførte prosjekter tilgjengelig i estimeringsprosessen fører til at eksperter vil benytte en mer analytisk estimeringsstrategi, og følgelig vil estimatene være mindre eksponert for menneskelig eller situasjonsbetinget estimeringsskjevhet. Man bør imidlertid være oppmerksom på at denne strategien fungerer godt dersom tilgjengelige data i stor grad er representative for estimeringsoppgaven, og dersom man korrigerer for tidligere estimeringsfeil. Dersom det på den annen side er et stort tilpasningsbehov vil resultatet bli skjeve og upresise estimater. Unøyaktigheten i estimat kan også bli stor dersom effekten av ”regression towards the mean” er betydelig. Da kan det være nødvendig å legge til grunn et estimat som ligger nærmere gjennomsnittet enn verdien på den siste observasjonen.<sup>13</sup>
- **Finn eksperter med relevant bakgrunn og gode estimeringsresultater.** Ulike eksperter vil være dyktige på ulike områder. Noen kan være dyktige til å estimere usikkerhet, noen kan ha sakskunnskap i dybden, mens andre vil ha kunnskap i bredden. Det vil være viktig å velge riktig ekspert til aktuelle estimeringsoppgaver. På samme måte er det viktig å velge eksperter som har vist evne til å levere presise estimater på relevante områder. Eksperter som har vist at de er dyktige til å levere gode estimater bør konsulteres dersom det er *punktestimater* som etterspørres. Likeledes bør eksperter som har vist at de er dyktige til å estimere usikkerhet konsulteres det er *estimat usikkerhet* som etterspørres.
- **Benytt estimater fra flere eksperter og gjennom ulike tilnærminger.** Estimeringsprosessen kan understøttes på ulike måter, bl.a. med hensyn til hvordan estimeringsproblemet dekomponeres. I Jørgensen (2002) anbefales det at bottom-up

---

<sup>13</sup> Ikke-representative observasjoner vil kunne føre til et fenomen som noen ganger blir kalt ”regression toward the mean.” Opprinnelig var fenomenet eksemplifisert ved sammenhengen mellom foreldre og barnas høyde og ”regression towards mediocrity”; ekstreme verdier på en variabel vil, i gjennomsnitt, sammenfalle med mindre ekstreme verdier på den andre variabelen. Uttrykket benyttes gjerne til å forklare hvorfor det ofte er slik at de som lå øverst eller nederst på en indikator hadde en tendens til å ligge nærmere gjennomsnittet ved en ny måling; Et eksempel kan benyttes for å forklare dette nærmere: Anta studenter som tar en eksamen midt i semesteret og så en endelig eksamen. Studenter som fikk en ekstremt god karakter ved midtveiseeksamenen vil sannsynligvis få en god karakter ved den endelige eksamenen, men vi forventer at deres karakter likevel vil ligge nærmere gjennomsnittet. Årsaken til dette er at flaks trolig var noe av grunnen til at de gjorde det dårlig på den første prøven (i tillegg til gode kunnskaper), og det er lite sannsynlig at alle har like mye flaks to ganger på rad. Det vil derfor være en tendens til at neste måling viser at de beste gjør det noe dårligere og at de dårligste gjør det noe bedre. Dette er et rent statistisk fenomen, og det er misvisende å tilskrive denne tendensen eventuelle tiltak man har gjennomført i mellomtiden. På samme måte er det misvisende å anta at man har hatt en forbedring (eller forverring) mellom to målinger dersom en av målingene gir et gjennomsnitt som ikke er representativt for hele populasjonens underliggende gjennomsnitt.

tilnærminger kombineres med top-down. Disse to ulike angrepsmåtene har sine respektive styrker og svakheter, men dersom begge tilnærmingene gjennomføres uavhengig av hverandre, bl.a. for å unngå såkalte ”anchoring” effekter, dvs. at estimater blir påvirket av utgangspunktet (tradisjon, historikk eller tilgjengelige data). Videre kan estimeringsprosessen understøttes ved at det benyttes flere eksperter. Dette reiser igjen flere spørsmål, bl.a. hvor mange eksperter som bør benyttes, hvordan ekspertenes estimater kan kombineres, og hvordan man kan korrigere enkelteksperters estimater i den grad man har informasjon som tyder på at vedkommende systematisk gir skjeve estimater. Dette er spørsmål som vi vil komme nærmere tilbake til.

Antall eksperter som bør benyttes i estimeringsprosessen avhenger av deres forventede presisjon, estimeringsskjevhet og inter-korrelasjon. Jørgensen (2002) peker på at 3-5 forskjellige eksperter med ulik bakgrunn som oftest vil være tilstrekkelig.

I den grad det er tilgjengelig informasjon om ekspertens tidligere estimeringspresisjon, bør denne, i den grad historikken er relevant, benyttes til å redusere estimeringsskjevhet. Hvis det for eksempel er kjent at en ekspert har en tendens til å være systematisk overoptimistisk i sine estimater, bør denne informasjonen benyttes. Hvordan denne korreksjonen kan gjennomføres kommer vi tilbake til senere.

Det vil også være en fordel å velge eksperter med ulik erfaringsbakgrunn, fordi høy inter-korrelasjon kan forventes dersom man velger eksperter som i stor grad har samme bakgrunn.

## 5.4 Hva skal eksperten anslå?

Når det er lang tid frem til et prosjekt eventuelt besluttes gjennomført, vil det være betydelig usikkerhet knyttet til de underliggende variable som er relevant for å vurdere prosjektet. I tillegg vil det selvsagt være usikkerhet som har sitt opphav i selve estimeringsprosessen, dvs. mangel på presis kunnskap om det systemet man ønsker å estimere.

I enkelte tilfeller vil det foreligge et sett av et visst objektivt erfaringsmateriale, f.eks. med hensyn til konstruksjonskostnader ved lignende prosjekt, av relevans for et gitt prosjekt. På bakgrunn av dette datagrunnlaget kan det settes opp en sannsynlighetsfordeling for kostnadsutfall. Er prosjektet svært likt, og det er få andre faktorer som særpreger dette prosjekt, vil en slik sannsynlighetsfordeling være smal, dvs. ha liten varians, og vi kan feste en rimelig høy grad av sannsynlighet til forventningsverdien. Men dessverre er ikke alltid virkeligheten slik. Objektive erfaringsdata kan være av beskjedent omfang, og vårt prosjekt har flere særegne karakteristika som gjør at relevansen av historiske erfaringsdata er begrenset.

En alternativ tilnærming i slike situasjoner, er å legge opp til en subjektivistisk (Bayesiansk) tilnærming for sannsynlighets- og usikkerhetsvurdering. I følge denne tilnærmingen er sannsynligheten for et visst utfall (et verdiestimat), den graden av tro (”Degree of belief”) en person har på dette utfallet, gitt all relevant informasjon som denne eksperten i øyeblikket besitter. All relevant informasjon vil bestå av tidligere objektivt erfaringsmateriale, informasjon om hva som er spesielt og særpreget ved dette prosjektet, og ekspertens vurdering av hvordan dette samlet påvirker sannsynligheten for et bestemt utfall. Derfor kan selvfølgelig to ulike eksperter gi ulike sannsynligheter for samme utfall, og en og samme person kan gi forskjellige sannsynligheter ved ulike tidspunkter.

Det er likevel bestemte formelle krav til subjektivistiske sannsynligheter, de kan ikke tilordnes vilkårlig. Blant annet må de være konsistente med aksiomene i sannsynlighetsteori. For eksempel dersom  $P(A)$  tilordnes estimatet  $A$ , må  $1-P(A)$  tilordnes komplementære utfall til  $A$ . Et annet krav er at utfallet  $A$  er velspesifisert.

### Punkttestimat og aggregering

Dersom flere eksperter blir benyttet, og de er bedt om å gi sitt beste estimat på den samme variabelen, vil det ofte vil det være sprikende oppfatninger både når det gjelder estimater og anslag av sannsynligheter og sannsynlighetsfordelinger. Det er flere teknikker for å koordinere slike forskjeller, herunder informasjon over andres anslag med mulighet for justering av egne anslag, jf. Delfi-metoden eller gruppediskusjoner av eksperter. En vanlig tilnærming er også å benytte en gjennomsnittsverdi, evt. vektet gjennomsnittsverdi. En oversikt over de ulike teknikkene finnes i Chatterjee og Chatterjee (1987) og Meyer og Booker (2001). Vi vil se nærmere på ulike tilnærminger til å benytte gjennomsnittverdier.

Den vanligste tilnærmingen til å kombinere et sett av svar er å beregne en enkelt verdi basert på alle verdiene i datasettet. Ulike uttrykk for sentraltendensen i et datasett er gjennomsnitt, median, eller geometrisk gjennomsnitt; hver estimator med egenskaper som gjør dem hensiktsmessig å benytte i ulike sammenhenger. På bakgrunn av Kahneman, et al. (1982) anbefaler Meyer og Booker (2001) at medianverdien benyttes når eksperter har avgitt ulike numeriske estimater. Begrunnelsen for dette er at i når en ekspert gir et numerisk svar, estimeres medianverdien heller enn gjennomsnittsverdien. Dersom svarfordelingen hadde vært symmetrisk hadde det ikke spilt så stor rolle om man benyttet det ene eller det andre, men svarfordelingen er typisk skjev med modalverdien på den ene siden og en lang hale på den andre siden av fordelingen.

Ved beregning av gjennomsnitt, median, eller geometrisk gjennomsnitt vektes hver observasjon likt. Det er flere grunner til at dette kan være problem i en ekspertvurderingssammenheng, ikke minst at det kan være grunn til å tillegge enkelte eksperters vurdering høyere eller lavere vekt enn andres, bl.a. ut fra vedkommendes estimeringshistorikk. I praksis vil vektingen skje på følgende måte:

$$\text{Vektet gjennomsnitt} = \frac{\sum_{i=1}^n x_i w_i}{\sum_{i=1}^n w_i},$$

der  $w_i$  er vekten som tillegges den enkelte ekspert  $i$ 's estimat  $x_i$ . Ulempen med denne tilnærmingen er at vektene må bestemmes for hver enkelt ekspert, noe som utover den utfordringen det representerer i seg selv, også medfører at estimatene kan påvirkes av subjektive oppfatninger hos den eller de som fastsetter vektene.<sup>14</sup>

---

<sup>14</sup> I praksis er problemet enda mer intrikat fordi ekspertenes ekspertise vil variere. Vekten kan altså variere for hvert spørsmål. Dersom man bare har generalister i gruppen kan vektingen være stabil.

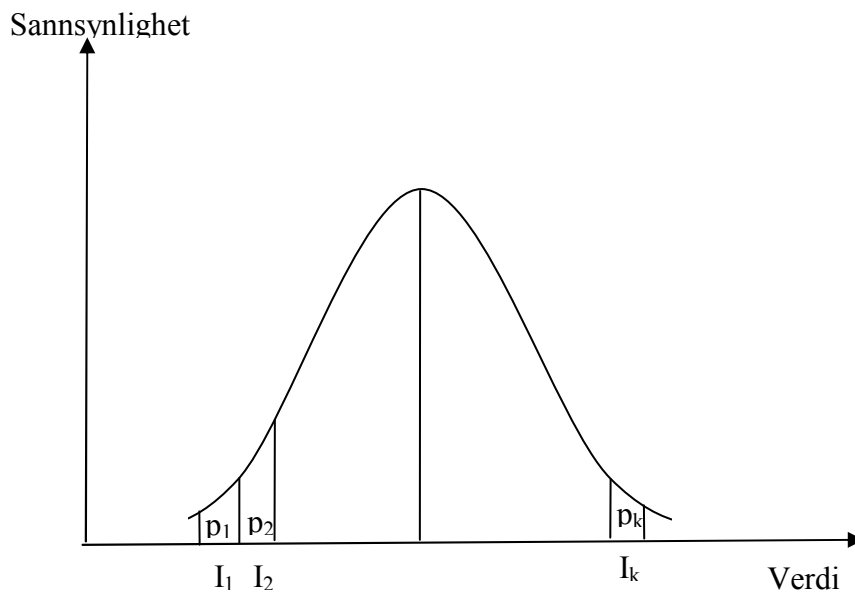


Meyer og Booker (2001) peker på at å bestemme vektene vil kreve kunnskap om den enkelte ekspert og hvordan denne kommer frem til sitt svar. Det er likevel ulike tilnærminger som kan benyttes. Noen av disse er:<sup>15</sup>

- **Databasert bestemmelse av vekter.** Fastsettelsen av vekter skjer på bakgrunn av datasettet, f.eks. slik at en ekspert som avgir svar som utgjør en ekstremverdi tillegges mindre vekt enn de andre.
- **Saaty-metoden.** Dette er en metode basert på informasjon om ekspertene, og der vektene fastlegges basert på parvise sammenligninger av ekspertene ut fra denne informasjonen. Metoden baserer seg altså på kvalitativ informasjon om ekspertene, men gir på dette grunnlaget kvantitative vekter.

### Fordelinger

Et alternativ til å be eksperten om et best mulig punkttestimat, er å be om et anslag på sannsynligheten for en hendelse eller hele fordelingen av mulige verdier (sannsynlighetsfordelingen). Her vil vi se kort på noen ulike tilnærminger som kan benyttes for å anslå fordelinger, samt noen feilkilder.

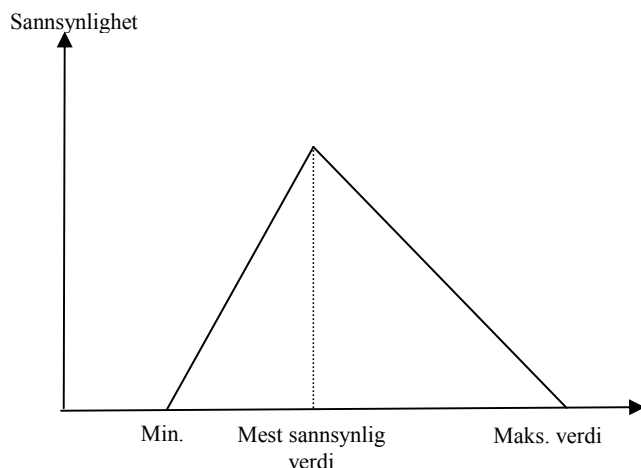


Figur 5-1. Intervallmetoden

- Intervall-metoden medfører å anslå sannsynligheten for at den størrelsen vi er interessert i ligger i visse intervaller. Intervallene  $I_1, I_2, \dots, I_k$  oppgis, og sannsynlighet  $p_1, p_2, \dots, p_k$  for hvert intervall anslås av eksperten, jf. figuren over.
- Fraktilmetoden går grovt sagt ut på at sannsynlighetene  $p_1, p_2, \dots, p_k$  er spesifisert på forhånd, og så er det ekspertenes oppgave å anslå de tilhørende intervaller.

<sup>15</sup> Mer utdypende informasjon om de ulike tilnærmingene kan finnes i Meyer og Booker (2001).

- Grafisk metode innebærer rett og slett ut på at eksperten tegner opp fordelingen slik han tror den er.
- Parametrisk metode innebærer å anta at fordelingen er av en viss type, for eksempel normal, lognormal, beta eller trekant, og eksperten gir inngangsverdier slik at fordelingen blir spesifisert. Typisk blir et trippelanslag benyttet, jf. figuren under.



Figur 5-2. Parametrisk metode

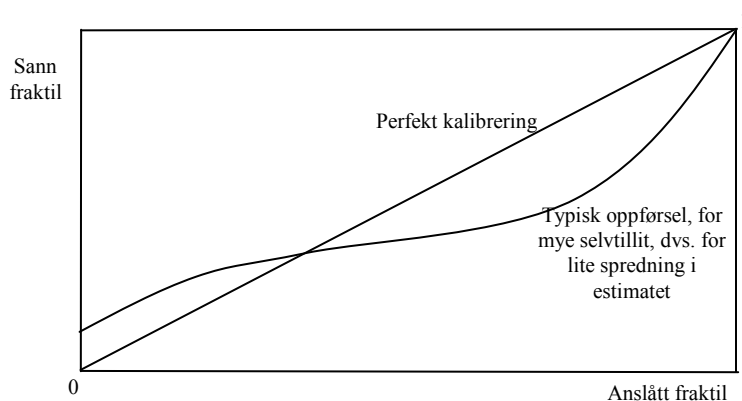
Eksperten må da anslå minimums, maksimums og mest sannsynlig verdi. Tilnærmingen kan tilpasses ulike fordelingstyper. Den enkleste tilnærmingen er å anta uniform fordeling. Eksperten anslår da bare  $[a, b]$ , der  $a$  er minimums- og  $b$  er maksimumsverdi, men lik (uniform) sannsynlighet for alle mellomliggende verdier. At tilnærmingen er enkel er også svakheten, den tillegger null sannsynlighet til verdier som ligger utenfor  $a$  og  $b$ . Det er heller ikke særlig sannsynlighet at alle mellomliggende verdier har samme sannsynlighet. Problemet reduseres noe ved å anta at den underliggende fordeling er triangulær. Her anslår eksperten i tillegg modus for å bestemme fordelings toppunkt, men fortsatt gjenstår problemet med null sannsynlighet for verdier utenfor maks- og minimumspunktet. En annen fordelingstype som ofte benyttes som utgangspunkt for å tilpasse ekspertenes estimater er betafordelingen. Denne benyttes ofte ut fra matematisk hensiktsmessighet. Eksperten kan ikke angi parametrene til betafordelingen direkte, men det er forholdsvis kurant å gå fra et anslag av fraktil- og forventningsverdier til å beregne de nødvendige parametrene.

Ved at eksperten får tilbake et skjermbilde av fordelingen kan vedkommende i ettertid justere sine estimater dersom det er ønskelig å gjøre dette. Denne metoden er kanskje den mest vanlige. Den er ikke minst anvendt i petroleumssammenheng, og da i sammenheng med reservoarmodellering. En oversikt og drøfting finnes i Gosling (2005).

Hvorvidt forutsetningen om fordelingstype er den rette kan alltid diskuteres. Ofte kan det virke som om fordelingstype velges mer ut fra matematisk hensiktsmessighet enn ut fra å representere usikkerhet på en mest mulig korrekt type. Det vil således være viktig å drøfte nøye hvilken type fordeling som er mest korrekt gitt den aktuelle rammen.

Videre er det også viktig å være oppmerksom på ulike kilder til estimeringsskjevhet. Det samme gjelder tendensen til at ekspertene har for stor tillit til presisjonen i egne estimat. Når metodene anvendes i situasjoner der de svaret kan evalueres, har de alle en tendens til å gi for

lite spredning. Dette gjelder spesielt fraktilmetoden og grafisk metode. Dette kan gjøre det nødvendig med kalibrering. En typisk kalibreringskurve vil kunne se ut som illustrert under.



Figur 5-3. En typisk kalibreringskurve

Ved den parametriske metoden er det meningen at eksperten skal kunne justere sitt anslag etter å ha sett skjerm bilde av det. Det kan her være av verdi at eksperten kjenner sin egen kalibreringskurve, slik at han har anledning til å rette opp eventuelle systematiske skjevheter. En annen viktig korrigeringsfaktor består i at eksperten blir bedt om å forklare sine estimater, noe som i seg selv kan føre til justering. Generelt vil det også settes store krav til intervjueren ved slike undersøkelser. Det finnes også andre tilnærminger til å gjennomføre, avdekke og korrigere ekspertvurderinger, f.eks. gruppeprosesser, jf. Øien, et al. (1996).

Subjektive sannsynlighetsfordelinger for at en gitt hendelse kan generelt sett oppnås på to ulike måter; enten å be eksperten om verdier som tilsvarer et bestemt prosentil i sannsynlighetsfordelingen, eller ved å be eksperten om å vurdere sannsynligheten for at realisert verdi overtiger en gitt størrelse. Tversky (1974) peker på at de to tilnærmingene formelt sett er ekvivalente, og skulle resultere i den samme fordelingen. Problemet er at *forankringspunktet* ved de to ulike tilnærmingene er forskjellig. Ved den første tilnærmingen gir eksperten et svar i form av et verdiesimat i den gitte prosentil, og det naturlige forankringspunkt for å foreta denne vurderingen er ekspertens egen vurdering av den mest sannsynlige verdi. Ved den andre tilnærmingen gis svaret i form av sannsynligheter eller odds, og det naturlige forankringsutgangspunktet er jevne odds eller en sannsynlighet på 50%. Tversky peker videre på at forankringsutgangspunktet ved den andre tilnærmingen vil gi konservative odds, dvs. odds som er for nærme 1:1 og en sannsynlighetsfordeling som er for flat. Som en løsning på dette forankringsproblemet, foreslår Tversky at en passende kalibrering kan oppnås ved å kombinere begge tilnærmingene.

Generelt er det viktig at tilnærmingen som benyttes er enkel og intuitiv å benytte. Det bør også være slik at man søker å benytte tilnærminger som eksperten er komfortabel med å benytte, heller enn å presse eksperten til å benytte en tilnærming som er ”ferdigsneket” ut fra behovet til de som skal bearbeide materialet frem til beslutningstaking. Hvis det benyttes en annen tilnærming enn den eksperten er kjent med, bør det settes av tid slik at eksperten blir fortrolig med metodikken.

En alternativ tilnærming, og ikke minst lovende, tilnærming er å benytte såkalt *fuzzy logikk*. Eksperten blir da bedt om å uttrykke kunnskapen sin i en form som vedkommende gjerne er mer kjent og komfortabel med. I stedet for å spesifisere verdier for ulike deler av fordelingen for en usikker størrelse, uttrykkes kunnskapen i lingvistisk form, for eksempel lav, medium, høy. En tilnærming basert på fuzzy logikk vil være særlig nyttig i sammenhenger der kvalitative erfaringsdata er begrenset, men der eksperten likevel kan uttrykke sin erfaring og kunnskap om sammenhenger mellom kostnader og prosjektgenskaper i lingvistisk, kvalitativ form. En praktisk anvendelse kan finnes i bl.a. Kaminaris, et al. (2006) og Mason og Sr. (1997). Dersom flere eksperter benyttes, oppstår aggregeringsutfordringer av samme type som vi vil komme nærmere inn på senere. Fuzzy logikk er for øvrig et tema for et eget kapittel i denne boken.

### Hvordan aggregere fordelinger

På samme måte som ved punkttestimat, vil ulike eksperter gi svar som resulterer i ulike fordelinger. I hovedsak er det to ulike tilnærminger til å kombinere fordelinger; den ene er å avdekke enkeltteksperternes fordelinger for så å aggregere disse gjennom ulike tilnærminger, og som vi skal komme nærmere inn på. Den andre er tilnærmingen er å prøve å identifisere en konsensusfordeling. Man kan også tenke seg kombinasjoner av disse to metodene. Her vil vi fokusere på problemstillinger knyttet til aggregering av fordelinger.

Vi antar altså at  $m$  ulike eksperter er konsultert, og hver konsultasjon resulterer i en sannsynlighetsfordeling  $\pi_i(\theta)$ , og som da oppsummerer ekspertens estimat for  $\theta$ .  $\theta$  kan være en gitt hendelse, for eksempel et kostnadselement eller et verdielement i et prosjekt. Vi har altså  $m$  ulike fordelinger. Spørsmålet er så hvordan de ulike fordelingene kan aggregeres slik at det kan utgjøre en del av beslutningsunderlaget. Berger (1985) viser til at det er mange *ad hoc* tilnærminger som har blitt foreslått, og at disse to kan deles i to kategorier<sup>16</sup>:

- **Lineær sammenslåing av oppfatninger.** Ved denne tilnærmingen tildeles hver informasjonskilde  $\pi_i(\theta)$  en vekt  $w_i$  (der  $\sum_{i=1}^m w_i = 1$ ), der vekten typisk reflekterer påliteligheten til den gitte informasjonskilden. Så benyttes  $\pi(\theta) = \sum_{i=1}^m w_i \pi_i(\theta)$  som den aggregerte sannsynlighetsfordelingen for  $\theta$ .
- **Uavhengig sammenslåing av oppfatninger.** Denne tilnærmingen blir benyttet dersom informasjonskildene tilsynelatende er uavhengige. Ved denne tilnærmingen finnes  $\pi(\theta)$  som  $\pi(\theta) = k \prod_{i=1}^m \pi_i(\theta)$ , der  $k$  er en normaliseringskonstant.

Vi ser tilnærmingene bl.a. benyttet i Meyer og Booker (2001)), der disse aggregeringsmetodene riktig nok kombineres med en bayesianske tilnærminger, dvs. at den aggregerte fordelingen kombineres med beslutningsfatterens egen oppfatning ved hjelp av Baye's teorem.

---

<sup>16</sup> Se også en god oversikt i Jenkinson (2005).

Berger (*op.cit*) peker på at til tross for at dette er tilnærminger som er i alminnelig bruk så er det flere svakheter. Et eksempel tilpasset fra Meyer kan benyttes for å illustrere et sentralt problem: Anta at vi har to geologer som anslår sannsynligheten for at det vil komme henholdsvis intet, middels eller mye vanngjennombrudd i et oljereservoar til henholdsvis  $\pi_1 = (0,1; 0,7; 0,2)$  og  $\pi_2 = (0,0; 0,8; 0,2)$ . Det vil si at det første ekspertens anslår at det er 10% sannsynlig at det ikke vil bli vanngjennombrudd i reservoaret. Hvis vi antar uavhengighet og tillegger begge ekspertene lik vekt, vil lineær sammenslåing av sannsynlighetene innebære at  $\pi = (0,05; 0,75; 0,2)$ . Vi ser imidlertid at den andre ekspertens har anslått at det er utenkelig at et visst vanngjennombrudd ikke vil skje, dvs. at sannsynligheten for at det ikke skjer vanngjennombrudd er lik null. Det er rimelig å gå ut fra at denne vurderingen er basert på informasjon som ikke er tilgjengelig for den første ekspertens, for eksempel erfaringer med akkurat denne reservoartypen på denne aktuelle dybden. Det er derfor urimelig at begge ekspertene skal tillegges like stor vekt. Hvis den første ekspertens ikke er klar over dette erfaringsmaterialet, kan en bedre tilnærming heller være å betrakte  $\pi_1$  som apriorifordelingen, og  $\pi_2$  som grunnlag for sannsynlighetsoppdatering. Bayes teorem gir da, som *a posteriori*-fordeling:

$$\pi = \frac{(0,1 * 0; 0,7 * 0,8; 0,2 * 0,2)}{[0,1 * 0 + 0,7 * 0,8 + 0,2 * 0,2]} = (0; 0,9333; 0,0667)$$

I dette tilfellet har ekspertenes oppfatning blitt forsterket av hverandre. At dette ikke er mulig, er nettopp en svakhet ved metoden ”lineær sammenslåing av oppfatninger”. Den samme fordelingen ville riktignok ha resultert hvis vi hadde benyttet den andre tilnærmingen; ”uavhengig sammenslåing av oppfatninger”, men problemet her er at det ikke er mulig med en delvis forsterkning av oppfatninger, noe som ofte vil være ønskelig. Det vil også svært ofte være slik at ekspertenes oppfatninger ikke kan betraktes som uavhengige. De vil i varierende grad ha samme utdannelse, til dels samme erfaringsbakgrunn osv. Dette vil resultere i avhengighet, og skape teoretiske problemer ved tilnærmingen ”uavhengig sammenslåing av oppfatninger”.

Berger (1985) peker at det ikke finnes noen universalløsning på aggregeringsproblemet, men at alternativet til de nevnte *ad hoc* tilnærminger til å kombinere ulike eksperters oppfatninger, er en nøye probabilistisk modellering av situasjonen, kombinert med probabilistisk behandling, hvorav Bayes teorem er ett aspekt. Dette innebærer å beregne *a posteriori*fordelingen for  $\theta$  på bakgrunn av all kjent informasjon, herav ekspertenes (aggregerte) oppfatning er ett element. Beslutningstaker blir ved en slik tilnærming gjerne benevnt en ”supra-Bayesian” (Jenkinson (2005)). Beslutningstaker oppdaterer da en apriorifordeling basert på informasjon bl.a. fra ekspertene.

## 5.5 Ekspert er også mennesker

Det er en del typiske feilkilder ved ekspertvurderinger ut over de som vi allerede har vært inne på. Omfattende psykologisk forskning har vist at selv eksperter finner det vanskelig å anslå sannsynligheter for en visse hendelse. For å gjøre oppgaven enklere, benyttes tommelfingerregler (heuristikk). Slike tommelfingerregler kan imidlertid også føre til

skjevheter i estimatene. De mest vanlige årsakene til estimatskjevhet er for stor grad av selvtilit, forankring og tilgjengelighet. Disse vil vi komme litt nærmere inn på her.<sup>17</sup>

Det finnes også metodikk for å forsøke å unngå disse skjevhetene når ekspertene gir sine anslag. Det vil vi komme inn på etterpå.

### **Tilgjengelighet**

Estimering av verdier og tilhørende sannsynligheter bestemmes i noen grad av hvorvidt eksperten kan tenke eller huske på tidligere forekomster av begivenheten (frekvenstankegang) eller i hvilken grad han kan forestille seg begivenheten, dvs. av *tilgjengelighet*.

Det vil fungere bra dersom ekspertens erfaring og hukommelse stemmer noenlunde overens med virkelige frekvenser for hendelsen. En får imidlertid en tendens til overestimering dersom søkingen i hukommelsen blir forsterket (for eksempel ved mediaomtale) og fører til underestimering dersom søkingen i hukommelsen er vanskelig (lite omtalt, involverer abstrakt begrep osv.). Sannsynligheten for hendelser som har fått en omfattende mediaomtale tenderer å bli overvurdert, mens sannsynligheten for hendelser som skjer sjelden gjerne blir undervurdert. Dersom en ekspert nylig har vært involvert i et prosjekt som på grunn av svært gunstige omstendigheter ble gjennomført på langt kortere tid enn det som er vanlig for tilsvarende prosjekt, kan det være at det er disse nylige erfaringene som preger estimatene som gis for et nytt lignende prosjekt. Ekspertene som har et problem med tilgjengelighetsskjevhet, vil legge for stor vekt på hendelser som de lett kan forestille seg eller huske.

I scenariotenkning viser det seg at personer har en tendens til å overestimere sannsynligheter for forekomst av scenarier som knytter sammen ulike sannsynligheter. Sammenknyttede begivenheter er lettere å forestille seg enn individuelle isolerte begivenheter.

### **Representativitet**

Personer kan la seg forvirre av i) detaljer som ikke er representative for fenomenet og/eller ii) ekstra informasjon som er irrelevant. Dette er blitt undersøkt blant andre av Kahneman og Tversky, se f.eks. Kahneman og Tversky (1982).

Et eksempel på dette er følgende: I et myntkasteksperiment med 6 myntkast vil sekvensen M K M M K M oftest bli antatt å være mer sannsynlig enn sekvensen M M M M M M eller M M M K K K, selv om de alle er like sannsynlige. Sekvensen M K M M K M blir her oppfattet som mer representativ for den tilfeldige mekanismen.

Et annet kjent eksempel er følgende: Det er oppgitt at et utvalg av personer inneholder 30 ingeniører og 70 jurister. En person blir karakterisert uten at hans profesjon oppgis, og oppgaven er å anslå sannsynligheten for at vedkommende er ingeniør. Dersom det ikke gis annen informasjon vil sannsynlighet som oftest estimeres korrekt til 30%. Anta så at den som blir bedt om å vurdere sannsynligheten blir presentert for følgende karakteristik av den som er trukket ut av utvalget: ”Vedkommende trukket ut er dyktig, motivert, godt likt og har

---

<sup>17</sup> Den interesserte leser henvises til Gilowitch, et al. (2002), som oppsummerer den mest innflytelsesrike forskningen som har blitt gjennomført etter den initiale artikkelsamlingen på området i Kahneman, et al. (1982). En god oversikt finnes også i Bazerman (2001).

gode sjanser for en vellykket karriere”. Denne irrelevante karakteristikkk fører ofte til at sannsynligheten feilaktig anslås til rundt 50%. Videre, dersom karakteristikkk som blir presentert er: ”Ingen interesse i politiske og sosiale saker. Hobbies: Hjemmesnekring, seiling, matematiske nøtter”, fører denne informasjonen ofte til at sannsynligheten for at vedkommende er ingeniør settes nær 100%, selv om dette er tvilsomt i lys av utvalgets sammensetning.

### **Forankring og justering**

Når det finnes noe informasjon om et fenomen, vil denne påvirke vår oppfatning og vurdering av annen informasjon om samme fenomenet. Dette kan forårsake store skjevheter i vurderingene.

Ofte starter vurderingsarbeidet ut fra et forankringspunkt, som ofte kan være relativt tilfeldig satt. Som et eksempel kan vi tenke oss at en ekspert, før han foretar sin egen vurdering, leser tidligere eksperters vurderinger. Informasjonen i disse papirene, som alltid er farget av diverse briller, vil nå danne et forankringspunkt som lett vil justere fortolkningen av ekspertens egne observasjoner.

Metoden (og svakheten) består altså i at et anslag først velges som en første approksimasjon og dette justeres så for å ta hensyn til eventuell supplerende informasjon. Typisk feil er at det justeres for lite, slik at en får en skjevhet i estimatet mot forankringen.

### **For stor selvtilit**

Eksperters med for stor grad av selvtilit resulterer i sannsynlighetsfordelinger som er for smale. Det er i de undersøkelser som er foretatt, vist at det er en generell tendens mot for stor selvsikkerhet. Fischhoff (1982) peker på at undersøkelser tyder på at sannsynligheter på 75% knyttes til en ”treffrate” på bare 60%, og at i tilfeller der det gis uttrykk for sikkerhet, dvs. 100% sannsynlighet, treffes målet bare i 85% av tilfellene. I tilfeller der det blir bedt om et estimat i form av numeriske verdier, er det ikke uvanlig å finne at ”fasiten” ligger utenfor det estimerte 98% konfidensintervall i 20 til 40% av tilfellene. Fischhoff mener naturlig nok at dette er resultater som er bekymringsfulle, både for de som er avhengige av gode estimater, og for dem som direkte eller indirekte overdriver hvor mye de vet.

Estimatskjevhet på bakgrunn av dette problemet kan reduseres ved at ekspertene gis tilbakemelding på sine estimater, eller enda bedre; at de blir bedt om å begrunne sine estimater eller eventuelt resonnerer rundt årsaker til at estimatet deres kan være galt.

Ulike psykologiske årsaker til estimatskjevheter er tema som det er forsket en del på også i Norge. Problemstillingen er dekket nærmere i kapittel 9.

## **5.6 Hvordan korrigere**

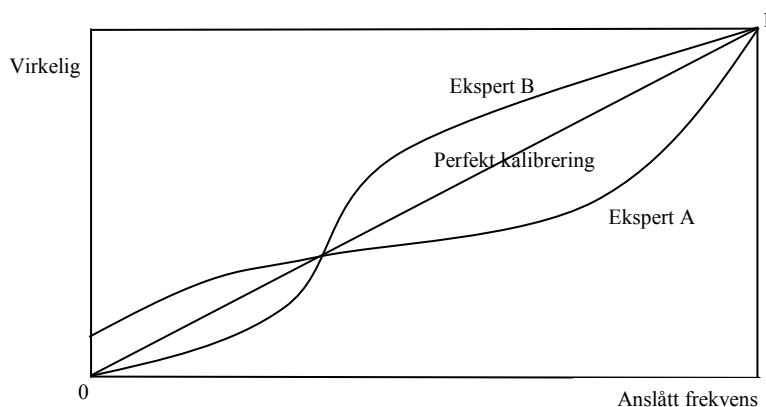
Det finnes ulike tilnærminger til å korrigere skjevheter i ekspertenes estimater, alt etter hva årsaken til skjevheten er. Det kan for eksempel være at eksperten er stilt overfor en for vanskelig estimeringsoppgave eller at estimeringsoppgaven er misforstått. I så fall bør oppgaven klargjøres bedre eller gjøres enklere. Semantikk er viktig i denne sammenhengen, og det som er innlysende for den som stiller spørsmålene kan oppfattes på en helt annen måte

av den som blir spurt. Oppgaven kan videre gjøres enklere ved å klargjøre instruksene, stille færre spørsmål, stille spørsmålene på en annen måte eller ved å illustrere med eksempler.

Til slutt kan det også være mismatch mellom kunnskapen til ekspert(ene) og estimeringsoppgave. I så fall kan problemet dekomponeres eller man kan vurdere alternative tilnærminger.

Det kan også være at det er eksperten det er noe feil med. Hvis eksperten har læringsevne, kan løsningen være å advare eller beskrive problemet, gi tilbakemeling, be eksperten kritisere eget estimat og gjennomføre grundig trening i forkant. I motsatt fall kan løsningen være å foreta en kalibrering, gitt at vi har informasjon som tillater en fornuftig kalibrering av avgitte svar.

Det finnes ulike muligheter for å kalibrere en eksperts sannsynlighetsvurderinger. Det kan for eksempel gjøres ved at eksperten anslår sannsynligheten for utfall der sannsynlighet kan verifiseres. En sammenlikning mot faktisk sannsynlighet kan så foretas. Eventuelt kan vi ha informasjon om vedkommendes estimeringshistorikk. Dette resulterer i et kalibreringsdiagram som vist under.



Figur 5-4. Kalibrering av subjektive sannsynligheter fra to ulike eksperter A og B.

Ekspert A er her en person med for mye selvtillit, dvs. for lite spredning i anslaget. Det er for mange sannsynligheter nær 0 og 1. Ekspert B derimot har for lite selvtillit. Dette er informasjon som så kan benyttes for å justere de ulike eksperters anslag, se f.eks. Morgan og Henrion (1990).

Vi har allerede vært inne på ulike tilnærminger til å redusere for stor grad av selvtillit, for eksempel ved at eksperten blir gjort oppmerksom på fenomenet, at eksperten blir bedt om å begrunne og kritisere egne estimat eller at eksperten kan revidere eget estimat i lys av andres estimat. Kalibrering av eksperten gjort i forkant av den undersøkelsen en ønsker å gjennomføre, kan også føre til at systematiske feil reduseres, bl.a. gjennom at eksperten blir oppmerksom på eventuelle egne skjevheter. En svakhet ved kalibreringsmetoden er at det ofte er vanskelig å kalibrere direkte på ekspertens eget spesialområde pga. manglende data.



## 5.7 Et praktisk eksempel

På bakgrunn av store kostnadsoverskridelser i store offentlige prosjekter i USA på nittitallet, foreslår Dillon, et al. (2002) en tilnærming til kostnadsestimering basert på ekspertestimer. Prosjektene som skulle kostnadsestimeres var alternative forsyningskilder for tritium. Vi vil se nærmere på denne anvendelsen av ekspertestimering, nettopp fordi den representerer en praktisk tilnærming til, og drøfting av mange av de problemstillingene vi har vært inne på så langt. En annen interessant artikkel med to praktiske eksempler, og en drøfting av metodiske utfordringer er presentert i O'Hagan (1998).

I grove trekk bestod prosedyren av følgende elementer:

1. Identifiser og velg en variert gruppe med kostnadsestimeringseksperter.
2. Velg hensiktsmessige kostnadskomponenter og en kostnadsmodell.
3. Estimer base-case kostnader for hver kostnadskomponent.
4. Gjennomfør opplæring og trening i sannsynlighetsvurdering og be ekspertene vurdere base-case kostnader.
5. Fastsett hver eksperts sannsynlighetsfordeling for hver kostnadskomponent.
6. Aggreger ekspertenes estimer til et estimat for de totale kostnader.
7. Dokumenter prosessen.

Dillon, et al. (2002) valgte 22 kostnadseksperter fra ulike institusjoner og kontraktører. Ekspertene hadde særskilt kompetanse på området, men representerte likevel ulike innfallsvikler. Det ble lagt vekt på å sikre at ekspertene var villig til å tenke i form av sannsynligheter og usikkerhet, og var komfortable med å estimere de viktigste kostnadskomponentene for de ulike alternativene. Ekspertene ble samlet i to arbeidsseminarer med to ukers mellomrom. I perioden mellom arbeidsseminarene jobbet ekspertene med estimatene individuelt.

Som utgangspunkt for kostnadsestimeringen ble det utviklet base-case estimer, uten at det var tatt høyde for alle eventualiteter. Ekspertene var oppmerksom på at dette var estimer uten høyde for eventualiteter, at det var optimistiske estimer, og at det var en rekke ulike forutsetninger som lå til grunn. Dokumentasjon for disse forutsetningene forelå.

Treningssesjonen ble ledet av analytiker med kunnskap i sannsynlighetsteori, statistikk, beslutningsanalyse og kognitiv psykologi. Dette ble betraktet som viktig for å gi ekspertene trening i å foreta vurderinger, informere dem om årsaker til mulige estimatskjevheter, og å motivere dem. De øvelsene som ble gjennomført ble evaluert, og det ble gitt tilbakemelding på hvordan de vanlige årsakene til estimatskjevheter kunne unngås.

Hver ekspert ble så bedt om å gi forhåndsbestemte fraktiler<sup>18</sup> for sannsynlighetsfordelingen for hver kostnadskomponent. Strukturen i skjemaet er gjengitt under:

Tabell 5-1. Kostnadsestimeringsskjema

Kostnadskomponent	Base-case estimat	MIN	0,10	0,50	0,90	MAX	Kommentar
Komponent 1	xxxUSD						
Komponent 2	yyyUSD						
.....	...						
Komponent n							

Fraktilene lav (10%), median (50%) og høy (90%) ble valgt fordi det ble antatt å være lettere å anslå disse, enn de mer ekstreme og mer spesielle tilfellene 1 og 99% eller for eksempel 5 og 95%. Ekspertene måtte begrunne sine estimater, i tillegg til at de ble bedt om å vise sine beregninger og oppgi de datakilder som underbygde estimatene. For viktige kostnadskomponenter ble de i tillegg også bedt om å oppgi plausible maksimums og minimumsestimater.

På grunnlag av disse estimatene ble det utarbeidet sannsynlighetsfordelinger (Weibull<sup>19</sup>) for det enkelte kostnadsestimat. Monte Carlo simulering ble brukt for å produsere overordnet kostnadsfordeling basert på den enkelte eksperts delkostnadsestimat og så for å aggregere kostnadsestimat på tvers av ekspertene.<sup>20</sup>

Ikke uventet var det stor variasjon i de ulike ekspertenes estimater og resulterende fordelinger. Etter aggregering av fordelingene, ble den endelige fordelingen gjennomgått av ekspertgruppen for å avdekke enighet eller uenighet, og for å sikre at konsensusfordelingen var representativ for hele gruppens oppfatning. Etter at ekspertene hadde estimert grunnlaget

<sup>18</sup> En fraktil  $f_z$  er den verdien på kostnadsestimatet som er slik at det er en  $z\%$  sannsynlighet for at den realiserte verdien ligger under  $f_z$  og en  $100-z\%$  sannsynlighet for at verdien ligger over dette nivået.

<sup>19</sup> Hvilken fordeling som er den riktige kan diskuteres. Betafordelingen er en annen som også benyttes i slike sammenhenger (Meyer og Booker (2001)). Weibullfordelingen er svært mye benyttet i sammenheng med vurdering av driftsforstyrrelser og lignende. Fordelingen defineres vanligvis med utgangspunkt i parameter for lokalisering, skala samt form. Når man har med ekspertvurderinger å gjøre, er imidlertid parameterverdiene ofte uttrykt i form av høyeste verdi, laveste verdi, eller en viss prosentilverdi. Det er imidlertid forholdsvis kurant å konvertere disse parameterverdiene til parametere for lokalisering, skala og form.

<sup>20</sup> Vi merker oss at prosessen består av en rekke steg, der hvert steg er avhengig av ulike forutsetninger. Den resulterende fordeling bør således tolkes med all mulig forsiktighet.

for sin initiale fordeling ble altså det andre møtet lagt opp som en iterativ prosess, der individuelle og gjennomsnittlige resultater ble presentert og diskutert blant deltagerne, og særlig kostnadskategorier der det var størst uenighet ble diskutert. Ekspertene ble så oppfordret til å revidere sine estimater. Prosessen med Monte Carlo simulering ble så gjentatt for å finne den reviderte sannsynlighetsfordelingen for totalkostnadsestimatet. Ekspertene gjennomgikk så kostnadsmodellen for å identifisere ulike positive og negative korrelasjoner mellom ulike kostnadselementer og mellom eksterne forhold og de ulike kostnadselementene. Korrelasjoner ble identifisert ved at ekspertene ble spurt hvor mye variansen i en gitt kostnadskategori A ville bli redusert som en konsekvens av økt kunnskap om den presise verdien for den korrelerte variabelen B. Endelig ble hele estimeringsprosessen og resultatene dokumentert for oppdragsgiver.

Dillon, et al. (2002) peker på at prosessen avdekket noen metodologiske spørsmål som utgjorde en særlig utfordring. Dette var i) korrelasjoner mellom kostnadskomponenter; ii) aggregering av ekspertvurderinger; iii) informasjonsdeling mellom eksperter og iv) hensiktsmessighetene av å diskontere kostnadsestimater. Diskonteringsspørsmålet er egentlig litt perifert i forhold til det vi fokuserer på her, så her henvises den interesserte leser til artikkelen. Men de øvrige temaene kan vi komme litt nærmere inn på.

Med hensyn til korrelasjoner mellom kostnadskomponenter peker forfatterne på at det på den ene side er slik at det i de aller fleste tilfellene er galt å anta at kostnadene er uavhengige (ukorrellert), på den annen side er det uklart hvor sensitivt totalkostnadsestimatet er med hensyn til estimatene på korrelasjoner mellom de ulike estimatene. Generelt vil positive korrelasjoner bidra til en bredere fordeling for totalkostnadene (økt usikkerhet). For å få et inntrykk av konsekvensene av ulike tilnærminger, ble spredningen i totalkostnadsestimatet sammenlignet med og uten de estimerte korrelasjoner, og basert på ulike tilnærminger til aggregering. Når det gjelder aggregering (pkt. ii) , anbefaler forfatterne at det først produseres en overordnet kostnadsfordeling basert på den enkelte eksperts delkostnadsestimat og at de ulike eksperters totalkostnadsestimat så aggregeres på tvers av ekspertene. Alternativet er at de ulike eksperters delkostnadsestimat aggregeres, og at det så etableres et totalkostnadsestimat basert på aggregering av disse. Bakgrunnen for denne anbefalingen er bl.a. at dette i størst mulig grad ivaretar integriteten i den tenking og det rasjonale som ligger bak den enkelte eksperts estimater. Eksempelvis vil enkelte eksperter være mer optimistiske enn andre, og aggregering på tvers av delkostnadsestimater vil maskere denne informasjonen.

Spørsmålet om interaksjon og kommunikasjonsmuligheter mellom eksperter var også et viktig tema (pkt. iii). Tema som var diskutert var i hvilken grad ekspertene skulle bli presentert for et base-case, som jo ville kunne utgjøre et forankringspunkt og dermed en årsak til estimatskjevhet. Videre var det et spørsmål om ekspertene skulle kunne kommunisere, skulle de kunne kommunisere ansikt til ansikt eller anonymt og skulle de presentere et basisestimat og så tillates å kommunisere osv. Forfatterne peker på at ved den tilnærming som ble valgt, ble det gjennom diskusjonen avdekket usikkerheter som enkelte eksperter hadde oversett. På den ene siden førte kommunikasjon til større grad av konsensus, likere individuelle fordelinger, og således også en smalere fordeling for totalkostnader. På den annen side førte fokuseringen på oversette usikkerhetsfaktorer til en større spredning i det endelige totalkostnadsestimat, og den siste effekten dominerte den første.

Forfatterne konkluderer med at prosessen avslørte stor uenighet mellom ekspertene, men at resultatet gjennom aggregering likevel ble et usikkerhetsspenn rundt kostnadsestimatene som gruppen aksepterte som representative for sin oppfatning.

## 5.8 Ekspertvurdering i tidligfasen

I engelskspråklig litteratur benyttes begrepet "elicitation" om prosessen med å avdekke ekspertkunnskap. "Elicit" betyr å lokke frem, og det er nettopp det som er av de store utfordringene. Prosesslederen står også overfor en stor utfordring etter "fremlokkingen". Det er å fremstille denne kvalitative kunnskapen på en kvantitativ form som gjør at kunnskapen kan benyttes i en beslutningssammenheng.

Men før man kommer til selve fremlokkingen, er det også en utfordring å identifisere et tilstrekkelig antall eksperter med kunnskap som representerer både dybde og bredde; bredde også i den forstand at de ulike ekspertene har ulik bakgrunn, slik at estimeringsproblemet betraktes fra ulike perspektiv. Den neste utfordringen er å skape en ramme rundt prosessen som er slik at eksperten får insentiv til å fremlegge kunnskapen sin på dette området på en måte som vedkommende er komfortabel med, og at dette skjer i en slik form at kunnskapen kan benyttes i det aktuelle evalueringsprosjektet. Dette er ingen liten utfordring, og det er mange feilkilder underveis som man må være oppmerksom på, ikke minst viser forskningen at mennesket, når det skal foreta vurderinger benytter ulike heuristikker som forankring, representativitet og tilgjengelighet. I disse heuristikkene ligger det også potensielle feilkilder, jf. diskusjonen i kapittel 9.

Det kan være nødvendig å benytte god tid og gå flere runder i denne delen av "fremlokkingsprosessen", ikke minst kan det være nødvendig å trene ekspertene opp i å arbeide på denne måten, og i unngå en del av de feilkildene som kan oppstå. Det kan også være nødvendig å gå flere runder, bl.a. for at ekspertene kan få innsyn i andres vurdering, slik at egne estimater kan korrigeres i lys av denne informasjonen.

Når kunnskapen er avdekket blir den neste utfordringen å behandle den på en slik måte at den kan benyttes i en beslutningssammenheng. Sentrale utfordringen er bl.a. å representere avdekket informasjon i en sannsynlighetsfordeling, samt å aggregere sannsynlighetsfordelinger. Her er det en rekke mulige feilkilder, ikke minst med hensyn til valg av fordeling og med hensyn til hvordan aggregeringen skjer. Kunnskap om bayesiansk metodikk vil likevel være viktig ved anvendelse av denne metodikken, og de siste årenes utvikling i datateknologi har muliggjort mer sofistikerte og, ikke minst, mer korrekte tilnærminger som ikke var mulig for 10-15 år siden. En alternativ lovende tilnærming til å avdekke, representere og analysere ekspertkunnskap er fuzzy logikk.

Gitt disse utfordringene, kan man spørre seg om det har noen hensikt å benytte ekspertvurderinger, ikke minst i en tidligfase. Det er da viktig å huske på at en viktig målsetting med bruk av ekspertvurdering er å fange opp hovedtrekkene i ekspertens eller ekspertenes oppfatning. Detaljene, for eksempel med hensyn til den nøyaktige formen på fordelingen, vil ikke spille så stor rolle i forhold til den beslutningen som skal fattes, avveining mellom alternativer og disses respektive forventede nytte.

Denne gjennomgangen har imidlertid også vist at det til tross for de metodologiske utfordringene, ikke minst gjennom det avsluttende praktiske eksempelet, så bør ekspertvurderinger kunne ha en viktig potensiell anvendelse ved tidligfasevurdering for store offentlige prosjekter i Norge. I England har det ved University of Sheffield blitt satt opp en forskningsgruppe som skal drive forskning og utprøving av metodikk for ekspertvurderinger;

med benevnelsen ”Bayesian Elicitation of Experts' Probabilities”; BEEP.<sup>21</sup> Denne gruppen, som også er en del av et internasjonalt nettverk for bruk av ekspertvurderinger, vil særlig fokusere på anvendelser innenfor helse. I Norge er det ingen tilsvarende forskningsaktiviteter eller miljøer på området.<sup>22</sup> Kanskje tiden er moden for et program for bruk av ekspertvurderinger ved store offentlige investeringsprosjekter?

---

<sup>21</sup> Se <http://www.shef.ac.uk/beep/index.html>.

<sup>22</sup> Ved Simula er det riktig nok blitt arbeidet en del med bruk av ekspertvurdering i forhold til store software utviklingsprosjekter.

## 5.9 Referanser

- Bazerman, M. H. (2001): *Judgment in managerial decision making*. New York: John Wiley & Sons.
- Berger, J. O. (1985): *Statistical Decision Theory and Bayesian Analysis*.: Springer.
- Chatterjee, S., og S. Chatterjee (1987): "On combining expert opinion," *American journal of mathematical and management science*, **7**(3&4): 271-295.
- Dillon, R. L., R. John, og D. v. Wintrefeldt (2002): "Assessment of Cost Uncertainties for Large Technology Projects: A Methodology and an Application," *Interfaces*, **32**(4): 52-66.
- Fischhoff, B. (1982): "Debiasing," i *Judgment under uncertainty: Heuristics and biases*, red. D. Kahneman, P. Slovic, og A. Tversky. Cambridge: Cambridge University Press.
- Gilowitch, T., D. Griffin, og D. Kahneman (2002): *Heuristics and biases - The psychology of Intuitive Judgment*. Cambridge: Cambridge University Press.
- Gosling, J. P. (2005): "Elicitation: A Nonparametric View," Sheffield: University of Sheffield.
- Jenkinson, D. (2005): *The Elicitation of Probabilities-A Review of the Statistical Literature*. The University of Sheffield; BEEP working paper
- Jørgensen, M. (2002): "A Review of Studies on Expert Estimation of Software Development Effort," *The Journal of Systems and Software*, **70**(1-2): 37-.
- Kahneman, D., P. Slovic, og A. Tversky (1982): "Judgment under uncertainty: Heuristics and biases," Cambridge: Cambridge Univ. Press.
- (1982): *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Kahneman, D., og A. Tversky (1982): "Subjective probability: A judgment of representativeness," i *Judgment under uncertainty: Heuristics and biases*, red. P. S. a. A. T. Daniel Kahneman. New York: Cambridge Univ. Press, 32-47.
- Kaminaris, S. D., T. D. Tsoutsos, D. Agoris, og A. V. Machias (2006): "Assessing renewables-to-electricity systems: a fuzzy expert system model," *Energy Policy*, **1357-1366**.
- Mason, A. K., og D. J. K. Sr. (1997): "Estimating costs with fuzzy logic," *AACE International Transactions*, **1997**: 122-127.
- Meyer, M. A., og J. A. Booker (2001): *Eliciting and Analyzing Expert Judgment - A Practical Guide*.: ASA-SIAM.

- Morgan, M. G., og M. Henrion (1990): *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge: Cambridge University Press.
- Morgan, M. G., og D. W. Keith (1995): "Subjective judgments by climate experts," *Environmental Science & Technology*, **29**(10).
- O'Hagan, A. (1998): "Eliciting Expert Beliefs in Substantial Practical Applications," *The Statistician*, **47**(1): 21-35.
- Tversky, A. (1974): "Assessing Uncertainty," *Journal of the Royal Statistical Society, Series B (Methodological)*, **36**(2): 148-159.
- Øien, K., O. J. Klakegg, P. R. Hokstad, og R. Rosness (1996): *Inngangsdata til LCP og bruk av ekspertvurderinger - Håndbok for gjennomføring av ekspertvurderinger*. NTNU: 96004.

## **6 ELIMINATING BIAS IN EARLY PROJECT DEVELOPMENT THROUGH REFERENCE CLASS FORECASTING AND GOOD GOVERNANCE**

BENT FLYVBJERG\*

### **6.1 Introduction**

Chapter 3 identified as a main problem in large-scale investment projects the existence of optimism bias and strategic misrepresentation with project promoters. A consequence of such bias is a high incidence of cost overruns and benefit shortfalls in projects. This chapter develops a number of measures aimed at eliminating, or at least reducing, optimism bias and strategic misrepresentation in project development. The measures include changed governance structures and better planning methods. The aim is to ensure that decisions on whether to build projects or not are based on valid information about costs and benefits, instead of being based on misinformation as is often the case today.

This is not to say that costs and benefits are or should be the only basis for deciding whether to build large projects. Clearly, forms of rationality other than economic rationality are at work in most projects and are balanced in the broader frame of public deliberation and decision making. But the costs and benefits of large-scale projects often run in the hundreds of millions of dollars, with risks correspondingly high. Without knowledge of such risks, decisions are likely to be flawed.

When contemplating what planners can do to improve decision making, we need to distinguish between two fundamentally different situations: (1) planners and promoters consider it important to get forecasts of costs, benefits, and risks right, and (2) planners and promoters do not consider it important to get forecasts right, because optimistic forecasts are seen as a necessary means to getting projects started. The first situation is the easier one to deal with and here better methodology will go a long way in improving planning and decision making. The second situation is more difficult. Here changed incentives are essential in order to reward honesty and punish deception, where today's incentives often do the exact opposite.

Thus two main measures of reform will be considered below: (1) better forecasting methods, and (2) improved incentive structures, with the latter being more important.

Like in the previous paper, the emphasis will be on transportation infrastructure projects, because the best data are available for these. But comparative research again shows that the issues raised for transportation apply to a wide range of other project types including power

---

\* Bent Flyvbjerg er professor ved Aalborg Universitet, Department of Development and Planning (bentflyvbjerg@stofanet.dk), og ved Delft University of Technology, Nederland.



plants, dams, water projects, concert halls, museums, sports arenas, convention centers, IT systems, oil and gas extraction projects, aerospace projects, and weapons systems (Flyvbjerg, Bruzelius, and Rothengatter, 2003: 18-19; Flyvbjerg, Holm, and Buhl, 2002: 286; Flyvbjerg, 2005; Altshuler and Luberoff, 2003).

## 6.2 Better Methods: Reference Class Forecasting

If planners and promoters genuinely consider it important to get forecasts right, this may be achieved by using a new forecasting method called "reference class forecasting" to reduce inaccuracy and bias. Reference class forecasting is based on theories of decision making under uncertainty that won Princeton psychologist Daniel Kahneman the Nobel prize in economics in 2002 (Kahneman and Tversky 1979a, b; Kahneman 1994). Reference class forecasting promises more accuracy in forecasts by taking a so-called "outside view" on prospects being forecasted, while conventional forecasting takes an inside view. The outside view on a given project is based on knowledge about actual performance in a reference class of comparable projects.

Reference class forecasting bypasses human bias--including optimism bias and strategic misrepresentation--by cutting directly to outcomes. In experimental research carried out by Daniel Kahneman and others, this method has been demonstrated to be more accurate than conventional forecasting methods (Kahneman and Tversky 1979a, 1979b; Kahneman 1994; Lovallo and Kahneman 2003).

In April 2005, based on a study by Flyvbjerg, Holm, and Buhl (2005), the American Planning Association (2005) officially endorsed reference class forecasting:

"APA encourages planners to use reference class forecasting in addition to traditional methods as a way to improve accuracy. The reference class forecasting method is beneficial for non-routine projects such as stadiums, museums, exhibit centers, and other local one-off projects. Planners should never rely solely on civil engineering technology as a way to generate project forecasts."

Below, first the theoretical and methodological foundations for reference class forecasting are explained, then the first instance of reference class forecasting in practice is presented.

## 6.3 The Planning Fallacy and the Outside View

The theoretical and methodological foundations of reference class forecasting were first described by Kahneman and Tversky (1979b) and later by Lovallo and Kahneman (2003). Reference class forecasting was originally developed to compensate for the type of cognitive bias identified in pioneering work on planning and decision making carried out by Kahneman and Tversky (1979a) and Kahneman 1994. Kahneman and Tversky found that errors of judgment are often systematic and predictable rather than random, manifesting bias rather than confusion, and that any corrective prescription should reflect this. They also found that many errors of judgment are shared by experts and laypeople alike.

Finally Kahneman and Tversky found that errors remain compelling even when one is fully aware of their nature. Thus awareness of a perceptual or cognitive illusion does not by itself

produce a more accurate perception of reality, according to Kahneman and Tversky (1979b: 314). Awareness may, however, enable one to identify situations in which the normal faith in one's impressions must be suspended and in which judgment should be controlled by a more critical evaluation of the evidence. Reference class forecasting is a method for such critical evaluation. Human judgment, including forecasts, are biased. Reference class forecasting is a method for unbiasing forecasts.

Kahneman and Tversky (1979a, b) argue that human judgment is generally optimistic due to overconfidence and insufficient regard to distributional information. Thus people will underestimate the costs, completion times, and risks of planned actions, whereas they will overestimate the benefits of the same actions. Lovallo and Kahneman (2003: 58) call such common behavior the "planning fallacy" and they argue that it stems from actors taking an "inside view" focusing on the constituents of the specific planned action rather than on the outcomes of similar actions that have already been completed.

Kahneman and Tversky (1979b) further argue that the prevalent tendency to underweigh or ignore distributional information is perhaps the major source of error in forecasting. "The analysts should therefore make every effort to frame the forecasting problem so as to facilitate utilizing all the distributional information that is available," say Kahneman and Tversky (1979b: 316). This may be considered the single most important piece of advice regarding how to increase accuracy in forecasting through improved methods. Using such distributional information from other ventures similar to that being forecasted is called taking an "outside view" and it is the cure to the planning fallacy. Reference class forecasting is a method for systematically taking an outside view on planned actions.

More specifically, reference class forecasting for a particular project requires the following three steps:

- (1) Identification of a relevant reference class of past, similar projects. The class must be broad enough to be statistically meaningful but narrow enough to be truly comparable with the specific project.
- (2) Establishing a probability distribution for the selected reference class. This requires access to credible, empirical data for a sufficient number of projects within the reference class to make statistically meaningful conclusions.
- (3) Comparing the specific project with the reference class distribution, in order to establish the most likely outcome for the specific project.

Thus reference class forecasting does not try to forecast the specific uncertain events that will affect the particular project, but instead places the project in a statistical distribution of outcomes from the class of reference projects. In statisticians vernacular, reference class forecasting consists of regressing forecasters' best guess toward the average of the reference class and expanding their estimate of credible interval toward the corresponding interval for the class (Kahneman and Tversky 1979b: 326).

Daniel Kahneman relates the following story about curriculum planning to illustrate how reference class forecasting works (Lovallo and Kahneman 2003: 61). Some years ago, Kahneman was involved in a project to develop a curriculum for a new subject area for high schools in Israel. The project was carried out by a team of academics and teachers. In time, the team began to discuss how long the project would take to complete. Everyone on the team

was asked to write on a slip of paper the number of months needed to finish and report the project. The estimates ranged from 18 to 30 months.

One of the team members--a distinguished expert in curriculum development--was then posed a challenge by another team member to recall as many projects similar to theirs as possible and to think of these projects as they were in a stage comparable to their project. "How long did it take them at that point to reach completion?", the expert was asked. After a while he answered, with some discomfort, that not all the comparable teams he could think of ever did complete their task. About 40 percent of them eventually gave up. Of those remaining, the expert could not think of any that completed their task in less than seven years, nor of any that took more than ten. The expert was next asked if he had reason to believe that the present team was more skilled in curriculum development than the earlier ones had been. The expert said no, he did not see any relevant factor that distinguished this team favorably from the teams he had been thinking about. His impression was that the present team was slightly below average in terms of resources and potential.

The wise decision at this point would probably have been for the team to break up, according to Kahneman. Instead, the members ignored the pessimistic information and proceeded with the project. They finally completed the project eight years later, and their efforts went largely wasted--the resulting curriculum was rarely used.

In this example, the curriculum expert made two forecasts for the same problem and arrived at very different answers. The first forecast was the inside view; the second was the outside view, or the reference class forecast. The inside view is the one that the expert and the other team members adopted. They made forecasts by focusing tightly on the project at hand, considering its objective, the resources they brought to it, and the obstacles to its completion. They constructed in their minds scenarios of their coming progress and extrapolated current trends into the future. The resulting forecasts, even the most conservative ones, were overly optimistic. The outside view is the one provoked by the question to the curriculum expert. It completely ignored the details of the project at hand, and it involved no attempt at forecasting the events that would influence the project's future course. Instead, it examined the experiences of a class of similar projects, laid out a rough distribution of outcomes for this reference class, and then positioned the current project in that distribution. The resulting forecast, as it turned out, was much more accurate.

The contrast between inside and outside views has been confirmed by systematic research (Gilovich, Griffin, and Kahneman, 2002). The research shows that when people are asked simple questions requiring them to take an outside view, their forecasts become significantly more accurate. For example, a group of students enrolling at a college were asked to rate their future academic performance relative to their peers in their major. On average, these students expected to perform better than 84% of their peers, which is logically impossible. The forecasts were biased by overconfidence. Another group of incoming students from the same major were asked about their entrance scores and their peers' scores before being asked about their expected performance. This simple diversion into relevant outside-view information, which both groups of subjects were aware of, reduced the second group's average expected performance ratings by 20%. That is still overconfident, but it is much more realistic than the forecast made by the first group (Lovallo and Kahneman 2003: 61).

However, most individuals and organizations are inclined to adopt the inside view in planning new projects. This is the conventional and intuitive approach. The traditional way to think

about a complex project is to focus on the project itself and its details, to bring to bear what one knows about it, paying special attention to its unique or unusual features, trying to predict the events that will influence its future. The thought of going out and gathering simple statistics about related projects seldom enters a manager's mind. This is the case in general, according to Lovallo and Kahneman (2003: 61-62). And it is certainly the case for cost and demand forecasting in transportation infrastructure projects. Of the several hundred forecasts reviewed in Flyvbjerg, Bruzelius, and Rothengatter (2003) and Flyvbjerg, Holm, and Buhl (2002, 2005), not one was a reference class forecast.<sup>23</sup>

While understandable, project managers' preference for the inside view over the outside view is unfortunate. When both forecasting methods are applied with equal skill, the outside view is much more likely to produce a realistic estimate. That is because it bypasses cognitive and political biases such as optimism bias and strategic misrepresentation and cuts directly to outcomes. In the outside view project managers and forecasters are not required to make scenarios, imagine events, or gauge their own and others' levels of ability and control, so they cannot get all these things wrong. Human bias is bypassed. Surely the outside view, being based on historical precedent, may fail to predict extreme outcomes, that is, those that lie outside all historical precedents. But for most projects, the outside view will produce more accurate results. In contrast, a focus on inside details is the road to inaccuracy.

The comparative advantage of the outside view is most pronounced for non-routine projects, understood as projects that managers and decision makers in a certain locale or organization have never attempted before--like building new plants or infrastructure or catering to new types of demand. It is in the planning of such new efforts that the biases toward optimism and strategic misrepresentation are likely to be largest. To be sure, choosing the right reference class of comparative past projects becomes more difficult when managers are forecasting initiatives for which precedents are not easily found, for instance the introduction of new and unfamiliar technologies. However, most projects are both non-routine locally and use well-known technologies. Such projects are, therefore, particularly likely to benefit from the outside view and reference class forecasting.

## 6.4 First Instance of Reference Class Forecasting in Practice

The first instance of reference class forecasting in practice may be found in Flyvbjerg and Cowi (2004): *Procedures for Dealing with Optimism Bias in Transport Planning*.<sup>24</sup> Based on this study, in the Summer of 2004 the UK Department for Transport and HM Treasury decided to employ the method as part of project appraisal for large transportation projects under their jurisdiction.

---

<sup>23</sup> The closest thing to an outside view in large infrastructure forecasting is Gordon and Wilson's (1984) use of regression analysis on an international cross section of light-rail projects to forecast patronage in a number of light-rail schemes in North America.

<sup>24</sup> The fact that this is, indeed, the first instance of practical reference class forecasting has been confirmed with Daniel Kahneman and Dan Lovallo, who also knows of no other instances of practical reference class forecasting (personal communications with Daniel Kahneman and Dan Lovallo, author's archives).

The immediate background to this decision was the revision to "The Green Book" by HM Treasury in 2003, that identified for large public procurement a demonstrated, systematic tendency for project appraisers to be overly optimistic:

"There is a demonstrated, systematic, tendency for project appraisers to be overly optimistic. To redress this tendency appraisers should make explicit, empirically based adjustments to the estimates of a project's costs, benefits, and duration ... [I]t is recommended that these adjustments be based on data from past projects or similar projects elsewhere" (HM Treasury 2003b: 1).

Such optimism was seen as an impediment to prudent fiscal planning, for the government as a whole and for individual departments within government. To redress this tendency HM Treasury recommended that appraisers involved in large public procurement should make explicit, empirically based adjustments to the estimates of a project's costs, benefits, and duration. HM Treasury recommended that these adjustments be based on data from past projects or similar projects elsewhere, and adjusted for the unique characteristics of the project at hand. In the absence of a more specific evidence base, HM Treasury encouraged government departments to collect valid and reliable data to inform future estimates of optimism, and in the meantime use the best available data. The Treasury let it be understood that in future the allocation of funds for large public procurement would be dependent on valid adjustments of optimism in order to secure valid estimates of costs, benefits, and duration of large public procurement (HM Treasury 2003a, b).

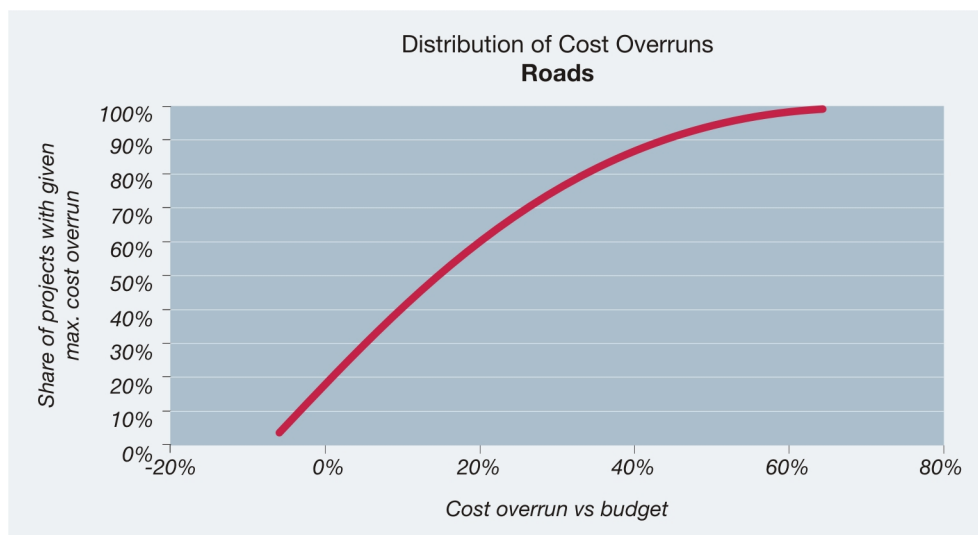
In response to the Treasury's Green Book and its recommendations, the UK Department for Transport decided to collect the type of data, which the Treasury recommended, and on that basis to develop a methodology for dealing with optimism bias in the planning and management of transportation projects. The Department for Transport appointed Bent Flyvbjerg in association with Cowi to undertake this assignment as regards costing of large transportation procurement. The main aims of the assignment were two; first, to provide empirically based optimism bias uplifts for selected reference classes of transportation infrastructure projects, and, second, to provide guidance on using the established uplifts to produce more realistic forecasts of capital expenditures in individual projects (Flyvbjerg and Cowi 2004). Uplifts would be established for capital expenditures based on the full business case (time of decision to build).

The types of transportation schemes under the direct and indirect responsibility of the UK Department for Transport were divided into a number of distinct categories where statistical tests, benchmarkings, and other analyses showed that the risk of cost overruns within each category may be treated as statistically similar. For each category a reference class of projects was then established as the basis for reference class forecasting, as required by step 1 in the 3-step procedure for reference class forecasting described above. The specific categories and the types of project allocated to each category are shown in Table 6-1.

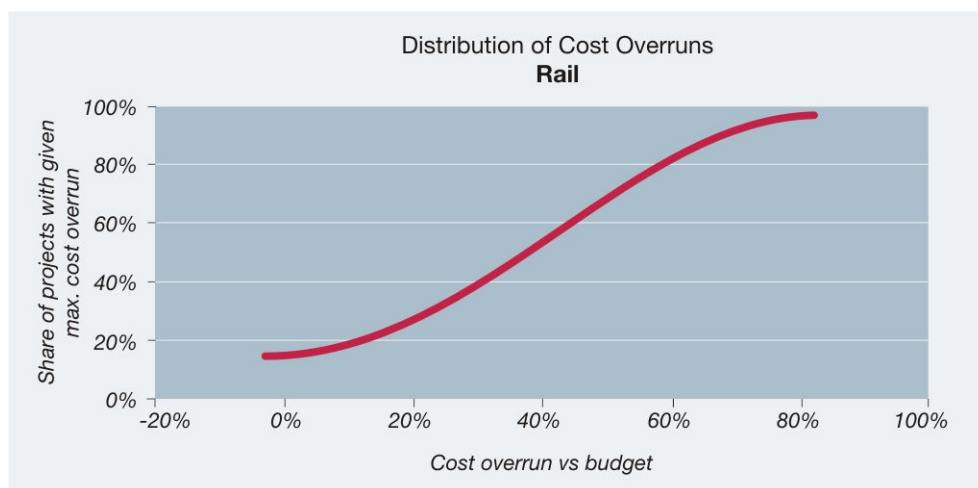
Table 6-1 Categories and types of projects used as basis for reference class forecasting

Category	Types of projects
Roads	Motorway Trunk roads Local roads Bicycle facilities Pedestrian facilities Park and ride Bus lane schemes Guided buses on wheels
Rail	Metro Light rail Guided buses on tracks Conventional rail High speed rail
Fixed links	Bridges Tunnels
Building projects	Stations Terminal buildings
IT projects	IT system development
Standard civil engineering	Included for reference purposes only
Non-standard civil engineering	Included for reference purposes only

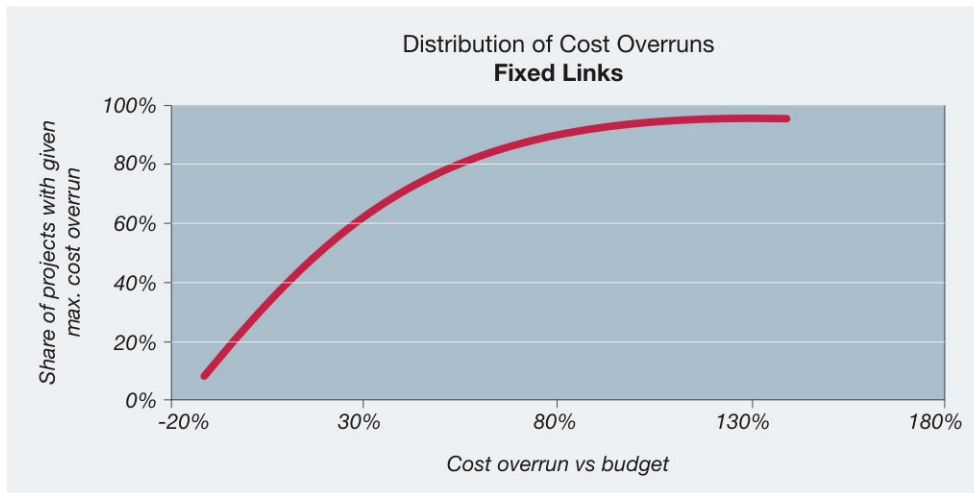
For each category of projects, a reference class of completed, comparable transportation infrastructure projects was used to establish probability distributions for cost overruns for new projects similar in scope and risks to the projects in the reference class, as required by step 2 in reference class forecasting. For roads, for example, a class of 172 completed and comparable projects was used to establish the probability distribution of cost overruns shown in Figur 6-1. The share of projects with a given maximum cost overrun is shown in the figure. For instance, 40% of projects have a maximum cost overrun of 10%; 80% of projects a maximum overrun of 32%, etc. For rail, the probability distribution is shown in Figur 6-2, and for bridges and tunnels in Figur 6-3. The figures show that the risk of cost overrun is substantial for all three project types, but highest for rail, followed by bridges and tunnels, and with the lowest risk for roads.



Figur 6-1 Probability distribution of cost overrun for roads, constant prices (N=172). Source: Flyvbjerg database on large-scale infrastructure projects.



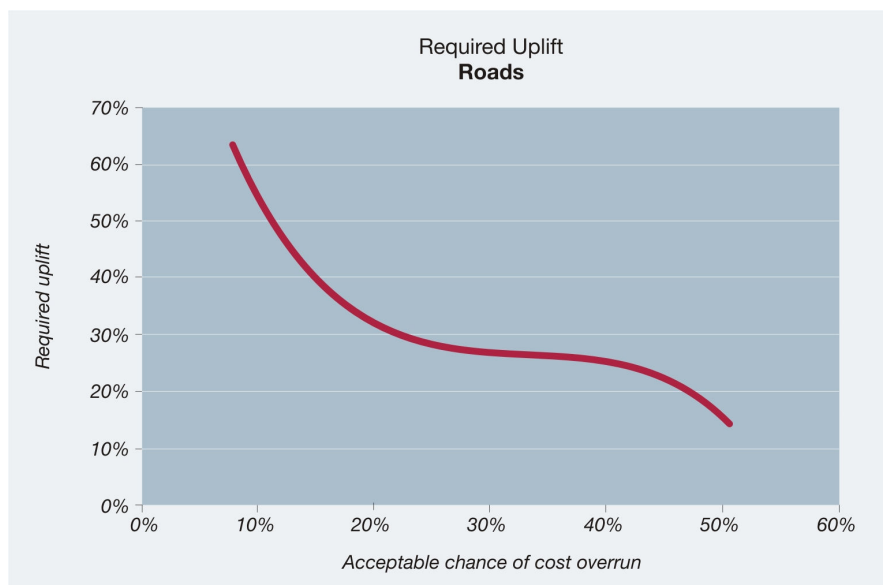
Figur 6-2 Probability distribution of cost overrun for rail, constant prices (N=46). Source: Flyvbjerg database on large-scale infrastructure projects.



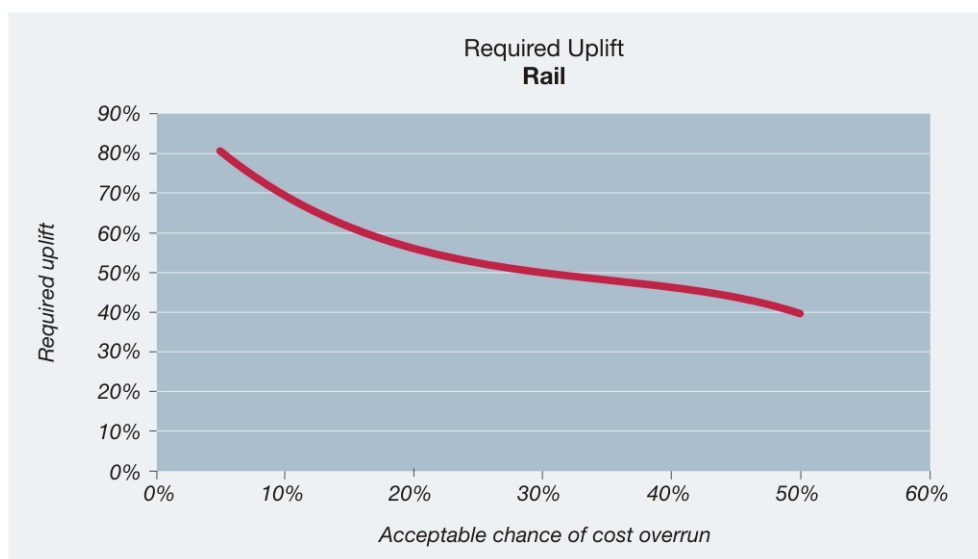
*Figur 6-3* Probability distribution of cost overrun for fixed links, constant prices ( $N=34$ ).  
 Source: Flyvbjerg database on large-scale infrastructure projects.

Based on the probability distributions described above the required uplifts needed to carry out step 3 in a reference class forecast may be calculated as shown in Figures 6-4 to 6-6. The uplifts refer to cost overrun calculated in constant prices. The lower the acceptable risk for cost overrun, the higher the uplift. For instance, with a willingness to accept a 50% risk for cost overrun in a road project, the required uplift for this project would be 15%. If the Department for Transport were willing to accept only a 10% risk for cost overrun, then the required uplift would be 45%. In comparison, for rail with a willingness to accept a 50% risk for cost overrun, the required uplift would be 40%. If the Department for Transport were willing to accept only a 10% risk for cost overrun, then the required uplift would be 68% for rail. All three figures share the same basic S-shape, but at different levels, demonstrating that the required uplifts are significantly different for different project categories for a given level of risk of cost overrun. The figures also show that the cost for additional reductions in the risk of cost overrun is different for the three types of projects, with risk reduction becoming increasingly expensive (rising marginal costs) for roads and fixed links below 20% risk, whereas for rail the cost of increased risk reduction rises more slowly, albeit from a high level.

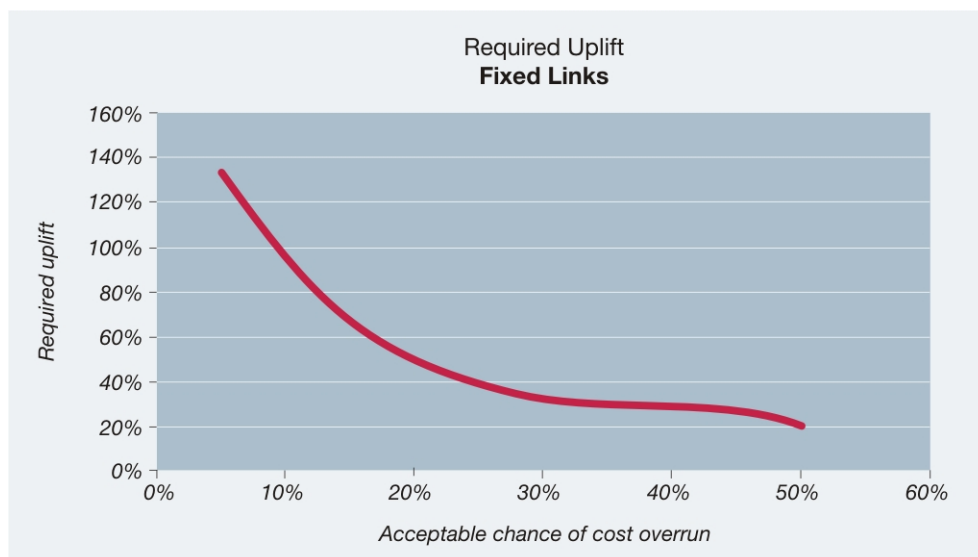




Figur 6-4 Required uplift for roads as function of the maximum acceptable level of risk for cost overrun, constant prices (N=172). Source: Flyvbjerg database on large-scale infrastructure projects.



Figur 6-5 Required uplift for rail as function of the maximum acceptable level of risk for cost overrun, constant prices (N=46). Source: Flyvbjerg database on large-scale infrastructure projects.



*Figur 6-6 Required uplift for fixed links as function of the maximum acceptable level of risk for cost overrun, constant prices (N=34). Source: Flyvbjerg database on large-scale infrastructure projects.*

Table 6-2 presents an overview of applicable optimism bias uplifts for the 50% and 80% percentiles for all the project categories listed in Table 6-1. The 50% percentile is pertinent to the investor with a large project portfolio, where cost overruns on one project may be offset by cost savings on another. The 80% percentile--corresponding to a risk of cost overrun of 20%--is the level of risk that the UK Department for Transport is typically willing to accept for large investments in local transportation infrastructure.

Table 6-2 *Applicable capital expenditure optimism bias uplifts for 50% and 80% percentiles, constant prices.*

Category	Types of projects	Applicable optimism bias uplifts	
		50% percentile	80% percentile
Roads	Motorway Trunk roads Local roads Bicycle facilities Pedestrian facilities Park and ride Bus lane schemes Guided buses on wheels	15%	32%
Rail	Metro Light rail Guided buses on tracks Conventional rail High speed rail	40%	57%
Fixed links	Bridges Tunnels	23%	55%
Building projects	Stations Terminal buildings	4-51%*	
IT projects	IT system development	10-200%*	
Standard civil engineering	Included for reference purposes only	3-44%*	
Non-standard civil engineering	Included for reference purposes only	6-66%*	

\*) Based on Mott MacDonald (2002: 32); no probability distribution available.

The established uplifts for optimism bias should be applied to estimated budgets at the time of decision to build a project. In the UK, the approval stage for a large transportation project is equivalent to the time of presenting the business case for the project to the Department for Transport with a view to obtaining the go or no-go for that project.

If, for instance, a group of project managers were preparing the business case for a new motorway, and if they or their client had decided that the risk of cost overrun must be less than 20%, then they would use an uplift of 32% on their estimated capital expenditure budget.

Thus, if the initially estimated budget were £100 million, then the final budget--taking into account optimism bias at the 80%-level--would be £132 million (£1 = \$1.8). If the project managers or their client decided instead that a 50% risk of cost overrun was acceptable, then the uplift would be 15% and the final budget £115 million.

Similarly, if a group of project managers were preparing the business case for a metro rail project, and if they or their client had decided that with 80% certainty they wanted to stay within budget, then they would use an uplift on capital costs of 57%. An initial capital expenditure budget of £300 million would then become a final budget of £504 million. If the project managers or their client required only 50% certainty they would stay within budget, then the final budget would be £420 million.

It follows that the 50% percentile should be used only in instances where investors are willing to take a high degree of risk that cost overrun will occur and/or in situations where investors are funding a large number of projects and where cost savings (underruns) on one project may be used to cover the costs of overruns on other projects. The upper percentiles (80-90%) should be used when investors want a high degree of certainty that cost overrun will not occur, for instance in stand-alone projects with no access to additional funds beyond the approved budget. Other percentiles may be employed to reflect other degrees of willingness to accept risk and the associated uplifts as shown in Figures 6-4 to 6-6.

Only if project managers have evidence to substantiate that they would be significantly better at estimating costs for the project at hand than their colleagues were for the projects in the reference class would the managers be justified in using lower uplifts than those described above. Conversely, if there is evidence that the project managers are worse at estimating costs than their colleagues, then higher uplifts should be used.

The methodology described above for systematic, practical reference class forecasting for transportation projects was developed 2003-2004 with publication by the Department of Transport in August 2004. From this date local authorities applying for funding for transportation projects with the Department for Transport or with HM Treasury were required to take into account optimism bias by using uplifts as described above and as laid out in more detail in guidelines from the two ministries.

## **6.5 Forecasting Costs for the Edinburgh Tram**

In October 2004, the first instance of practical use of the uplifts was recorded, in the planning of the Edinburgh Tram Line 2. Ove Arup and Partners Scotland (2004) had been appointed by the Scottish Parliament's Edinburgh Tram Bill Committee to provide a review of the Edinburgh Tram Line 2 business case developed on behalf of Transport Initiatives Edinburgh. Transport Initiatives Edinburgh is project promoter and is a private limited company owned by the City of Edinburgh Council established to deliver major transport projects for the Council. The Scottish Executive is a main funder of the Edinburgh Tram, having made an Executive Grant of £375 million (\$670 million) towards lines 1 and 2 of which Transport Initiatives Edinburgh proposed spending £165 million towards Line 2.

As part of their review, Ove Arup assessed whether the business case for Tram Line 2 had adequately taken into account optimism bias as regards capital costs. The business case had estimated a base cost of £255 million and an additional allowance for contingency and

optimism bias of £64 million--or 25%--resulting in total capital costs of app. £320 million. Ove Arup concluded about this overall estimate of capital costs that it seemed to have been rigorously prepared using a database of costs, comparison to other UK light rail schemes, and reconciliations with earlier project estimates. Ove Arup found, however, that the following potential additional costs needed to be considered in determining the overall capital costs: £26 million for future expenditure on replacement and renewals and £20 million as a notional allowance for a capital sum to cover risks of future revenue shortfalls, amounting to an increase in total capital costs of 14.4% (Ove Arup and Partners Scotland 2004: 15-16)

Using the UK Department for Transport uplifts for optimism bias presented above on the base costs, Ove Arup then calculated the 80<sup>th</sup> percentile value for total capital costs--the value at which the likelihood of staying within budget is 80%--to be £400 million (i.e., £255 million x 1.57). The 50<sup>th</sup> percentile for total capital costs--the value at which the likelihood of staying within budget is 50%--was £357 million (i.e., £255 x 1.4). Ove Arup remarked that these estimates of total capital costs were likely to be conservative, that is, low, because the UK Department for Transport recommends that its optimism bias uplifts be applied to the budget at the time of decision to build, which typically equates to business case submission, and Tram Line 2 had not yet reached even the outline business case stage, indicating that risks would be substantially higher at this early stage as would corresponding uplifts. On that basis Arup concluded that "it is considered that current optimism bias uplifts [for Tram Line 2] may have been underestimated" (Ove Arup and Partners Scotland 2004: 27).

Finally, Ove Arup mentioned that the Department for Transport guidance does allow for optimism bias to be adjusted downward if strong evidence of improved risk mitigation can be demonstrated. According to Ove Arup, this may be the case if advanced risk analysis has been applied, but this was not the case for Tram Line 2. Ove Arup therefore concluded that "the justification for reduced Department for Transport optimism bias uplifts would appear to be weak" (Ove Arup and Partners Scotland 2004: 27-28). Thus the overall conclusion of Ove Arup was that the promoter's capital cost estimate of app. £320 million was optimistic. Most likely Tram Line 2 would cost significantly more.

By framing the forecasting problem to allow the use of the empirical distributional information made available by the UK Department for Transport, Ove Arup was able to take an outside view on the Edinburgh Tram Line 2 capital cost forecast and thus debias what appeared to be a biased forecast. As a result Ove Arup's client, The Scottish Parliament, was provided with a more reliable estimate of what the true costs of Line 2 was likely to be.

## 6.6 The "Dark Side" of Project Development

The present section considers the situation where planners and promoters do not find it important to get forecasts right and where planners, therefore, do not help to clarify and mitigate risk but, instead, generate and exacerbate it. Here planners are part of the problem, not the solution.

This situation may sound counterintuitive to some. After all, it may be agreed that planners ought to be interested in being accurate and unbiased in forecasting as part of their normal professional conduct. This is often even stated as an explicit requirement in the code of ethics of planners, for instance the American Institute of Certified Planners' (AICP) Code of Ethics and Professional Conduct, which states that "A planner must strive to provide full, clear and

accurate information on planning issues to citizens and governmental decision-makers" (American Planning Association, 1991: A.3). The British Royal Town Planning Institute (RTPI) has laid down similar obligations for its members (Royal Town Planning Institute, 2001).

However, the literature is replete with things planners and planning "must" strive to do, but which they don't. Planning must be open and communicative, but often it is closed. Planning must be participatory and democratic, but often it is an instrument of domination and control. Planning must be about rationality, but often it is about power (Flyvbjerg, 1998; Watson, 2003). This is the "dark side" of planning and planners identified by Flyvbjerg (1996) and Yiftachel (1998), which is remarkably underexplored by planning researchers and theorists.

Forecasting, too, has its dark side. It is here that "planners lie with numbers," as Wachs (1989) has aptly put it. Planners on the dark side are busy not with getting forecasts right and following the AICP Code of Ethics but with getting projects funded and built. And accurate forecasts are often not an effective means for achieving this objective. Indeed, accurate forecasts may be counterproductive, whereas biased forecasts may be effective in competing for funds and securing the go-ahead for construction. "The most effective planner," says Wachs (1989: 477), "is sometimes the one who can cloak advocacy in the guise of scientific or technical rationality."

Such advocacy would stand in direct opposition to AICP's ruling that "the planner's primary obligation [is] to the public interest" (American Planning Association, 1991: B.2). Nevertheless, seemingly rational forecasts that underestimate costs and overestimate benefits have long been an established formula for project approval as shown in Chapter XXX. Forecasting is here mainly another kind of rent-seeking behavior, resulting in a make-believe world of misrepresentation which makes it extremely difficult to decide which projects deserve undertaking and which do not. The consequence is, as even one of the industry's own organs, the Oxford-based Major Projects Association, acknowledges, that too many projects proceed that should not. We would like to add that many projects don't proceed that probably should, had they not lost out to projects with "better" misrepresentation (Flyvbjerg, Holm, and Buhl, 2002).

In this situation, the question is not so much what planners can do to reduce inaccuracy and risk in forecasting, but what others can do to impose on planners the checks and balances that would give planners the incentive to stop producing biased forecasts and begin to work according to their Code of Ethics. The challenge is to change the power relations that govern forecasting and project development. Better forecasting techniques and appeals to ethics won't do here; institutional change with a focus on transparency and accountability is necessary.

## 6.7 Good Governance in Major Projects

As argued in Flyvbjerg, Bruzelius, and Rothengatter (2003), two basic types of accountability define liberal democracies: (1) public sector accountability through transparency and public control, and (2) private sector accountability via competition and market control. Both types of accountability may be effective tools to curb planners' misrepresentation in forecasting and to promote a culture which acknowledges and deals effectively with risk. In order to achieve accountability through *transparency and public control*, the following would be required as

practices embedded in the relevant institutions (the full argument for the measures may be found in Flyvbjerg, Bruzelius, and Rothengatter, 2003, chapters 9-11):

- National-level government should not offer discretionary grants to local infrastructure agencies for the sole purpose of building a specific type of infrastructure. Such grants create perverse incentives. Instead, national government should simply offer "infrastructure grants" or "transportation grants" to local governments, and let local political officials spend the funds however they choose to, but make sure that every dollar they spend on one type of infrastructure reduces their ability to fund another. This prevents the situation where "free money" seems to be available, seen from the promoter's perspective, a situation likely to result in waste. Instead a situation is created where the promoter's own money is at stake, which should always be the case.
- Forecasts should be made subject to independent peer review. Where large amounts of taxpayers' money are at stake, such review may be carried out by national or state accounting and auditing offices, like the General Accounting Office in the US or the National Audit Office in the UK, who have the independence and expertise to produce such reviews. Other types of independent review bodies may be established, for instance within national departments of finance or with relevant professional bodies.
- Forecasts should be benchmarked against comparable forecasts, for instance using reference class forecasting as described in the previous section.
- Forecasts, peer reviews, and benchmarkings should be made available to the public as they are produced, including all relevant documentation.
- Public hearings, citizen juries, and the like should be organized to allow stakeholders and civil society to voice criticism and support of forecasts. Knowledge generated in this way should be integrated in planning and decision making.
- Scientific and professional conferences should be organized where forecasters would present and defend their forecasts in the face of colleagues' scrutiny and criticism.
- Projects with inflated benefit-cost ratios should be reconsidered and stopped if recalculated costs and benefits do not warrant implementation. Projects with realistic estimates of benefits and costs should be rewarded.
- Professional and occasionally even criminal penalties should be enforced for planners and forecasters who consistently and foreseeably produce deceptive forecasts. An example of a professional penalty would be the exclusion from one's professional organization if one violates its code of ethics. An example of a criminal penalty would be punishment as the result of prosecution before a court or similar legal set-up, for instance where deceptive forecasts have led to substantial mismanagement of public funds (Garett and Wachs, 1996). Malpractice in planning should be taken as seriously as it is in other professions. Failing to do this amounts to not taking the profession of planning seriously.

In order to achieve accountability in forecasting via *competition and market control*, the following would be required, again as practices that are both embedded in and enforced by the relevant institutions:

- The decision to go ahead with a project should, where at all possible, be made contingent on the willingness of private financiers to participate without a sovereign

guarantee for at least one third of the total capital needs.<sup>25</sup> This should be required whether projects pass the market test or not, that is, whether projects are subsidized or not or provided for social justice reasons or not. Private lenders, shareholders, and stock market analysts would produce their own forecasts or would critically monitor existing ones. If they were wrong about the forecasts, they and their organizations would be hurt. The result would be more realistic forecasts and reduced risk.

- Full public financing or full financing with a sovereign guarantee should be avoided.
- Forecasters and their organizations must share financial responsibility for covering cost overruns and benefit shortfalls resulting from misrepresentation and bias in forecasting.
- The participation of risk capital should not mean that government gives up or reduces control of the project. On the contrary, it means that government can more effectively play the role it should be playing, namely as the ordinary citizen's guarantor for ensuring concerns about safety, environment, risk, and a proper use of public funds.

Whether projects are public, private, or public-private (PPPs), they should be vested in one and only one project organization with a strong governance framework. The project organization may be a company or not, public or private, or a mixture. What is important is that this organization enforces accountability vis-à-vis contractors, operators, etc., and that, in turn, the directors of the organization are held accountable for any cost overruns, benefits shortfall, faulty designs, unmitigated risks, etc. that may occur during project planning, implementation, and operations.

If the institutions with responsibility for developing and building major investment projects would effectively implement, embed, and enforce such measures of accountability, then the misrepresentation in cost, benefit, and risk estimates, which is widespread today, may be mitigated. If this is not done, misrepresentation is likely to continue, and the allocation of funds for major projects is likely to continue to be wasteful and undemocratic.

## 6.8 Towards Better Practice

Fortunately, after decades of widespread mismanagement of the planning and design of large investment projects, signs of improvement have recently appeared. The conventional consensus that deception is an acceptable way of getting projects started is under attack, as will be apparent from the examples below. This is in part because good governance is generally getting more attention around the world. The Enron scandal and its successors have triggered a war on corporate deception that is spilling over into government with the same objective: to curb financial waste and promote good governance. Although progress is slow, good governance is gaining a foothold even in large investment projects.

The conventional consensus is also under attack for the practical reason that the largest projects are now so big in relation to national economies that cost overruns, benefit shortfalls, and risks from even a single project may affect the finances of a whole country or region, as happened in Greece with the cost overruns on the Athens 2004 Olympics and in Hong Kong

---

<sup>25</sup> The lower limit of a one-third share of private risk capital for such capital to effectively influence accountability is based on practical experience. See more in Flyvbjerg, Bruzelius, and Rothengatter (2003: 120-123).



with large initial demand shortfalls for the new Chek Lap Kok airport. Lawmakers and governments begin to see that national fiscal distress is too high a price to pay for the conventional way of planning and designing large projects. The main drivers of reform lie outside the agencies and industries conventionally involved in major project development, which increases the likelihood of success.

In 2003 the Treasury of the United Kingdom required, for the first time, that all ministries develop and implement procedures for large public projects that will curb what it calls--with true British civility--"optimism bias." Funding will be unavailable for projects that do not take into account this bias, and methods have been developed for how to do this (Mott MacDonald, 2002; HM Treasury, 2003; Flyvbjerg and Cowi, 2004). In the Netherlands in 2004, the Parliamentary Committee on Infrastructure Projects for the first time conducted extensive public hearings to identify measures that will limit the misinformation about large infrastructure projects given to the Parliament, public, and media (Tijdelijke Commissie Infrastructuurprojecten, 2004). In Boston, the government sued to recoup funds from contractor overcharges for the Big Dig related to cost overruns.

More governments and parliaments are likely to follow the lead of the UK, the Netherlands, and Boston in coming years. It's too early to tell whether the measures they implement will ultimately be effective. It seems unlikely, however, that the forces that have triggered the measures will be reversed, and it is those forces that reform-minded groups need to support and work with in order to curb misinformation and waste. This is the "tension-point" where convention meets reform, power-balances change, and new things may happen.

The key weapons in the war on misinformation and waste is accountability and critical questioning. The professional expertise of planners, engineers, architects, economists, and administrators is certainly indispensable to constructing the infrastructures that make society work. Studies show, however, that the claims about costs, benefits, and risks made by these groups usually cannot be trusted and should be carefully examined by independent specialists and organizations. The same holds for claims made by project-promoting politicians and officials. Institutional checks and balances--including financial, professional, or even criminal penalties for consistent and unjustifiable biases in claims and estimates of costs, benefits, and risks--should be developed and employed. The key principle is that the cost of making a wrong forecast should fall on those making the forecast, a principle often violated today.

If done right, many of the public-private partnerships currently emerging in major projects--including increased involvement of banks and large capital funds--may involve more and better checks and balances than previous institutional setups, as has been demonstrated by the UK National Audit Office (2003). This can be a step in the right direction but should be no cause for repose. All available measures for improvement must be employed. There is no easy fix. The conventional mode of planning and designing major projects has long historical roots and is deeply ingrained in professional and institutional practices. It would be naive to think it is easily toppled. Given the stakes involved--saving taxpayers from billions of dollars of waste, protecting citizens' trust in democracy and the rule of law, avoiding the destruction of spatial and environmental assets--this shouldn't deter us from trying.

## 6.9 Summary

A main problem in major investment projects is a high incidence of cost overruns and benefit shortfalls caused by optimism bias and strategic misrepresentation in early project development. This paper presents measures aimed at reducing and eliminating optimism bias and strategic misrepresentation. The measures include the use of a new forecasting method called "reference class forecasting," based on ideas of decision making under uncertainty that won the Nobel Prize in economics 2002. The measures also include improved governance structures with incentives that better reward valid estimates of costs, benefits, and risks and punish deceptive estimates. The aim is to ensure that decisions on whether to build projects or not are based on valid information about costs, benefits, and risks instead of being based on misinformation as is often the case today.

## 6.10 References

- Altshuler, A. and D. Luberoff, 2003, *Mega-Projects: The Changing Politics of Urban Public Investment* (Washington, DC: Brookings Institution).
- American Planning Association, 1991, AICP Code of Ethics and Professional Conduct. Adopted October 1978, as amended October 1991, <http://www.planning.org>
- American Planning Association, 2005, "JAPA Article Calls on Planners to Help End Inaccuracies in Public Project Revenue Forecasting" <http://www.planning.org/newsreleases/2005/ftp040705.htm>, April 7
- Flyvbjerg, B., 1996, The dark side of planning: Rationality and *Realrationalität*. In S. Mandelbaum, L. Mazza, and R. Burchell (Eds.), *Explorations in Planning Theory* (New Brunswick, NJ: Center for Urban Policy Research Press), pp 383-394
- Flyvbjerg, B., 1998, *Rationality and power: Democracy in practice* (Chicago: University of Chicago Press)
- Flyvbjerg, Bent, 2005, "Design by Deception: The Politics of Megaproject Approval." *Harvard Design Magazine*, no. 22, Spring/Summer, pp 50-59
- Flyvbjerg, Bent, 2007, "How Optimism Bias and Strategic Misrepresentation In Early Project Development Undermine Implementation." Chapter XXX in this volume.
- Flyvbjerg, Bent, Nils Bruzelius, and Werner Rothengatter, 2003, *Megaprojects and Risk: An Anatomy of Ambition* (Cambridge University Press)
- Flyvbjerg, Bent and Cowi, 2004, *Procedures for Dealing with Optimism Bias in Transport Planning: Guidance Document* (London: UK Department for Transport)
- Flyvbjerg, Bent, Mette K. Skamris Holm, and Søren L. Buhl, 2002, "Underestimating Costs in Public Works Projects: Error or Lie?" *Journal of the American Planning Association*, vol. 68, no. 3, Summer, pp 279-295

- Flyvbjerg, Bent, Mette K. Skamris Holm, and Søren L. Buhl, 2005, "How (In)accurate Are Demand Forecasts in Public Works Projects? The Case of Transportation." *Journal of the American Planning Association*, vol. 71, no. 2, Spring, pp 131-146
- Garett, M. and Wachs, M., 1996, *Transportation Planning on Trial: The Clean Air Act and Travel Forecasting* (Thousand Oaks, CA: Sage)
- Gilovich, T., Griffin, D., and Kahneman, D., 2002, Eds., *Heuristics and Biases: The Psychology of Intuitive Judgment* (Cambridge University Press, 2002)
- Gordon, P. and Wilson, R., 1984, "The Determinants of Light-Rail Transit Demand: An International Cross-Sectional Comparison." *Transportation Research A*, 18A(2), pp 135-140
- HM Treasury, 2003a, *The Green Book: Appraisal and Evaluation in Central Government, Treasury Guidance* (London: TSO)
- HM Treasury, 2003b, *Supplementary Green Book Guidance: Optimism Bias* (London: Author).
- Kahneman, D., 1994, New challenges to the rationality assumption. *Journal of Institutional and Theoretical Economics*, 150, pp 18-36
- Kahneman, D. and Tversky, A., 1979a, "Prospect theory: An analysis of decisions under risk." *Econometrica*, 47, pp 313-327
- Kahneman, D. and Tversky, A., 1979b, "Intuitive Prediction: Biases and Corrective Procedures." In S. Makridakis and S. C. Wheelwright, Eds., *Studies in the Management Sciences: Forecasting*, 12 (Amsterdam: North Holland).
- Lovullo, Dan and Daniel Kahneman, 2003, "Delusions of Success: How Optimism Undermines Executives' Decisions," *Harvard Business Review*, July, pp 56-63
- Mott MacDonald, 2002, *Review of Large Public Procurement in the UK*, study for HM Treasury (London: HM Treasury)
- Ove Arup and Partners Scotland, 2004, *Scottish Parliament, Edinburgh Tram Line 2 Review of Business Case* (West Lothian, Scotland: Author).
- Royal Town Planning Institute, 2001, Code of professional conduct. As last amended by the Council on 17 January 2001, <http://www.rtpi.org.uk>
- Tijdelijke Commissie Infrastructuurprojecten, *Grote Projecten Uitvergroot: Een Infrastructuur voor Besluitvorming* (The Hague: Tweede Kamer der Staten-Generaal, 2004).
- UK National Audit Office, 2003, *PFI: Construction Performance*, report by the Comptroller and Auditor General, HC 371 Session 2002–2003: February 5, 2003 (London: National Audit Office)

- Wachs, M., 1989, When Planners Lie with Numbers. *Journal of the American Planning Association*, 55(4), pp 476-479
- Watson, V., 2003, Conflicting Rationalities: Implications for Planning Theory and Ethics. *Planning Theory and Practice*, 4(4), pp 395-408
- Yiftachel, O., 1998, Planning and Social Control: Exploring the Dark Side. *Journal of Planning Literature*, 12(4), pp 395-406

## 7 DATAANALYSE OG PROSJEKTRISIKO VED SPARSOM INFORMASJON

JOSTEIN LILLESTØL\*

### 7.1 Innledning

Dette kapittelet inneholder tre hovedavsnitt som kan leses uavhengig av hverandre:

- Kvalitetssikring av datagrunnlaget for prosjekter
- Dataanalyse av usikkerhet og risiko
- Bayesiansk sannsynlighetsvurdering

Avsnittene setter stigende krav til leseren mht. fortrolighet med begreper og symboler. Det første hovedavsnittet om kvalitetssikring forutsetter ingen spesielle forkunnskaper. Det neste hovedavsnittet tar opp temaet dataanalyse i forbindelse med kostnadsvurderinger. Her er det en fordel å ha gjennomgått elementærkurs i sannsynlighetsregning og statistikk.<sup>26</sup> Vi skal imidlertid presentere sentrale begreper og analysemåter stort sett uten formler, slik at avsnittet også skal kunne være lett tilgjengelig for lesere uten slik bakgrunn. I det siste avsnittet tar vi for oss prosjektvurdering basert på subjektive sannsynligheter og oppdatering av slike (ved Bayes lov), et tema som har fått stigende oppmerksom de senere år, i situasjoner der det foreligger lite data. Det vil her typisk være tale om ekspertvurderinger, et tema som tas opp i et eget kapittel. For å presentere ideene i dette hovedavsnittet, er det dessverre ikke lett å unngå formler. Disse bør likevel ikke være avskrekkende, og det er vårt håp at den interesserte leser kan finne noen stimulerende ideer her.

### 7.2 Kvalitetssikring av datagrunnlaget for prosjekter - Perspektiv

Kvalitetssikring knyttet til datainnsamling og analyse dreier seg om kunne se muligheter og forhindre feil. Blant potensielle risikoer finner vi:

- formålet er dårlig definert
- innsamlede data samsvarer ikke med behovet
- analysen oppfattes ikke som relevant eller objektiv av beslutningstaker
- analysen oppfattes ikke som relevant eller objektiv av eventuell motpart
- tvilsomme analysemetoder benyttes

---

\* Jostein Lillestøl er professor ved Institutt for foretaksøkonomi, NHH.

<sup>26</sup> En aktuell lærebok er Ubøe og Jørgensen: Statistikk for økonomifag. Gyldendal 2004.

- staben er ikke trent til å analysere eller forsvare resultatet av analysen
- data er feilkodet
- data går tapt
- kvalitetskontrolltiltak er ikke virksomme

Spørsmål om hva en trenger av data til et prosjekt vil typisk skje ved en gruppeprosess, der involverte parter og statistisk ekspertise er representert. En kan med fordel benytte de råd for gruppeprosesser som er beskrevet i litteraturen for kvalitetsforbedringsarbeid, enten det er kvalitetssirkler, forbedringsgrupper eller ”process reengineering”.

### 7.3 Kriterier for vurdering av datakvalitet

En vurdering av datakvalitet kan skje ut fra følgende kriterier, der en har hele tiden har prosjektets formål som ”bakteppe”:

**Relevans:**

- I hvilken grad kan denne typen data bidra til å belyse viktige aspekter ved prosjektet?

**Nøyaktighet:**

- I hvilken grad beskriver dataene korrekt det som skal måles eller beskrives?

**Tilgjengelighet:**

- Er dataene lett å få tak i, og på egnet medium?

**Dokumenterbarhet:**

- Er dataene tilgjengelige fra pålitelige kilder og kan dette dokumenteres?

**Tolkbarhet:**

- Er dataene enkle å analysere, fortolke og formidle?

**Kombinerbarhet:**

- Lar dataene seg kombinere med øvrige data, kvantitative og kvalitative?

**Supplerbarhet**

- Lar dataene seg supplere i løpet av prosjektet ved behov for mer detaljering?

**Gyldighet:**

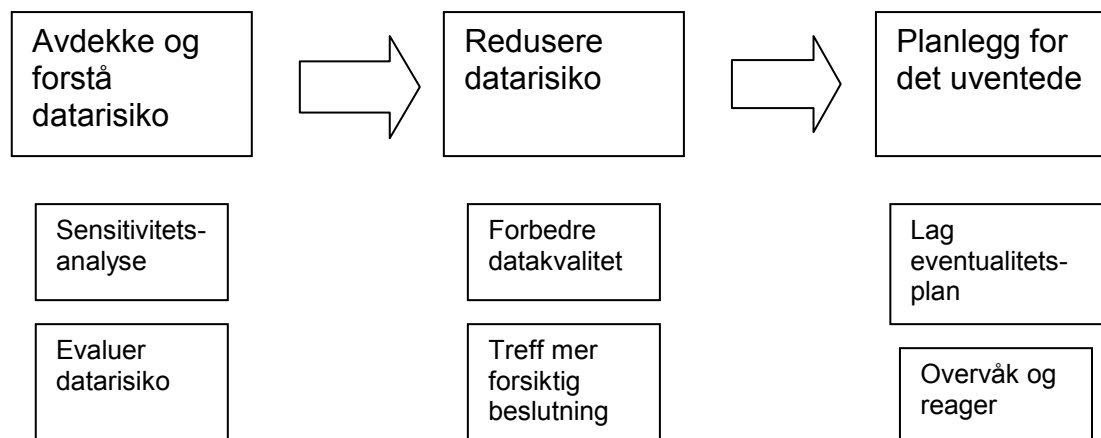
- I hvilken grad er data fortsatt nyttige når de viktige beslutninger skal tas?
- Oppfyller dataene de krav som tilsynsmyndighetene stiller, f.eks. personvern?

### 7.4 Risikostyring av datagrunnlag og bruken av data

Kvalitetssikring av datagrunnlaget for et prosjekt bør være like selvfølgelig som sikring av andre aspekter. Dersom datakvaliteten er dårlig risikerer en ikke bare at feil beslutninger treffes, men også at fremdriften i prosjektet forsinkes eller stoppes, fordi det på et senere tidspunkt reises tvil om datagrunnlaget. I noen situasjoner kan nok en prosjektledelse bevisst holde tilbake de mest relevante og pålitelige data for å kunne få prosjektet realisert, f.eks. lave

kostnadsestimater til en bevilgende myndighet. I så måte kan rådene her være til hjelp for medspillere og motspillere som er på vakt.

Kvalitetssikring av datagrunnlaget kan deles i tre trinn, som illustrert i Figur 7-1:



Figur 7-1 Risikostyring av datagrunnlaget

### Trinn 1 - Forstå datarisiko

Det første trinnet - å forstå risikoen - omfatter to elementer (i) sensitivitetsanalyse (ii) evaluering av datarisiko.

Hensikten med sensitivitetsanalysen er å identifisere risikoen ved å bruke en aktuell datakilde for å dekke et bestemt databehov. Dette kan skje ved å stille seg hvert av spørsmålene knyttet til kriteriene i forrige avsnitt, og så se på den mulige innvirkning bruk av dataene har for prosjektet, og de beslutninger som skal tas.

Evaluering av datarisiko kan skje ved å sjekke samsvaret mellom dataene og databehovet for hvert tema, og for eksempel måle det på en skala fra ”disse data dekker overhodet ikke noen del av behovet” til ”disse data dekker langt mer enn behovet”, fortrinnsvis ved bruk av et ferdig utarbeidet vurderingsskjema.

Sensitivitetsanalysen kan skje ved å forsøke å besvare følgende to spørsmål:

- Hvor avvikende må data være for at den aktuelle beslutning skal bli en annen?
- Hvor sannsynlig er det at dataene vil være såpass avvikende som følge av manglende samsvar mellom databehovet og datakilden?

## Trinn 2 - Redusere datarisiko

Dette trinn omfatter undersøkelse av muligheten for å redusere risikoen. Det foreligger to muligheter:

- a. Forbedre datakvaliteten.
- b. Treffe en mer forsiktig beslutning

### **a. Forbedre datakvaliteten**

Forbedring av datakvaliteten dreier seg i hovedsak om å oppnå bedre samsvar mellom databehov og datakilde. Blant de muligheter som kan vurderes er:

- Bruke den aktuelle datakilden til en mindre del av prosjektet enn opprinnelig tenkt.
- Bruke flere datakilder, enten til validering av eksisterende data, eller til supplering.
- Skaffe mer informasjon om og fra den aktuelle datakilden.
- Endre den aktuelle datakilde: innsamlingsmåte, datalevering og tilgjengelighet
- Forlate eksisterende datakilde, og finne en ny (som må vurderes tilsvarende).

Dersom data fra den eksisterende datakilden er utilstrekkelig mht. omfang, kan det være et spørsmål å utvide materialet, for eksempel øke antall observasjoner/stikkprøve.

I noen tilfeller dekker data mer enn behovet, og det er mulig å redusere datamengden. For mye data kan innebære at man ikke ser skogen for bare trær, og at man trenger mer avanserte statistiske metoder for å avsløre strukturer og sammenhenger. Ovenstående kan brukes til å forbedre enhver datakilde, men knapphet på ressurser og tid kan innebære at må nøyse seg med den kvalitet man ser seg råd med. Er man nødt til å inngå kompromisser må man vurdere:

- Hvor skal man inngå kompromisser?
- Dersom man må inngå kompromisser, vil data fortsatt møte kravene?

### **b. Treffe mer forsiktige beslutninger**

Beslutningene må ikke bare reflektere data, men også kvaliteten i data. Er man klar over at datakvaliteten er dårlig eller i tvil om denne, kan det i noen tilfeller være aktuelt å treffe mer konservative beslutninger. Noen prosjekter kan gjennomføres i mindre omfang eller som prøveprosjekter. For mange prosjekter, for eksempel innen offshore dreier det seg imidlertid om en skal realisere et større prosjekt eller ikke, og når beslutningen om å sette i gang er tatt, er det ingen vei tilbake. Muligheten for mer konservative beslutninger avhenger sterkt av område og typen prosjekt, og det er vanskelig å gi allmenngyldige anbefalinger. Det hyppigst forekommende konservative beslutning er trolig utsetting, som gjerne innebærer krav om ny eller bedre informasjon.



### Trinn 3 - Planlegge for det uforutsette

#### a. Lag en eventualitetsplan

Når risikoer er avdekket og klassifisert iht. Trinn 1 og risiko redusert iht. Trinn 2, vil det fortsatt kunne være betydelig risiko knyttet til data. Det gjelder alt fra ekspertvurderinger basert (mer eller mindre) på data, til statistiske estimater som kan ha skjevhet ("bias") og feilmargin. En eventualitetsplan ("contingency plan") dreier seg om å planlegge for mulige situasjoner der forventninger ikke slår til. Noen slike situasjoner kan være knyttet til selve data, for eksempel dersom man har utført en etterspørselsanalyse i en tidlig fase, og at det ganske snart skjer ting i markedet som gjør denne mindre relevant.

#### b. Overvåk og reager

Overvåking er nøkkelen til planlegging for det uforutsette. Det er viktig at den informasjonen fins som eventualitetsplanen krever, og at den blir lagt merke til. Det gjelder ikke bare overvåking av budsjetter og andre økonomiske nøkkeltall, men også tidsbruk og andre siden ved virksomheten. I denne sammenheng kan ideer fra statistisk prosessstyring og bruk av styringsdiagrammer ("control charts") kunne være til nytte.

Sluttkommentar: Mange lar forventninger inngå som faste tall i planleggingsfasen, uten å ta omsyn til tilfeldig variasjon omkring forventningsverdien. Dette vil typisk kunne medføre undervurdering av prosjektrisikoen. Datagrunnlaget bør derfor omfatte kunnskap om risikoen for avvik fra forventninger, og slike data må kvalitetssikres på linje med data om forventningen selv. Dette er hovedtema i neste avsnitt.

## 7.5 Dataanalyse av usikkerhet og risiko - Allmenne betraktninger

Prosjektvurdering i tidligfasen innebærer som regel stor usikkerhet mht. nytte, kostnad og realisering. Forsøk på tallfesting er i beste fall en forventning, og hvor velbegrunnet den er, kan ofte diskuteres. En forventning kan i mer eller mindre grad være forankret i erfaring. På den ene side kan den ha preg av ren subjektivitet, mens på den annen side kan den være basert på omfattende erfaring fra tidligere prosjekter, men da kan det reises tvil om hvor relevant denne erfaring er for det nye prosjekt.

Mange bruker (dessverre) en slik forventning direkte i den videre analyse som om de var sikre tall. Her er det to forhold en bør ha i mente:

- i. risiko for avvik fra forventningen, selv når denne er kjent
- ii. usikkerheten i anslaget for forventningen når denne er ukjent

Når forventningen er ukjent og anslås, kommer den første usikkerheten på toppen av den andre. Statistisk teori og metode kan bidra til å få et bedre grep på begge utfordringene.

Ved planlegging av prosjekter underlagt usikkerhet hører det med å skaffe best mulig kunnskap om variasjon, dvs. risikoen for avvik fra forventninger, og herunder spesielt ekstremrisiko vedrørende prosjektets gjennomføring. Det hjelper lite å ha gode estimater på

forventningen, dersom den genuine variasjon omkring denne, som planleggeren ikke kan gjøre noe med, er stor og ukjent.

En annen utfordring i tidligfasen av prosjekter går på forståelsen av deler og helhet i prosjektet. På den ene side har vi at prosjektet som helhet vurderes langs ulike dimensjoner, noen økonomiske andre ikke-økonomiske. På den annen side har vi studier av sammenhengen mellom deler og helhet, og som kan være kritiske for gjennomføringen. Et godt eksempel er utbyggingen av ny hovedflyplass i Norge. I tillegg til kostnad-nytte vurderingen knyttet til selve flytrafikken, herunder tilkomst for passasjerer, kommer konsekvenser for bosetning i nærhet og region, med sine fordeler og ulemper. Selv om enkelte av disse aspekter kan oversettes til økonomiske mål og "legges sammen", vil det typisk være mange som det verken er naturlig eller ønskelig å oversette. I praksis blir det et spørsmål om egnet valg av attributter ved prosjektet, valg av måleenhet (ofte ordinal skala) og vektlegging i den sammenfattende vurdering. Også her er det en utfordring å kunne ta omsyn til usikkerheten. En annen utfordring ved "å legge sammen" deler til helhet, for eksempel kostnader ved ulike deler eller trinn i et prosjekt, er å skaffe innsikt i den rolle mulig samvariasjon kan spille, og å kunne ta omsyn denne når beslutninger skal treffes. Selv om forventningen til en sum av kostnader er lik summen av forventningene, kan en ikke legge sammen usikkerhetene på samme måte. Her spiller eventuell samvariasjonen inn. Vi skal belyse dette tema i et senere avsnitt.

Kostnader sammen med nytte er de viktigste tema ved prosjekter, i tidligfasen som senere. Kostnadsanslag i tidligfasen har som kjent (minst) to formål

1. Grunnlag for beslutningstakere om valg av prosjekt.
2. Kostnadsdempende rettesnor ved prosjektgjennomføring

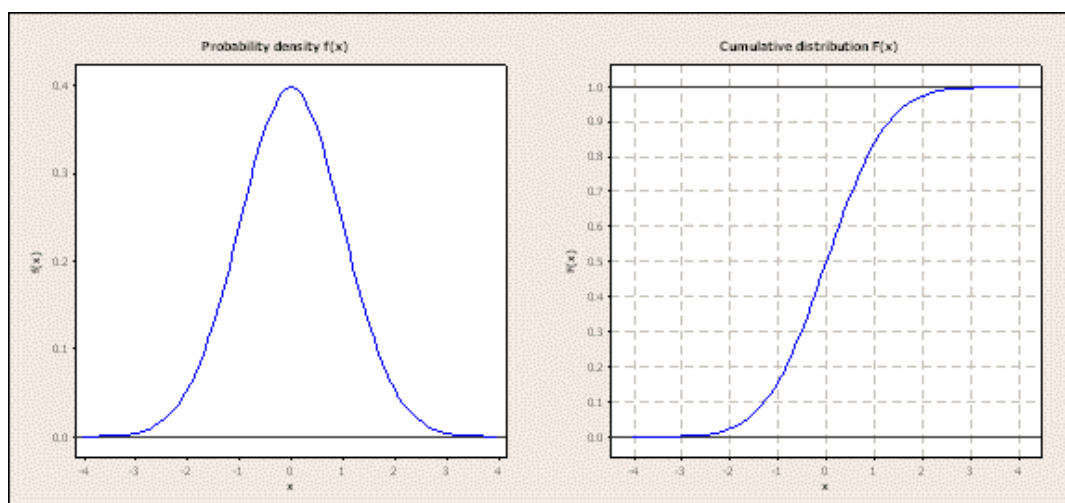
For formål 1 er det ønskelig at kostnadsanslaget er realistisk, mens for formål 2 gjør det mindre om anslaget er lavt for å kunne holde kostnadene nede. I praksis hender det imidlertid ofte at beslutningstakere mottar lave anslag for noen eller alle prosjektalternativer, fordi planlegger har interesse av et bestemt prosjekt eller unngå at alle skrinlegges. Når prosjektet er valgt og under gjennomføring, blir det ofte fort klart at kostnadsanslagene var urealistiske, og at kostnadsøkninger er uunngåelige. Dette er eksempel systematisk kostnadsavvik som skyldes taktiske forhold, noe som entreprenører ofte vet (og beslutningstakere burde vite) å ta omsyn til. I praksis vil en selvfølgelig også kunne ha undervurdering av kostnadene (sjelden overvurdering) pga. uforutsette forhold og prisøkninger osv.

Ved sammenligning av flere prosjekter med sikte på valg, der et skal gjennomføres, dvs. det ikke er aktuelt å skrinlegge alle, er det ikke så nødvendig at de absolutte kostnadsanslag er riktige. Det er nok at prosjektene sammenlignes langs de samme dimensjoner, med de samme variable, målt etter de samme prinsipper og analysert på samme måte. Når dette ikke er mulig, kan statistisk metode i noen tilfeller bidra til å gjøre ting sammenlignbare, ting som i utgangspunktet ikke er det.

## 7.6 Kostnader som tilfeldige variable

Vi trenger en egnet terminologi for å håndtere variable underlagt usikkerhet. Vi betrakter en kostnad  $X$  som en tilfeldig variabel, ofte kalt stokastisk variabel (engelsk: random variable, stochastic variable) med verdiområde de ikke-negative reelle tall. Den tilfeldige variable  $X$  har en tenkt sannsynlighetsfordeling uttrykt ved dens såkalte sannsynlighetstetthet  $f(x)$  eller ved dens kumulative fordeling  $F(x)$ .<sup>27</sup> Fordelingen er i praksis mer eller mindre kjent, i noen tilfeller er den rent subjektiv. Fordelingen kan karakteriseres ved sin forventningsverdi  $E(X)$  og varians  $V(X)$ . Noen saksforhold uttrykkes bedre med standardavvik  $SD(X)$  som er kvadratroten av variansen. Dette er et (ufullkomment) uttrykk for risikoen for avvik fra forventningen. Denne terminologi er generell, og gjelder derfor for andre variable enn kostnader, for eksempel tidsforbruk. For presise definisjoner av disse begrepene, se lærebøker i statistikk.<sup>28</sup>

Mange vil trolig ha kjennskap til normalfordelingen, også kalt Gauss-fordelingen. Den er ”klokkeformet” og symmetrisk omkring sin forventning: I figuren vises tettheten og den kumulative fordelingen dersom  $X$  er normalfordelt med forventning 0 og standardavvik (og dermed varians) lik 1.<sup>29</sup>



Dette er standardnormalfordelingen, som ofte betegnet  $N(0, 1)$ . Her er arealet under tettheten fra -1 til +1 lik 0.68. Den kumulative sannsynlighet opp til 1 er 0.84.

<sup>27</sup> En sannsynlighetstetthet er en ikke-negativ funksjon  $f(x)$  som definerer en kurve som avgrensar totalt areal lik 1, og der sannsynligheten for at variabelen faller i et gitt område er lik arealet under kurven over dette området. En kumulativ fordelingsfunksjon  $F(x)$  gir sannsynligheten for at variabelens verdi er høyst lik  $x$ , som derfor er lik arealet under tetthetskurven opp til  $x$ .

<sup>28</sup> For eksempel Ubøe og Jørgensen: Statistikk for økonomifag. Gyldendal 2004.

<sup>29</sup> Beregninger og grafiske illustrasjoner i dette kapittelet er utført med den statistiske programpakken Minitab™.

For normalfordelingen er sammenhengen mellom standardavvik og risiko enkel, vi har nemlig at sannsynligheten for at utfallet avviker høyst  $k$  ganger standardavviket fra forventningen er gitt ved  $A(k)$ , der vi for noen valg av  $k$  har:

$k$	1	1.5	2	2.5	3	4
$A(k)$	68.3%	86.6%	95.4 %	98.8%	99.7%	99.99%

Disse sannsynlighetene finner vi ut fra arealer under standardnormalfordelingen mellom  $-k$  og  $k$ . Merk imidlertid at de gjelder for alle normalfordelinger.

Eksempel: En usikker kostnad antas å kunne bli en verdi mellom 85 og 115 (enhet mill. kr.), med 100 som forventet verdi. Antar vi normalfordeling, slik at variasjonsområdet (med stor sannsynlighet) tilsvarer  $\pm 3$  standardavvik, må standardavviket være 5. Altså har vi normalfordelingen  $N(100, 5)$ . Vi ønsker å vite sannsynligheten for at kostnaden overstiger 110, og bruker programvare til å beregne sannsynligheten for høyst 110, dvs. en kumulativ sannsynlighet. Vi får

### **Cumulative Distribution Function**

Normal with mean = 100 and standard deviation = 5

```
x      P( X <= x )
110    0.977
```

Den søkte sannsynlighet er derfor  $1-0.977 = 0.023$ . Vi kunne alternativt funnet dette ved bemerke at 110 ligger 2 ganger standardavviket over forventningen, slik at det er tale om å finne arealet under standardnormalkurven over 2, som er 2.3%

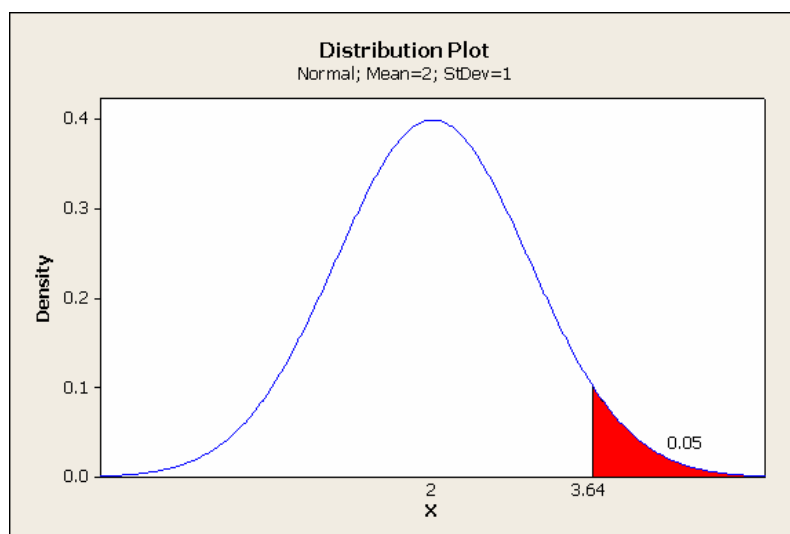
Med kjennskap til eller et godt anslag på forventning og standardavvik, er vi derfor i stand til å uttale oss om risiko for avvik, men altså under forutsetning av normalfordelt variasjon.

For andre fordelinger gir disse sannsynlighetene i beste fall bare en grov antydning på risiko, Når mange i praksis regner med at all variasjon skjer innenfor 3 ganger standardavviket fra forventningen, kan det være for optimistisk, og i ekstreme tilfeller er standardavviket ikke lenger særlig informativt. Det kan synes urimelig med en slik antakelse for kostnader, som ikke kan anta negative verdier. Det behøver imidlertid ikke bekymre oss, da den teoretiske sannsynligheten for dette er neglisjerbar. Mer bekymringsfullt er symmetrien og høyrehalen i fordelingen. Man vil typisk være interessert i ekstremrisiko. Da spiller også fordelings skjevhet (skewness) og kurtose (kurtosis) en rolle. Det siste er et uttrykk for fordelings toppethet/haletyngde. For ekstremrisiko er imidlertid sannsynlighetsfordelingen i seg selv et bedre utgangspunkt. Noe kjennskap til fordelingsteori, aktuelle fordelinger og deres egenskaper vil være nyttig, men vi kan ikke gå i detalj her, utover å kommentere grafiske illustrasjoner.

En god kilde til kunnskap er appendikset i boken Aven (2003), en bok som anvender fordelingsteorien i risikoanalyse.

Det fins en rekke skjeve sannsynlighetsfordelinger som er aktuelle som modell for kostnader evt. merkostnader. Blant disse er triangulær, lognormal-, Gamma-, Weibull-fordeling, som vi skal ta for oss nedenfor. La oss ta utgangspunkt i et eksempel, der vi i utgangspunktet fortsatt antar normalitet.

Eksempel: Vi studerer merkostnad i forhold til budsjett, og tror at forventet merkostnad er 2 millioner kroner. Vi er primært interessert i sannsynligheten for ekstreme merkostnader. I denne sammenheng vurderes standardavviket til 1 mill. kr. Dette innebærer en betydelig sjanse for negativ merkostnad, som i så fall vil bety kostnad mindre enn budsjettet. Med mill. kroner som enhet viser figuren nedenfor normalfordelingen  $N(2, 1)$ .



Det høyrearealet som utgjør 5% av totalarealet under fordelingskurven er markert, slik at vi kan lese av hvilken merkostnad som har 5% sjanse for å overskrides (den såkalte øvre 5% fraktilen i fordelingen), her 3.64. Ved bruk av programvare må en spesifisere at en ønsker det punkt som gir kumulativ sannsynlighet lik 0.95, og vi får

### ***Inverse Cumulative Distribution Function***

Normal with mean = 2 and standard deviation = 1

$P( X \leq x )$	$x$
0.95	3.64485

Vi kan imidlertid finne dette straks vi kjenner øvre fraktiler i standardnormalfordelingen, som er

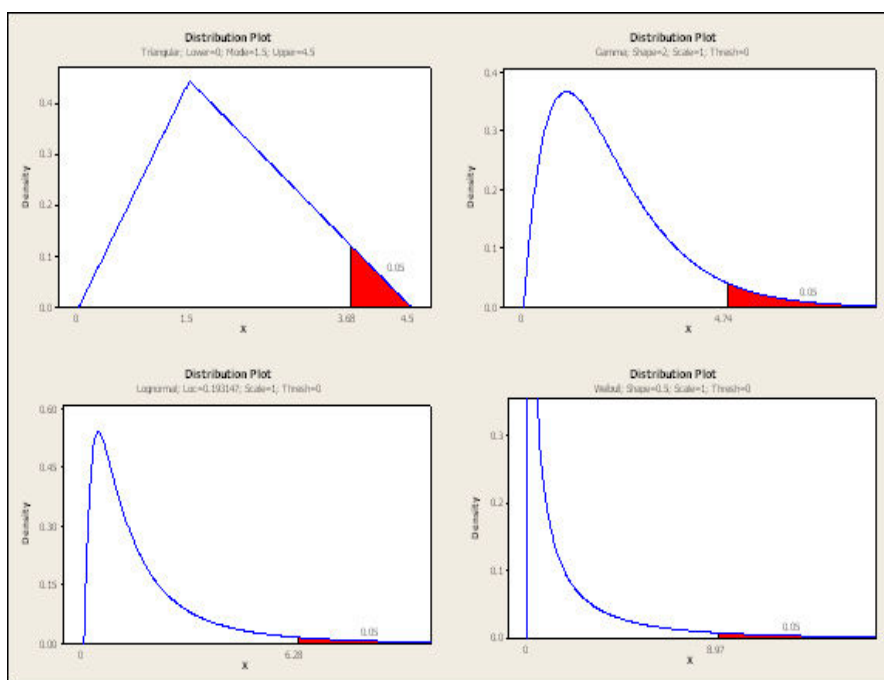
<i>Halesannsynlighet</i>	10%	5%	1%	0.5%	0.1%	0.05%	0.01%
<i>Øvre fraktil <math>N(0,1)</math></i>	1.28	1.64	2.33	2.58	3.09	3.29	3.72

For generell normalfordeling er øvre fraktil gitt ved

$$\text{Øvre fraktil} = \text{Forventning} + \text{''Standard'' øvre fraktil} \times \text{standardavviket}$$

I vårt eksempel blir derfor øvre 5% fraktilen lik  $2 + 1.64 \times 1 = 3.64$  som beregnet ovenfor. Tabellen ovenfor kan også brukes motsatt: Hvis en tror at det er kun 1 promilles sjans for at merkostnaden overskrider 5 mill. kr. betyr dette ca.  $3.09 \times$  standardavviket over forventningen på 2 mill.kr. må tilsvare 3 mill. kr. Dette skjer når standardavviket er ca. 1 mill. kr. som antatt. Dersom vi tror at denne sjansen er mye større, f.eks 5%, svarer det til  $1.64 \times$  standardavviket, som betyr at standardavviket ved normalitet er  $3/1,64=1.82$ .

Fordelingstyper som kan reflektere skjevhet med lang høyrehale er blant andre: triangulær, lognormal, Gamma og Weibull-fordeling. Figuren nedenfor gis eksempler på disse fire fordelingene for samme forventning lik 2, der det igjen dreier seg om merkostnader i millioner kroner.<sup>30</sup>

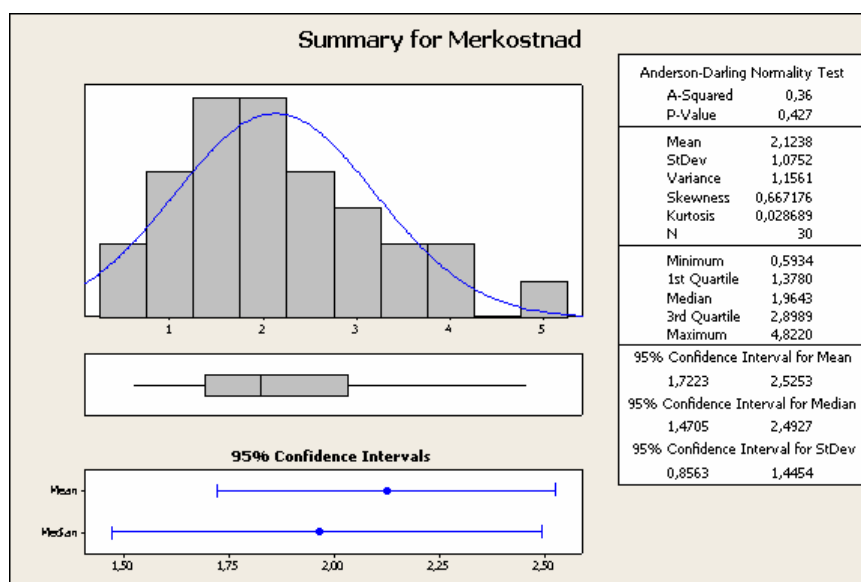


Vi ser at alle fordelingene starter i null, men ved å spesifisere terskelverdi (engelsk: "threshold"), kan den starte hvor som helst. Selv om forventningene er de samme i de fire eksemplene, er variansen ulik. Fordelingenes form og markeringene som uttrykker ekstremrisikoen, er derfor ikke direkte sammenlignbare. I praksis er valg av fordeling avhengig av empirisk erfaring og teoretisk kunnskap, kombinert med ønsket om enklest mulig modell. Statistisk programvare kan

<sup>30</sup> Nevnes bør også Beta-fordelinger, som kan reflektere både symmetriske og asymmetriske situasjoner. Disse har verdiområde fra 0 til 1, med den rektangulære fordeling som spesialtilfelle. Beta-fordelingen lar seg enkelt modifiseres til et hvilket som helt intervall på tallinjen.

være til hjelp for å identifisere modell. På den ene siden kan slik programvare brukes (som ovenfor) til å studere ulike teoretiske fordelinger grafisk for ulike valg av parametere (form, skala). På den annen side kan slik programvare brukes til å finne en brukbar modell direkte ut fra data. Man vil typisk erfare at flere fordelingstyper er forenlig med dataene, spesielt dersom man har lite data, der kanskje ekstreme tilfeller tilfeldigvis ikke forekom. Disse fordelingene kan derfor innebære nokså ulik ekstremrisiko. Kunnskap om de ulike fordelingers egenskaper er derfor verdifull når en skal gjøre realistiske risikovurderinger i de tilfeller erfaringsgrunnlag er svakt, slik det ofte er med prosjekter i tidligfasen.

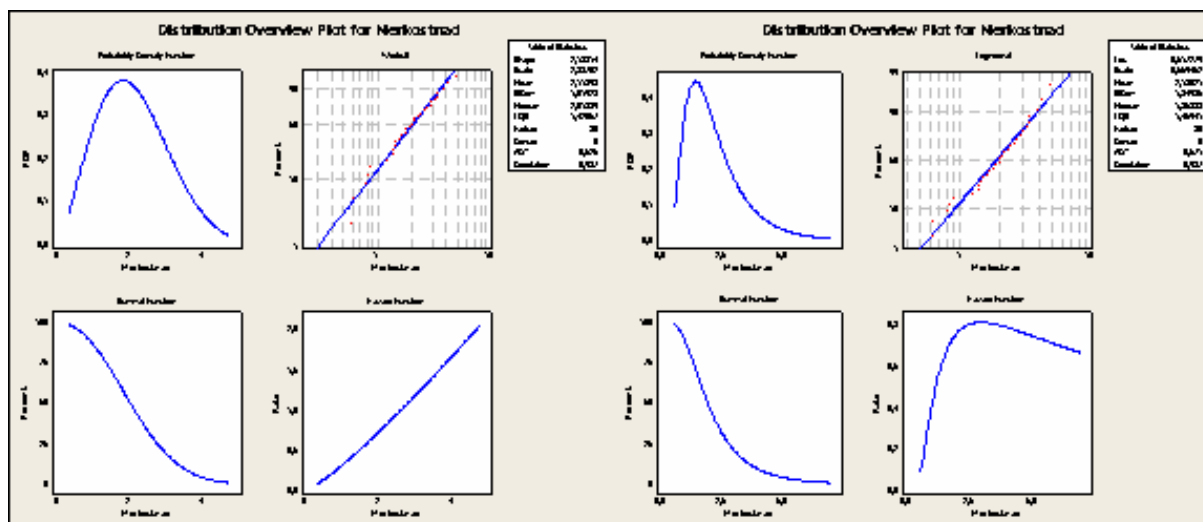
Eksempel: Anta at vi har registrert kostnader ved 30 likeverdige aktiviteter, som antas å innebære samme forventet merkostnad, og samme risiko for avvik fra forventningen. I statistiske termer antas de 30 registreringene å være uavhengige observasjoner generert fra samme sannsynlighetsfordeling. En sammenfattende statistisk analyse av dataene er gitt i tablået nedenfor, med grafikk, beskrivende mål og ulike andre beregninger.



Histogrammet antyder en fordeling omkring et gjennomsnitt på i overkant av 2. Av de beskrivende mål i kolonnen til høyre finner vi gjennomsnittet (mean) lik 2.1238 og (empirisk) standardavvik (StDev) lik 1.0752, som kan tas som estimater (anslag) for forventning og standardavvik i den teoretiske fordelingen. I figuren er tegnet inn en normalfordeling med forventning og standardavvik lik anslaget, og vi ser at det gir ganske god tilpasning til histogrammet, selv om dette nokså klart antyder en høyreskjev fordeling. Beregningene omfatter også en formell test for normalitet, som imidlertid innebærer at normalitet ikke kan forkastes. Dette er basert på den såkalte p-verdien, som her angir sannsynligheten for å observere minst like god tilpasning til normalitet i våre data, dersom normalitet er den ”korrekte” fordeling. Typisk vil heller ikke en valgt skjev fordelingstype bli forkastet, og man vil muligens foretrekke en slik isteden, for ikke å undervurdere risikoen for store merkostnader. Tablået omfatter for øvrig muligheter for å vurdere hvor usikker fastlegging av forventning og standardavvik i den teoretiske fordeling er (ved såkalte konfidensintervaller). Det er denne usikkerhet som kommer på toppen av den usikkerhet som ligger i risikoen for

avvik fra forventningen, dersom denne er kjent. Vi ser at selv med ganske mye erfaringsmateriale (hele 30 observasjoner), vil en kunne være i stor uvisse om risikoen for ekstremavvik ved en ny observasjon. Dette er en erkjennelse som er god å ha med seg ved risikovurderingen av prosjekter, der tilgangen på sammenlignbare og relevante data ofte er langt mindre. I praksis brukes ofte ekstremverdiene i den observerte fordeling som rettesnor. Med lite data har vi ofte ikke observert ekstremene, og derfor vil en lett undervurdere risikoen. Det er i slik sammenheng vi trenger teoretisk fordeling som rettesnor.

La oss så se nærmere på de nevnte skjeve fordelinger: triangulær, lognormal, Gamma og Weibull-fordeling. Den triangulære fordeling er forholdsvis enkel i praksis. En trenger å spesifisere tre verdier: den lavest mulige, den høyest mulige og den mest sannsynlige. Dersom data er tilgjengelig, som i eksemplet ovenfor, kan en lese rimelige verdier ut fra histogrammet. Sannsynligheter for store kostnader kan da finnes ved arealberegning i den etablerte trekant, og vi går ikke nærmere inn på dette. De tre andre fordelingstypene gir alle brukbar tilpassing til dataene i eksemplet, og det er hensiktsmessig å kjenne til noen egenskaper ved disse fordelingene. Det er spesielt viktig dersom man ikke har data i det hele tatt, og vil velge fordeling rent subjektivt. Med de gitte data har vi følgende tablåer med oversikt ved tilpassing av hhv. Weibull- og lognormal fordeling til dataene, mens Gamma-fordeling mangler.<sup>31</sup> For hver av de to fordelingene gis fire figurer og en boks med beregninger. De fire figurene er i rekkefølge (med urviseren): Anslått fordeling, et såkalt sannsynlighetsplott, sannsynligheter for overskridelse av gitte nivåer ("survival") og den såkalte hazardfunksjonen..



Vi har god tilpassing til data dersom punktene i sannsynlighetsplottet ligger noenlunde på rett line. Det er tilfelle både for Weibull-fordelingen (til venstre) og den log-normale fordeling (til høyre). Hazard-funksjonen gir her uttrykk for sannsynligheten for at "merkostnaden vil stoppe

<sup>31</sup> Slik beregning og grafikk som fins i statistisk programvare som omfatter analyse av pålitelighet og risiko.



på dette nivå, gitt at den ikke har stoppet på lavere nivå”. For Weibull-fordelingen er denne funksjonen enten stigende, konstant eller avtagende. For den anslåtte Weibull-fordeling ser vi at den stiger med merkostnaden, mens for den log-normale fordeling stiger den først, for så å avta. Den første situasjonen svarer til at når det først går galt, går det fra galt til verre, mens det andre betyr at dersom det inntreffer merkostnader, vil disse øke tilbøyeligheten for ytterligere merkostnader, men ikke utover alle grenser. Det siste er trolig mest realistisk. For Gamma-fordelingen vil hazard-funksjonen være økende eller avtagende alt ettersom formparameteren er større enn eller mindre enn 1. I tabellen til høyre finner vi estimerte parameterverdier, som fastlegger fordelingen entydig. En kan da foreta risikoberegninger. Eksempelvis dersom vi har valgt log-normal fordeling med det estimerte nivå (location) lik 0.617278 og skala (scale) lik 0.569462, får vi bestemt den merkostnad som har 1% sjanse for å overskrides til 6.973, som også grovt kan leses ut av figuren nederst til venstre.

### ***Inverse Cumulative Distribution Function***

Lognormal with location = 0,617279 and scale = 0,569462

P( X <= x )                    x  
                   0,99        6,9730

En fastlagt fordeling kan også brukes til simuleringer, f.eks dersom det dreier seg om delkostnader som skal aggregeres, og vi ikke greier å finne uttrykk for totalkostnadene på annen måte. Dette kommer vi tilbake til senere. Simuleringer forankret i erfaringsdata, eller antakelser som kan gjøres gjenstand for diskusjon, kan også ofte være en velegnet framgangsmåte for å formidle den rolle tilfeldigheter spiller for et resultat. La oss her bare se hva vi fikk derom vi simulerte ti trekninger fra den lognormale fordeling med det anslåtte nivå og skala. Vi fikk følgende merkostnader (forenlig med erfaringen):

5.0984.7950.6601.5601.0012.9272.3851.3370.9074.635

Kommentar: Ved valg av fordeling må man veie grad av realisme og enkelhet mot hverandre. Med lite erfaringsdata kan nok dette by på problemer. Man kan da se på om de ulike valg leder til vesensforskjellig konklusjon. Hvis så ikke er tilfelle, er saken grei. I motsatt fall, bør man trekke sin konklusjon ut fra det mest konservative valg av fordeling. Hvis dette innebærer for stor risiko, har man avdekket et behov for ytterligere informasjon.

I beslutningssituasjoner med lite data, vil selvsagt estimer av kostnader og beregnede sannsynligheter være usikre. En mulighet er å foreta en sensitivitetsanalyse, se avsnittet om kvalitetssikring av datagrunnlaget. Ofte må resultatene fra dataanalysen holdes opp mot subjektive vurderinger, se de andre bidragene i dette kapitlet. I enkelte sammenhenger hører man: ”Lag ditt beste estimat av kostnadene, så multipliserer vi det med tre” (eventuelt  $\pi=3.14$ , så ser det mer vitenskapelig ut). Riktignok tar dataanalyse sikte på å redusere det uformelle, subjektive og skjulte ved vurdering av kostnader. På den annen side kan synlig subjektivitet noen ganger være et gode, som rommer mye erfaring. Vi skal komme tilbake til en utradisjonell dataanalyse av kostnader, som fanger opp erfaring i et senere avsnitt. Om en ikke kan forsvare en analyse basert på sparsomme data fullt ut, vil hensikten i mange tilfeller være oppnådd, dersom analysen har gitt bedre innsikt som grunnlag for en beslutning.

## 7.7 Aggregering av usikre kostnader

La oss så etablere en ramme for å diskutere aggregering av usikre kostnader. Kostnader knyttet til et prosjekt kan studeres på ulike aggregeringsnivåer. På det laveste nivå har vi kostnader knyttet til en enkeltkomponent og/eller aktivitet (begge i vid forstand). På det midlere aggregeringsnivå har vi kostnader ved en enhet sammensatt av flere komponenter og/eller en kombinasjon av aktiviteter. På det høyeste aggregeringsnivå har vi kostnaden ved hele prosjektet med alle sine innsatsfaktorer. På alle aggregeringsnivåer kan en tenke seg totalkostnaden  $X$  som en sum av delkostnader

$$X = X_1 + X_2 + \dots + X_n$$

der  $X_i$  er kostnaden ved komponent nr.i. Vi bruker her ordet komponent i mer generell forstand, og kan omfatte alle typer kostnader. Vi kan også tenke oss at summen er en aggregering av kostnader over  $n$  tidsperioder, slik at  $X$  er totalkostnad over  $n$  perioder og  $X_i$  er en totalkostnad i periode nr.i for  $i=1,2,\dots,n$ . I dette tilfellet kan det også være aktuelle å tenke seg at hver  $X_i$  er en neddiskontert kostnad.

Vi har følgende regneregler for forventning:

$$E(X) = E(X_1) + E(X_2) + \dots + E(X_n)$$

og for varians, så framtr komponentene i summen (her delkostnadene) er ukorrelerte

$$V(X) = V(X_1) + V(X_2) + \dots + V(X_n)$$

Er komponentene i summen korrelerte må en legge til to ganger summen av de såkalte kovariansene mellom hvert par av komponenter. I praksis er ofte kostnadskomponenter positivt korrelerte, som vil medføre at varianser til summen er større enn summen av variansene. Overses dette kan en risikere for optimistiske risikovurderinger knyttet til kostnadene.

En kovarians har samme dimensjon som varians, men kan likevel være vanskelig å forholde seg til. Den kan imidlertid uttrykkes som korrelasjonskoeffisienten multiplisert med de respektive standardavvik (kvadratroten av variansene). En korrelasjonskoeffisient er et tall mellom null og en, der null uttrykker ingen korrelasjon og en uttrykker perfekt korrelasjon, dvs. at den ene kostnaden er eksakt en lineær funksjon av den andre. Hvor store korrelasjoner det er realistisk å regne med i praksis er vanskelig å si noe generelt om, men det hadde vært fint om erfaringsdata om dette kunne gjøres tilgjengelig, slik at alle kunne dra nytte av det.

Eksempel: Vi har to delkostnader som kan være korrelerte. Da er

$$V(X) = V(X_1) + V(X_2) + 2 \cdot C(X_1, X_2)$$

der kovariansen  $C(X_1, X_2)$  kan beregnes som  $\text{Corr}_{12} \cdot SD_1 \cdot SD_2$ . Antar vi variansene (og dermed standardavvikene) like store, vil forholdet mellom den virkelige variansen og variansen ved antatt ukorrelertethet være  $1 + \text{Corr}_{12}$ . Dersom korrelasjonen er lik 0.25, vil variansen til totalkostnaden øke med 25% sett i forhold til tilfellet med ukorrelertethet (standardavviket øker med ca. 12%). Nå er muligens korrelasjon lik 0.25 i overkant at hva man vil kunne oppleve i praksis, men dersom totalkostnaden er sammensatt av mange positivt

korrelerte delkostnader, vil en økning i variansen i forhold til summen av variansene av denne størrelsesorden kunne forekomme for langt lavere korrelasjoner.

I noen tilfeller er hver  $X_i$  en sum av delkostnader, hvor det ikke er hensiktsmessig å bryte ned ytterligere. Vi kan ofte uttrykke en delkostnad  $X_i$  ved andre størrelser. Noen eksempler:

I noen tilfeller kan  $X_i$  være et produkt av kvantum og pris, dvs.

$$X_i = Q_i \cdot P_i$$

der  $Q_i$  er kvantum i antall enheter og  $P_i$  er pris per enhet. Dersom kostnadskomponenten er knyttet til arbeid, kan en naturlig oppsplitting være

$$X_i = Q_i \cdot P_i \cdot T_i$$

der  $Q_i$  nå betyr antall i arbeid med aktivitet nr.i, og  $P_i$  betyr lønnsutgift per person per måned og  $T_i$  er antall måneder den aktuelle aktivitet krever. Hvis vi isteden studerer aggregering av kostnader over  $n$  etterfølgende perioder, som muligens krever ulikt antall i arbeid i hver periode og usikker sluttidspunkt, kan vi skrive

$$X_i = Q_i \cdot P_i \cdot I_i$$

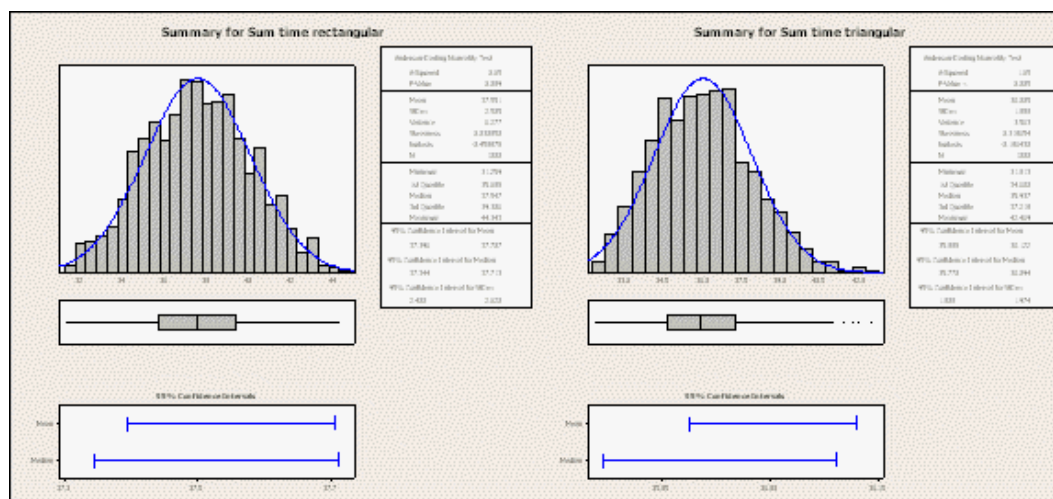
der  $Q_i$  nå betyr antall i arbeid i periode nr.i, såframt aktiviteten fortsatt pågår, og der  $I_i = 1$  dersom aktiviteten fortsatt pågår i periode nr.i og  $= 0$  dersom den avsluttet før periode nr.i.

I disse tre situasjonene tenker vi at faktorene i produktet selv er tilfeldige variable, hver med en tilhørende sannsynlighetsfordeling. I noen tilfeller er vi i stand til å gjøre antakelser om hver av disse faktorene som har konsekvenser for sannsynlighetsfordelingen til produktet. Dette er imidlertid vanskelig å studere analytisk. Imidlertid kan slike formler ofte være til stor hjelp som utgangspunkt for simuleringer. Et analytisk resultat bør nevnes: Forventningen til et produkt er lik produktet av forventningene, så framt faktorene er stokastisk uavhengige. I dagligdagse vendinger betyr uavhengighet her at kjennskap til en faktor ikke påvirker vår formening om noen av de øvrige. I praksis har vi ofte avhengighet, noe som overses og kan lede til urealistisk lave risikovurderinger.

Sentralt i simulering er en generator av tilfeldige reelle tall mellom 0 og 1, oppfattet som uavhengige trekninger fra den rektangulære fordeling over intervallet, kalt  $R[0,1]$ . Slike trekninger danner grunnlaget for trekninger fra andre fordelinger. For omfattende simuleringer og analyse av disse finnes egnet spesialprogramvare, men mulighetene er gode også med mer allmenn statistisk programvare, eller i enkelte regnearkprogrammer. Kjenner man sammenhengen mellom den rektangulære fordeling og andre fordelinger, er det ikke vanskelig å programmere enkle simuleringer i regneark. Enklest er å benytte den rektangulære fordeling over intervallet fra den (optimistiske) lavest mulig verdi  $a$  til den (pessimistiske) høyest mulige verdi  $b$ , dvs fordelingen kalt  $R[a, b]$ . Selv om det kan synes naivt å anta at alle verdier i mellom er like sannsynlige, kan dette likevel ofte gi nyttig innsikt i aggregering av risiko. Vi vil her nøye oss med et slikt enkelt eksempel, som isteden for kostnader dreier seg om tidsbruk.

Eksempel: Anta at et prosjekt skal gjennom tre faser, der neste trinn kan påbegynnes først når det foregående er ferdig. Anta, for enkelhets skyld, at for hvert trinn er den mest optimistiske tidsbruk 10, mens den mest pessimistiske er 15. Antas rektangulær fordeling  $R[10, 15]$ , betyr

det at forventet tidsbruk i hvert trinn er 12.5. Samlet tidsbruk i de tre trinn er summen av tre slike variable, som vi her antar er uavhengige av hverandre. For å avdekke fordelingen til summen simulerer tre uavhengige trekninger fra  $R[10, 15]$  fordelingen, hvoretter resultatet summeres. Dette gjentas 1000 ganger, med resultat som illustreres til venstre i figuren.



Vi ser at hyppighetsfordelingen til summen av de tre tidene er klokkeformet. For mange er det kanskje overraskende. Dette er imidlertid et resultat av den såkalte sentralgrensesetningen i statistisk teori. Grovt sagt sier denne at sannsynlighetsfordelingen til en sum av uavhengige tilfeldige variable (under visse tilleggsvilkår) nærmer seg normalfordelingen når antall ledd i summen øker. Nå var utgangsfordelingen rektangulær og er antall ledd i summen bare tre, men vi ser at effekten allerede er til stede. Anta at vi i utgangspunktet hadde valgt en triangulær med samme optimistiske og pessimistiske verdi 10 og 15, men med mest sannsynlig verdi 11. Dette tilsvarer en forventning lik 12, fordi forventningen i den triangulære fordeling faktisk er gjennomsnittet av de tre parameterverdiene 10, 11 og 15. Igjen foretas tre uavhengige simuleringer som summeres, og dette gjentas 1000 ganger. En oppsummering av resultatet er gitt på høyresiden i figuren. Vi ser den samme effekten, men siden utgangsfordelingen nå er skjev, er noe skjevhet fortsatt til stede i fordelingen til summen. Tilpasningen til normalfordelingen er likevel såpass god at grove vurderinger kan foretas på grunnlag av denne, selv om en hypotese om normalitet formelt må forkastes, når vi med mange observasjoner (1000) oppdager selv små avvik fra normalitet. Risikoen for at samlet tidsforbruk overstiger ulike nivåer kan leses direkte ut fra den empiriske fordeling i figuren. Eventuelt kan man bruke beregnet forventning og standardavvik og foreta vurderingen i lys av den tilpassede normalfordeling. I vårt eksempel har vi utregnet de teoretiske verdiene:

Fordeling	Forventning	Standardavvik	Forv. Sum	St.avvik Sum
Rektangulær	12.5	1.4433	37.5	2.5000
Triangulær	12.0	1.0540	36.0	1.8256

Med de antakelsene vi gjorde for den rektangulære fordeling, vil vi kunne slutte, i alle fall tilnærmet basert på normalfordelt sum med standardavvik 2.5, at sjansen for at samlet tid overstiger hhv 40.0, 42.5 og 45 er hhv. 15.9%, 2.3% og 0.14%

Eksemplet har lært oss noe viktig og svært nyttig: Dersom vi har med summer av uavhengige variable å gjøre, kan usikkerheten med fordel vurderes ut fra normalfordelt variasjon, uansett sannsynlighetsfordeling til leddene i summen. Det gjelder med visse forbehold også dersom leddene i summen ikke har den samme sannsynlighetsfordelingen. Dette innebærer at det er nok å kjenne eller anslå forventningen og standardavvik til summen. Den kan finnes ut fra forventning og standardavvik til de enkelte ledd i summen, og så bruke formlene innledningsvis i dette avsnittet. Vi trenger derfor regneformler for forventning og varians i de fordelinger som inngår. Imidlertid åpner det seg en annen mulighet: Det er nok å foreta direkte vurdering av forventning og standardavvik for hvert ledd i summen, vi trenger ikke anta noe om fordelingen. Dette fungerer rimelig bra unntatt i ekstreme tilfeller, som imidlertid nok kan forekomme.

Eksempel (fortsett): Anta at for de tre fasene i et prosjekt gjelder for forbruket av tid at

<i>Fase</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>Sum 1-3</i>
<i>Forventning</i>	<i>5</i>	<i>15</i>	<i>10</i>	<i>30</i>
<i>Standardavvik</i>	<i>1</i>	<i>2</i>	<i>2</i>	<i>3</i>

Standardavviket fremkommer som kvadratroten av variansen til summen av tidene for uavhengige trinn, som ifølge formelen først i dette avsnittet gir:  $V(X)=1^2 + 2^2 + 2^2=9$ .

## 7.8 Kostnader: Realisert kostnad vs. tidligfaseanslag

Vi skal se nærmere på å utnytte erfaring fra tidligfaseanslag til å predikere realisert kostnad.

La  $Y =$  Realisert kostnad

$X =$  Kostnad anslått i tidlig fase

som begge oppfattes som tilfeldige variable. For gitt  $X = x$  kan vi tenke oss at  $Y$  har sannsynlighetsfordeling  $f(y | x)$  med forventning  $E(Y | X = x)$ . Dersom man trenger mer realistiske kostnadsvurdering enn hva som er mulig å beregne seg fram til i tidligfasen, er det klart at kjennskap til ovennevnte fordeling kan være til stor hjelp. Dersom man kun trenger en prediksjon av realisert kostnad, er kjennskap til forventningen nok. Ønsker man innsikt i sjansene for store eller ekstreme avvik fra forventningen, trenger man å få vite mer om fordelingen. Noen fordelingstyper, som den normale, er entydig bestemt ved sin forventning og varians, slik at kunnskap om variansen i tillegg til forventningen vil kunne gi verdifull informasjon. I sin natur vil undervurdering av de realiserede kostnader være mer typisk enn overvurdering, slik at aktuelle fordelinger er skjeve (ikke-normale) med lang høyrehale. Eksempler på slike fordelinger er, som før nevnt lognormal-, Gamma- og Weibull-fordelinger. Alle disse har to parametere og er entydig bestemt ved sine respektive forventninger og varianser.

En mulighet er å betrakte  $Z = Y/X$ , dvs. forholdet mellom realisert kostnad og kostnaden anslått i tidligfasen, dvs.. Betrakket som tilfeldig variabel har  $Z$  sin sannsynlighetsfordeling som i prinsippet kan avledes av fordelingen til  $Y$  og  $X$ . Det kan imidlertid være enklere å gå direkte på fordelingen til  $Z$  med sin forventning og varians, som kan estimeres ut fra erfaringsdata fra prosjekter, ved gjennomsnittet av forholdene fra disse prosjektene.<sup>32</sup>

Med sammenhørende verdier av de to kostnadstallene for  $n$  tidligere prosjekter ( $X_i, Y_i$ ) beregner vi forholdet  $Z_i = Y_i / X_i$  for  $i=1,2,\dots,n$ . Dersom  $n$  ikke er for liten gir det mening å studere fordelingen i et histogram. Forventningen til slike forhold estimeres med gjennomsnittet  $\underline{Z}$ . Gitt et nytt prosjekt med anslått kostnad  $X_0$ . En prognose for den realiserede kostnaden er da  $\hat{Y}_0 = \underline{Z} \cdot X_0$ .

Fordelen med ovenstående er at en kan hente data fra svært ulike typer prosjekter. På grunnlag av data kan statistisk teori også gi mulighet til å beregne feilmarginer for prognosen. Dette forutsetter at prosjektene i en viss forstand er tilfeldige representanter fra en relevant populasjon av prosjekter, eksempelvis

- vi vet at de tidligere prosjekter og vårt/våre er strategisk undervurderte
- vi vet ikke noe om hvorvidt det foreligger strategisk undervurdering verken i de tidligere prosjekter eller i vårt/våre.

I det siste tilfellet vil selvsagt feilmarginer kunne bli betydelig større.

Merknad. Vi har her foreslått multiplikativ beregning istedenfor en additiv, fordi den stabile faktor typisk er relativ undervurdering, og prosjekter i tallmaterialet gjerne kan ha svært ulik størrelse. En mulig innvending er likevel at store prosjekter typisk er mer kompliserte og mer risikofylte også i relativ forstand. På den annen side blir det store absolutte beløp av det, som kan være en dempende faktor som ikke er til stede ved små prosjekter. Det er imidlertid også hevdet at når beløp blir riktig store, i milliardklassen, mister de fleste forankringen til det de vanligvis forholder seg til.

## 7.9 Realisert kostnad vs. observerbar indikator

La oss se på realistisk kostnadsvurdering ut fra de(t) aktuelle prosjekts egne karakteristika. Kilder til kostnadsavvik av mer eller mindre omfang kan være

- uteglemte kostnadskomponenter
- feilvurdering av kostnad og pris
- feilvurdering av tidsbruk
- uventede forhold (klimatekniske, geotekniske etc.)
- planendringer underveis
- teknologifremskritt underveis

---

<sup>32</sup> Problemstillingen over er bl.a. studert av Tore Knudsen (1976) i sin Ph.D avhandling fra University of California, Berkeley.

I praksis kan det være samspill mellom to eller flere av disse kildene til kostnadsavvik. Det er ikke lett å analysere hver av disse og slett ikke samspillet mellom dem.

For prosjekter i tidligfasen vil man sjelden være tjent med å gå i særlig detalj. At man heller ikke senere nødvendigvis er tjent med det, viser følgende eksempel fra norsk oljeindustri.

Eksempel: Kostnaden ved et pumpesystem i offshore virksomhet beregnes vanligvis ut fra prissetting av de enkelte komponenter, noe som var et omfattende og tidkrevende arbeid. Data fra 8 tidligere innkjøpte pumpesystemer forelå, bl.a. deres kostnad og vekt gitt i tabellen.

<i>Kostnad i 1000 kr</i>	47645	36276	4230	8474	18448	44157	5942	5400
<i>Vekt i tonn</i>	103,5	77,6	7,6	17,1	39,0	101,7	11,4	10,2
<i>Forholdet K/V</i>	460	467	556	495	473	434	521	529

Vi ser at de ulike pumpesystemene er svært ulike mht. størrelse målt med vekt, men at forholdet K/V er overraskende stabilt og varierer rundt et nivå nær 500. Utskrift av beregninger av beskrivende mål for forholdet foretatt med statistisk programvare gir følgende:

#### **Descriptive Statistics: Ratio**

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
Ratio	8	0	492,2	14,5	41,0	434,2	462,1	484,3	527,4	556,6

Vi ser at forholdet varierer fra 434,2 (minimum) til 556,6 (maksimum), med gjennomsnitt lik 492,2 og median lik 484,3. Median er gjennomsnittet av (de to) midterste når disse ordnet i stigende rekkefølge. Q1 og Q3 er såkalte kvartiler, der man regner at 25% av observasjonene vil ligge hhv. under og over. StDev (standardavviket) er 41,0. Det måler spredningen omkring gjennomsnittet i tallmaterialet, og kan tolkes som et sannsynlig avvik dersom vi bruker gjennomsnittet som prediksjon av forholdet for et nytt pumpesystem som ikke er med i tallmaterialet. SE Mean (standardavviket til gjennomsnittet) er 14,5. Det gir uttrykk for usikkerheten ved å bruke gjennomsnittet som anslag for forventet forhold mellom kostnad og vekt i en tenkt (uendelig) populasjon av pumpesystemer. Noen bruker feilaktig dette mål for risikoen ved prediksjon av ny observasjon, men det vil være å overse risikoen for avvik fra forventningen.

Ønsker en tilsvarende standardavvik for kostnaden for en gitt vekt, må man multiplisere ovennevnte tall med vekten. Eksempelvis blir standardavviket til kostnaden med et nytt system med vekt 50 anslått til  $50 \cdot 41,0 = 2050$ . Forøvrig kan det være nyttig å beregne forholdet mellom standardavvik og gjennomsnitt (den såkalte variasjonskoeffisienten), som her blir  $41,0/492,2=0,083$ , som grovt sagt betyr at inntil 8% avvik mellom prediksjon og virkelig kostnad ikke er usannsynlig.

Merknad. Ved antatt normalitet knyttes en 68% garanti til et utsagn om høyst ett standardavviks avvik fra gjennomsnittet, og 95% garanti for avvik på høyst to standardavvik fra gjennomsnittet. Slik garanti er ikke eksakt, hovedsakelig av to grunner: (i) Forholdene er neppe normalfordelte (noe som er vanskelig å avgjøre med få observasjoner), og (ii) garantier

forutsetter kjent (teoretisk) standardavvik, og vi har kun et estimat basert på 8 observasjoner. Statistisk teori kan imidlertid tilby mer presise sannsynlighetsutsagn, noe som vil føre for langt å komme inn på her.

Merk at for å kunne anvende dette i et nytt prosjekt i tidligfasen, trenger en i prinsippet vekten til systemet. Denne er kanskje ikke kjent eksakt, men man har størrelsesnivået. Dette medfører selvsagt økt usikkerhet, og statistisk teori kan i noen tilfeller bidra til også å kartlegge denne.

Vi stiller nå spørsmålet: Er ovenstående det beste vi kan få ut av dataene? Trolig ikke!

Ved ethvert pumpesystem vil det trolig være en grunnkostnad som ikke direkte har med størrelsen å gjøre. I stedet for å tenke seg at forholdet K/V variere rundt et forventet nivå, kan det være rimelig å anta følgende såkalte regresjonsmodell

$$K = a + b \cdot V + \text{tilfeldig avvik}$$

som uttrykker kostnaden som en lineær funksjon av vekten. Konstantleddet a gir uttrykk for før nevnte grunnkostnad, mens den såkalte regresjonskoeffisienten b gir uttrykk for forventet endring i kostnaden ved en enhets endring i vekten. Koeffisientene a og b kan vi anslå ut fra dataene, for eksempel ved den såkalte minste kvadraters metode, som ligger til grunn for beregningene i standard statistisk programvare. Dette gir

### **Regression Analysis: Kostnad versus Vekt**

The regression equation is Kostnad = 1000 + 442 Vekt

Predictor	Coef	SE Coef	T	P
Constant	1000,1	501,5	1,99	0,093
Vekt	441,650	8,304	53,19	0,000

S = 918,863    R-Sq = 99,8%    R-Sq(adj) = 99,8%

Unusual Observations

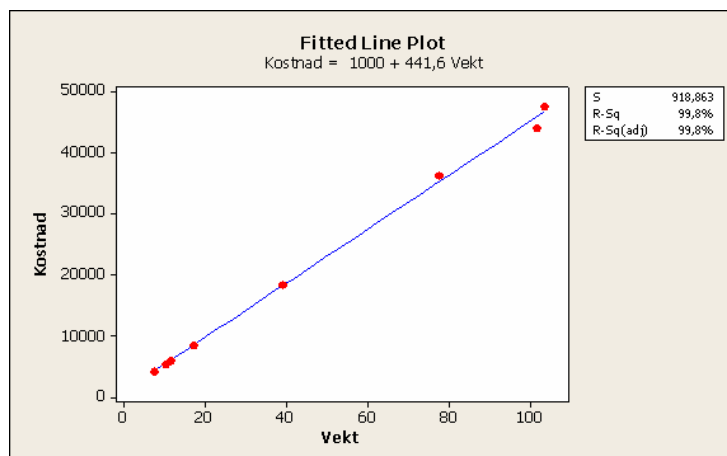
Obs	Vekt	Kostnad	Fit	SE Fit	Residual	St Resid
6	102	44157	45916	565	-1759	-2,43R

Utskriften gir oss prediksjonsformelen

$$K = 1000 + 441.65 \cdot V$$

Her blir standardavviket til feilen ved prediksjon av kostnaden ved nytt pumpeanlegg estimert til S=918, uansett hvilken vekt dette har. Forklaringsgraden er målt med R-Sq, ofte kalt determinasjonskoeffisienten. Vi ser det ganske oppsiktsvekkende at vekt forklarte over 99% av variasjonen i kostnader. Regresjonsutskriften gir også en rekke andre opplysninger, som vi ikke trenger komme inn på her, men er svært så relevante i andre sammenhenger. Følgende plott gir den informasjon vi trenger:





Tolkning av regresjonsutskrifter er basert på en modell som omfatter forutsetninger som er utelatt her, og som spiller liten rolle når sammenhengen er så klar som her. For andre problemer vil det kunne være nyttig å kjenne disse forutsetningene i detalj, og konsekvensen av brudd på disse. Standard lineær regresjonsmodellen antar: Stokastisk uavhengige feilledd med konstant varians, med tilleggsantakelsen om normalitet, dersom en ønsker eksakte sannsynlighetsvurderinger, se lærebøker i statistikk.

Dette eksemplet viser at det i noen situasjoner ikke er nødvendig å kjenne til detaljerte spesifikasjoner av en komponent i et prosjekt for å kunne gi et godt estimat på kostnaden. Det er nok å kjenne en vital karakteristikk ved elementet. Det var nokså overraskende at vekt kunne tjene som karakteristikk i dette tilfellet, noe som først ble klart ved dataanalysen. Man hadde også data om de enkelte pumpers kapasitet, og trodde på forhånd at dette var en bedre indikasjon på kostnaden. Det viste seg ikke å være tilfelle, dersom man skulle velge mellom de to karakteristikkene som prediksjonsgrunnlag. Man kan også spørre seg om kapasitet i tillegg til vekt kunne gi et bedre resultat, dvs. utføre en regresjon med to forklarende variable. Det viste seg at kapasitet ikke ga noen forbedring, dels fordi kapasitet var sterkt korrelert med vekt, og vekt allerede ga god forklaring. I andre situasjoner kan betydelige forbedringer oppnås ved å trekke inn flere forklarende variable, og utføre en multippel regresjon. For å vurdere flere aktuelle modeller opp mot hverandre brukes ofte forklaringsgraden R-Sq. Her må man utvise betydelig grad av forsiktighet. Med få data og flere forklarende variable kan en få svært høy forklaringsgrad for observerte data, uten at man likevel har noe godt prediksjonsgrunnlag. Man må alltid ha noe flere observasjoner enn forklarende variable, også for å kunne vurdere usikkerheten. Med sparsomt med data, nytter det derfor ikke å bringe inn mange forklarende variable.

I mange tilfeller forblir sammenhenger som i dette eksemplet ukjente og dermed uutnyttet. I den foreliggende situasjon var faktisk dataene ikke hentet frem med formål å gi bedre prediksjon av kostnader, men som ledd i å skaffe data til et kurs i statistisk metode!

## 7.10 Bayesiansk sannsynlighetsvurdering

### Bayesiansk klassifisering.

Lønnsomheten i et prosjekt avhenger av en underliggende tilstand  $Y$  som er ukjent, og først blir avdekket i en senere fase av prosjektet, hvis det overhodet blir videreført. Anta at antall mulige tilstander er endelig, og at man ønsker å planlegge ut fra hvilken som er mest sannsynlig. Vi kan ha a priori oppfatninger om  $Y$ , men ønsker å få ytterligere informasjon ved å observere en variabel  $X$  som i noen grad avhenger av  $Y$ . Vi vil betrakte  $Y$  og  $X$  som tilfeldige variable.

#### Eksempel Oljeboring

$X$  seismisk måling og  $Y$  kategorisk tilfeldig variabel med mulige "verdier":

$A$  = Borehull tomt eller ikke drivverdig

$B$  = Borehull har drivverdig olje

$C$  = Borehull har drivverdig gass

$E$  = Borehull har drivverdig olje og gass

For enkelhets skyld antar vi at  $X$  har et endelig antall mulige utfall, men perspektivet kan lett utvides til at  $X$  er en kontinuerlig variabel eller en vektor av flere variable.

Anta at sannsynlighetsfordelingen til observasjonen for gitt tilstand  $Y=y$  er gitt ved

$$P(X = x | Y = y)$$

Anta at denne er kjent for alle mulige tilstander  $x$  og  $y$ . Som prediksjon av "mest sannsynlig" tilstand kan en velge den  $y$  som for den observerte  $x$  gjør denne sannsynligheten størst mulig. Dette er iht. det såkalte "maksimum likelihood" prinsippet, som ikke gjør bruk av eventuell a priori oppfatning om sjansene for de ulike tilstandene. Dersom vi har uttrykt slik apriori oppfatning i form av en a priori sannsynlighetsfordeling

$$P(Y = y)$$

vil den såkalte a posteriori fordelingen være bestemt Bayes lov ved

$$P(Y = y | X = x) = \frac{P(Y = y) \cdot P(X = x | Y = y)}{P(X = x)}$$

der  $P(X = x)$  er bestemt ved summen av telleren over alle mulige  $y$ . Den  $y$  som maksimerer  $P(Y = y | X = x)$  kan finnes ved å maksimere telleren og dette gir

Bayes classifier: Maksimer  $P(Y = y) \cdot P(X = x | Y = y)$  mhp  $y$

I situasjoner der det er knyttet store økonomiske gevinster til bestemte utfall, er det ikke sikkert at det er lurt å innrette seg iht. det mest sannsynlige. Hvis  $Y$  er en numerisk variabel vil dens (aposteriori) forventning gitt  $X$  kunne være mer relevant.

### Bayes-nett.

Mange prosjekter i tidligfasen omfatter mange mulige hendelser, og man kan ha behov for en egnet terminologi og metodikk for å uttrykke og analysere hva man er sikker og usikker på.

Det fins flere muligheter for å kvantifisere og analysere usikkerhet: sannsynligheter, treverdi logikk (sant, usant, kanskje), fuzzy set logikk, "possability logic", "non- monotone reasoning".

Utfordringen er at det blir fort uoversiktlig dersom man ikke har en egnet systematikk for mange mulige hendelser. Bayes-nett er en mulighet som har fått betydelig gjennomslag i praksis de siste ti årene, og som baserer seg på sannsynligheter. Andre betegnelser på det samme er. "belief nets", "probabilty nets" og "causal nets".

Man forsøker å besvare spørsmål av typen:

- Hva er sjansen for at dette problem oppstår gitt disse symptomene?
- Er det noe rart ved denne observasjonen?
- Hva bør vi finne mer ut av når vi har observert dette?

Problemet er å kunne spesifisere simultane sannsynligheter på en systematisk og konsistent måte, slik at ikke de grunnleggende egenskaper og regnereglene for sannsynligheter brytes. Et Bayes-nett kan her være til stor hjelp. Det er de senere år utviklet programvare som er til hjelp.

### Informasjon og entropi

Anta at en tilfeldig variabel  $Y$  har  $m$  mulige utfall med tilhørende sannsynligheter gitt ved

<i>Utfall</i>	$y_1$	$y_2$	...	$y_m$
<i>Sannsynlighet</i>	$P(Y=y_1)$	$P(Y=y_2)$	...	$P(Y=y_m)$

Utfallene er typisk numeriske verdier, men kan også være kategorier, f.eks A, B, C...

Et nyttig mål knyttet til informasjon er entropien til (fordelingen til)  $Y$  er gitt ved

$$H(Y) = -\sum_{j=1}^m P(Y = y_j) \log P(Y = y_j)$$

Høy entropi betyr (nær) fordeling med like sannsynligheter (uniform fordeling), mens lav entropi betyr en fordeling med svært ulike sannsynligheter. Dette målet knyttes ofte til

overføring av informasjon i bits (0-1), og da brukes logaritmer med grunntall 2, men ellers spiller det ingen rolle om man isteden bruker naturlige logaritmer, dvs. med grunntall e.

### Eksempel

<i>Utfall</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>Sannsynlighet</i>	$1/4$	$1/4$	$1/4$	$1/4$

$$H(Y) = -\frac{1}{4} \log_2\left(\frac{1}{4}\right) - \frac{1}{4} \log_2\left(\frac{1}{4}\right) - \frac{1}{4} \log_2\left(\frac{1}{4}\right) - \frac{1}{4} \log_2\left(\frac{1}{4}\right) = 2$$

<i>Utfall</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
<i>Sannsynlighet</i>	$1/8$	$1/8$	$1/4$	$1/2$

$$H(Y) = -\frac{1}{8} \log_2\left(\frac{1}{8}\right) - \frac{1}{8} \log_2\left(\frac{1}{8}\right) - \frac{1}{4} \log_2\left(\frac{1}{4}\right) - \frac{1}{2} \log_2\left(\frac{1}{2}\right) = \frac{7}{4}$$

Vi ser at vi er mindre uviss på utfallet i det siste eksemplet, som medfører at entropien er mindre. Vi skjønner også at rekkefølgen av kategoriene er uten betydning.

Generelt gjelder at entropien er null dersom vi er sikker på utfallet og at maksimal entropi skjer når alle utfallene er like sannsynlige, med  $m$  utfall lik  $\log(m)$ .

I praksis kan dette brukes til å skille mellom situasjoner der vi er uviss på utfallet og situasjoner der bestemte utfall er mer sannsynlig. Ved ny informasjon endrer sannsynlighetene for de ulike utfallene, og dermed også entropien. Når entropien går ned reduseres uvissheten om hvilket utfall som inntreffer.<sup>33</sup>

I denne forbindelse trenger en begrepet betinget entropi. Anta at vi kan observere en tilfeldig variabel  $X$  med  $n$  mulige utfall som gir delvis informasjon om  $Y$ , slik at vi har en betinget sannsynlighetsfordeling for  $Y$  for hver mulig verdi  $x$  av  $X$ . For gitt  $X=x$  kan vi beregne (den betingede) entropien  $H(Y|X=x)$ , og informasjonsendringen ("information gain") kan da beregnes som

$$IG(Y|X=x) = H(Y) - H(Y|X=x)$$

<sup>33</sup> For øvrig kan entropien gis en konkret fortolkning, nemlig det minste antall bits man trenger i gjennomsnitt for å sende en strøm av symboler trukket uavhengig av hverandre fra den gitte fordeling

Et mål for entropien til  $Y$  betinget gitt  $X$ , sett før vi vet utfallet på  $X$ , er definert ved å veie med de respektive  $X$ -sannsynlighetene, dvs.

$$H(Y | X) = \sum_{i=1}^n H(Y | X = x_i) \cdot P(X = x_i)$$

og informasjonsendringen ved å observere  $X$  er da gitt ved

$$IG(Y | X) = H(Y) - H(Y | X)$$

### Eksempel

Anta at vi kan observere en indikator  $X$  med tre mulige verdier 1, 2, og 3, og at gitt hva vi observerte, så er de betingede sannsynlighetene for utfallene  $A$ ,  $B$ ,  $C$  og  $D$  som gitt i de tre linjene midt i tabellen. Da blir de betingede entropier (med grunntall 2) gitt  $X$ -utfallet som angitt i høyre kolonne i tabellen.

<i>Utfall</i>	<i>Sannsyn</i>	$Y=A$	$Y=B$	$Y=C$	$Y=D$	$H(Y X=x)$
$X=1$	$1/4$	$1/4$	$1/4$	$1/4$	$1/4$	2
$X=2$	$1/4$	$1/8$	$1/8$	$1/4$	$1/2$	$7/4=1.75$
$X=3$	$1/2$	$5/16$	$5/16$	$4/16$	$1/16$	1.2988

Antar vi at de tre mulige  $X$ -utfallene inntreffer med sannsynlighet  $1/4$ ,  $1/4$ ,  $1/2$  (gitt i første søyle i tabellen), blir den betingede entropi for  $Y$  gitt  $X$

$$H(Y | X) = \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 1.75 + \frac{1}{2} \cdot 1.2988 = 1.5869$$

Data i tabellen er i samsvar med en ubetinget fordeling for  $Y$ , der de fire utfallene er like sannsynlige, dvs.  $H(Y) = \log(4) = 2$ . Informasjonsgevinsten er dermed

$$IG(Y | X) = H(Y) - H(Y | X) = 2 - 1.5869 = 0.4131$$

Merk at hvis vi er så uheldig å observere  $X=1$ , får vi ingen ny informasjon om  $Y$ .

**Bayes oppdatering og ikke-informative a priori sannsynligheter**

La  $Y$  bety ukjent underliggende tilstand og  $X$  observerbar indikator for  $Y$ , der vi antar at de betingede sannsynligheter for ulike observasjonsverdier gitt underliggende tilstand er kjent

$$P(X = x | Y = y)$$

Dersom vi har a priori sannsynligheter  $P(Y = y)$  for de ulike mulige tilstander, er

a posteriori sannsynlighetene gitt at vi observerer  $X=x$  gitt ved Bayes lov slik

$$P(Y = y | X = x) = \frac{P(Y = y) \cdot P(X = x | Y = y)}{P(X = x)}$$

Man har ofte lite informasjon på forhånd, og da kan det virke rimelig å anta at alle mulige tilstander er like sannsynlige. Dette er blitt kalt "the principle of insufficient reason", og ble formulert allerede på 1700-tallet. Det vil også være forenlig med det såkalte "maximum entropy principle" omtalt ovenfor.

I noen situasjoner er man villig til å anta at a priori sannsynlighetene oppfyller en betingelse (utover at de summerer til en), men ut over denne er man ikke villig til å si noe. Eksempelvis kan man være villig til å si at  $Y$  har en sannsynlighetsfordeling av mulige verdier med en bestemt forventning. Maksimum entropi prinsippet innebærer da at velger en a priori fordeling som maksimerer entropien under den angitte bibetingelsen. Det kan vises at dette leder til et uttrykk av formen

$$P(Y = y_j) = \exp(q \cdot y_j) / \text{sum}(\exp(q \cdot y_j))$$

der  $\exp$  betyr eksponensialfunksjonen og  $q$  bestemmes slik at bibetingelsen er oppfylt.

Eksempel

Anta at  $Y$  kan anta verdiene 1, 2, og 3 dvs.  $y_j = j$  for  $j=1,2,3$ . Hvis vi ikke vet noe som helst på forhånd, kan vi sette de tre a priori sannsynlighetene like, dvs. lik  $1/3$ . Dersom vi tror at forventningen er  $f$ , og bruker prinsippet om maksimum entropi for å uttrykke vår uvitenhet ut over dette, får vi for  $j=1,2,3$  at

$$P(Y = y_j) = \exp(q \cdot j) / \text{sum}(q \cdot \exp(q \cdot j))$$

der  $\text{sum}(j P(Y=y_j)) = f$ . Dette leder til følgende annengradsligning til bestemmelse av  $q$ .

$(3-f)x^2 + (2-f)x + (1-f) = 0$ , der  $x = \exp(q)$ . Vi ser at dersom vi antar  $f = 2$ , blir ligningen  $x^2 - 1 = 0$  med løsningen  $x=1$ , dvs.  $q=0$ , som gir fordelingen med like sannsynligheter, som ventet. Dersom vi isteden antar  $f=2.5$ , og dermed en høyreskjev fordeling, blir ligningen  $x^2 - x - 3 = 0$  som har løsning  $x=(1+\sqrt{13})/2$ . Innsatt får vi fordelingen

$y$	1	2	3
$P(Y=y)$	0,1162	0,2676	0,6162

Med  $m$  mulige utfall får vi en polynomligning av grad  $m-1$ , som i praksis må løses numerisk. Dette kan skjules for bruker i programvare. Den samme generelle formel gjelder dersom  $Y$  har alle heltallene  $0,1,2,3, \dots$  som mulige verdier. Formelen kan modifiseres til å ta omsyn til annen bibetingelse enn for forventningen og modifiseres til å håndtere flere bibetingelser. Dette kan også skjules i programvare. Dersom  $Y$  er en kontinuerlig variabel, er teorien ikke fullt så enkel.

Vi har her nøydt oss med å presentere hovedideen ved maksimum entropi for å uttrykke vag kunnskap. Prosjekter i tidligfasen er nettopp preget av vag kunnskap, men noe kunnskap fins uttrykt i ulike termer. Maksimum entropi skulle således kunne ha et potensiale i denne sammenheng, men erfaringen i praksis er hittil begrenset.

## 7.11 Referanser

- Aven, T. (2003): Foundations of Risk Analysis. A Knowledge and Decision Oriented Approach, Wiley.
- Berger, J.O. (1980): Statistical Decision Theory and Bayesian Analysis. Springer Verlag.
- Carlin, D.F. & Louis T.A. (1996): Bayes and Empirical Bayes Methods for Data Analysis, Chapman & Hall.
- Cooper, D.F. & Chapman, C.B. (1987): Risk Analysis for Large Projects - Models, Methods & Cases, Wiley.
- Garvey, P.R. (1999): Probability Methods for Cost Uncertainty Analysis: A System Engineering Perspective, Marcel Dekker.
- Huang, K.T., Lee, W.L. & Wang, R.Y. (1998): Quality Information and Knowledge, Prentice Hall.
- Jensen, F. (1996): An introduction to Bayesian Networks, Springer Verlag .
- Vose, D. (1996): Quantitative Risk Analysis: A Guide to Monte Carlo Simulation Modelling, Wiley.
- Wang, R. Ziad, M. & Lee, Y. (2001): Data Quality, Kluwer Academic Publishers.
- Ubøe, J. & Jørgensen, K. (2004): Statistikk for økonomifag, Gyldendal.





# Del 3

## Kvalitetssikring av informasjon



## 8 KVALITETSSIKRING AV INFORMASJON I PRAKSIS

OLE JONNY KLAKEGG\*

### 8.1 Innledning

Kapitlet handler om kvalitetssikring av den informasjonen som bygges inn i planer og beslutningsunderlag. Det kan sies å drøfte på tre ulike nivå (eller i tre ulike perspektiv):

- Sikre at planlegger/rådgiver har tilgang til rett informasjon/data i sine beregninger.
- Sikre at prosjektet har komplett og riktig beslutningsunderlag
- Sikre at virksomheten har gyldig og pålitelig informasjon om tiltaket/samfunnsnyttene i livsløpsperspektivet for viktige beslutninger.

Problemstillingen har innvirkning på arbeidsmåten både for planleggerne som først utarbeider de dokumentene som beslutningen skal bygges på, og for kvalitetssikrerne som gjennomgår et allerede utarbeidet dokument.

Hva betyr kvalitetssikring i denne sammenhengen? Det betyr å sikre at det dokumentet som er eller blir utarbeidet inneholder riktig informasjon og korrekte resultater og fremstillinger, og at de er fullstendige og konsistente. Med konsistens menes at dokumentene ikke inneholder motsigelser og representerer en sammenhengende logisk sekvens gjennom hele og henger logisk sammen med overordnede mål eller beslutninger. I dette kapitlet er temaet begrenset til å omfatte sikring av riktig informasjon<sup>34</sup>. Den bakenforliggende problemstillingen handler om hvordan en kan gå fram i praksis for å sikre at beslutninger blir tatt på grunnlag av informasjon som er både gyldig og pålitelig.

Utgangspunktet for denne rapporten er drøfting av beslutningsunderlaget i situasjoner med tynt informasjonsgrunnlag. Altså når informasjonstilgangen er begrenset. Dette er også utgangspunktet for dette kapitlet, men det karakteristiske med drøftingen her er at det faktisk kreves et informasjonsunderlag, uansett hvor dårlig det måtte være. Derfor er oppgaven til planleggere, kvalitetssikrere og forskere å gjøre det beste ut av det som faktisk finnes. Med andre ord; hvordan unngå at et dårlig utgangspunkt blir dårligere enn nødvendig?

For å gi leseren et inntrykk av hva som er ment med kvalitet i dette kapitlet, tas med et par definisjoner, noe forkortet (Langlo et al, 2006):

- Datakvalitet består av fire elementer; 1) indre datakvalitet (troverdighet, nøyaktighet, objektivitet, rykte), 2) kontekstuell datakvalitet (forståelse av hvordan data er samlet, behandlet og presentert

---

\* Ole Jonny Klakegg er forskningssjef i Concept-programmet ved NTNU.

<sup>34</sup> I Concept-rapport nr 12 Usikkerhetsanalyse - Metoder (Austeng m. fl., 2006) finnes et kapittel som tar for seg kvalitetssikring av hele usikkerhetsanalysen.

for å kunne tolke/bruke data), 3) representasjonell kvalitet (format og mening i data), 4) tilgjengelighet (lett å komme til i fremtidig bruk).

- Informasjonskvalitet er et mer subjektivt mål på om nytte, objektivitet og integritet er ivaretatt i den informasjonen som er samlet/tolket ut av data.

For å være et så godt grunnlag som mulig må både data og informasjon være oppdatert, komplett, velegnet for formålet etc.

Første del omhandler ulike rammeverk for kvalitetssikring av underlag for beslutninger, slik de er introdusert i et par utvalgte land og store private bedrifter. Andre del identifiserer fallgruver som er aktuelle trusler mot god kvalitetssikring i praksis. Den tar for seg erfaringer med kvalitetssikring av beslutningsgrunnlaget slik det faktisk er utført og bør utføres i store statlige investeringsprosjekt. Det er i hovedsak basert seg på dokumentert praksis fra utførte kvalitetssikringer i perioden 2000-2006. Målet er å fastslå hvordan kvalitetssikring av informasjon faktisk er utført i store offentlige prosjekt i denne perioden, og spesielt innenfor ordningen med ekstern kvalitetssikring. Hensikten er å kunne lære av det.

Videre benyttes innspill fra litteraturen som supplement om beste praksis i dag og i fremtiden. Tredje del viser elementer av hvordan en beste praksis på dette området vil kunne se ut i dag. Også her hentes eksempler fra KS-rapporter. I denne delen er det vist gode prinsipper fra flere fagfelt og andre sammenhenger for å supplere bildet. Til slutt tas med en mer visjonær del som beskriver en løsning som vil forbedre evnen til å håndtere de fleste av fallgruvene presentert i andre del, samt tilføre en rekke andre fordeler. Løsningen er teknisk gjennomførbar allerede i dag.

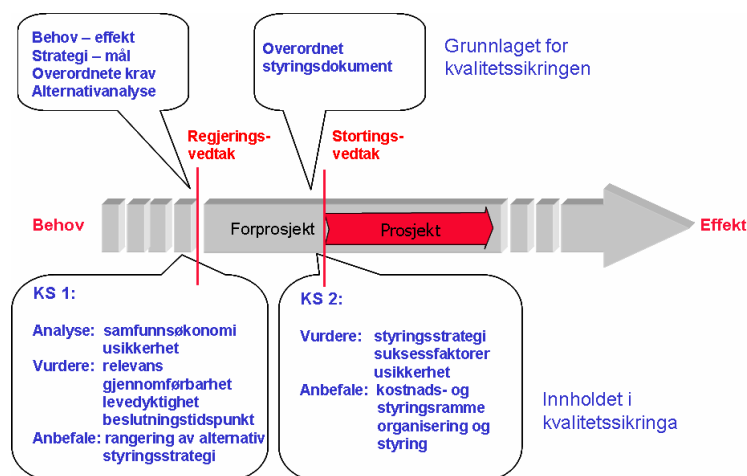
## 8.2 Rammeverk for kvalitetssikring av informasjon

Denne delen oppsummerer og karakteriserer 4 ulike rammeverk for (kvalitetssikring av) planlegging, beslutning og gjennomføring av investeringsprosjekt. Utvalget er ikke tilfeldig men konsentrerer seg om rammeverk som er velkjente, representerer det mange vil akseptere som nær beste praksis på sine områder, og er relevante for Concept-programmet. Utvalget er begrenset til to fra offentlig sektor og to fra privat sektor. Av disse er to norske og to utenlandske. I denne beskrivelsen er kun relevante hovedtrekk valgt å ta med. Nyanser og detaljer er det ikke plass til. Hensikten er å etablere en ramme for den videre drøftingen og vise at hvordan temaet kvalitetssikring av informasjon kan sees i en større sammenheng (forholdet til selve virksomheten). For en mer omfattende drøfting av slike rammeverk og anbefalinger om hvordan et slikt rammeverk bør formes henvises til Concept-rapport nr 3 Beslutningsunderlag og beslutninger i store statlige investeringsprosjekt (Holte et al., 2004).

### **Rammeverk A: Store statlige investeringsprosjekt i Norge**

Finansdepartementet introduserte i år 2000 ordningen med ekstern kvalitetssikring av store statlige investeringsprosjekt. Dette var et tiltak for å bedre på håndteringen av slike prosjekter i beslutningsfasen. Tiltaket var basert på erkjennelsene som et utvalg nedsatt tre år tidligere kom frem til, basert på gjennomgang av et utvalg av statlige investeringsprosjekt (Berg et al, 1999). Det overordnede formålet var at dette skal gi mer vellykkete prosjekter, reduserte kostnader for staten og mer nytte for hver krone. Fra 2005 er ordningen utvidet til å inkludere en ekstern kvalitetssikring av konseptvalg. Se figur 8-1 under. For mer informasjon om den

norske kvalitetssikringsordningen – se Concept-programmets hjemmesider på internett (Concept, 2006).



Figur 8-1 Illustrasjon av det norske rammeverket for kvalitetssikring av store statlige investeringsprosjekt.

De viktigste virkemidlene i den norske rammeverket er følgende:

- Ekstern gjennomgang og kontroll av grunnlaget for beslutninger og styring.
- Uavhengige analyser av samfunnsmessig lønnsomhet og investeringskostnad.
- Innføring av bruk av avsetninger innenfor kostnadsrammen.
- Felles begrepsbruk og krav til metodikk på tvers av sektor.
- Arena for tverrsektoriell erfaringsutveksling og etablering av omforente prinsipper.

En karakteristisk side ved den norske ordningen er at den griper fatt i det enkelte prosjektet, ikke først i statens rutiner. Den representerer dermed en slags nedenfra og opp angrepsmåte der erfaringer basert på de kvalitetssikrede prosjektene løftes opp til departementsnivå, der en blir enige om prinsipper som skal benyttes. Deretter får etatene utfordringen med å iverksette nødvendige tiltak for å imøtekomme kravene.

Et annet karakteristisk trekk er den konsekvente bruken av eksternt fokus for å sikre uavhengige vurderinger. Gjennom rammeavtaler med Finansdepartementet er utvalgte rådgivere fra privat sektor godkjent for å gjennomføre ekstern kvalitetssikring av statlige investeringsprosjekter utenom oljesektoren. Kontrollaspektet er dominerende i KS-rådgivernes arbeid og dermed et kjennetegn ved hele ordningen.

Ordnningen gjelder bare for prosjekt med forventet investeringskostnad over 500 mill. NOK, men prinsippene som utvikles forventes å ha innvirkning på hele statens portefølje av investeringsprosjekt.

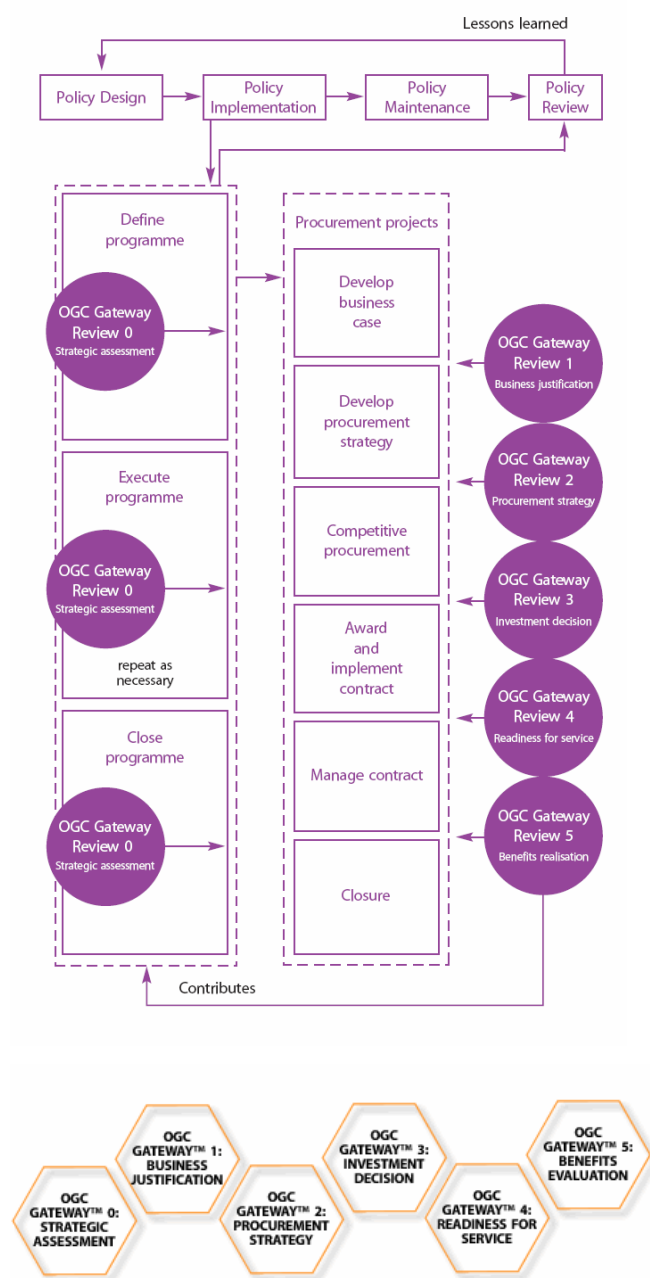
### **Rammeverk B: Offentlige investeringsprosjekt i Storbritannia**

Det Britiske finansdepartementet (Her Majesty's Treasury) etablerte på 1990-tallet en egen frittstående etat kalt Office of Government Commerce (OGC) for å støtte opp under offentlige innkjøp og gjennomføring av offentlige prosjekt<sup>35</sup>. Sitat fra OGC sin hjemmeside på internett: *"OGC's mission is to work with the public sector as a catalyst to achieve efficiency, value for money in commercial activities and improved success in the delivery of programmes and projects."* (OGC, 2006). Den overordnede målsettingen kan sies å være nokså parallell med tilsvarende for den norske kvalitetssikringsordningen. Bakgrunnen for initiativet er imidlertid nokså forskjellig og det politiske bakteppet annerledes enn i Norge.

---

<sup>35</sup> Militære anskaffelser sorterer i UK under National Audit Office (NAO).

## The wider context of the OGC Gateway™ Process



Figur 8-2 Illustrasjoner fra rammeverk for offentlige investeringsprosjekt i UK (Office of Government Commerce).

De viktigste virkemidlene:

- Etablering av et helhetlig og heldekkende system for policy-utvikling, planlegging, kontroll og gjennomføring av prosjekter. Elementene i systemet er mange og allsidige. De kan oppfattes som felles standarder, men kan kombineres og tilpasses den enkelte offentlige virksomhet og type prosjekt.

- Etablering av et internt organ i staten for å drive utvikling og oppfølging av systemet, opplæring, rådgivning og revisjon. Lokale "Centre of excellence" etablert rundt om i landet for tett oppfølging i ulike sektorer.
- Gateway-modell med klare krav til beslutningsunderlag og beslutningsprosess.
- Ekstern revisjon (innenfor staten) utført av spesielt kvalifiserte revisorer opplært av OGC.
- Utstrakt informasjonsarbeid og aktiv rådgivning overfor offentlige instanser i rollen som prosjekteier, innkjøper og prosjektledelse etc.

Karakteristisk ved det britiske opplegget er at det er en omfattende standardisering av metoder og systemer introdusert ovenfra, med utgangspunkt i et krav om effektivisering av offentlig sektor. Det kan sammenlignes med å innføre et felles kvalitetssikringssystem for hele "Staten AS". Hensikten er omtrent som den norske; mer verdi for pengene og frigjøring av offentlige ressurser til viktige oppgaver i det offentlige. Rammeverket er i Storbritannia eksplisitt knyttet til en fornying av styringen i offentlig sektor (med referanser tilbake til Thatcher-tiden i britisk politikk). "Supporting your business" er et slagord som synes å være flittig brukt i brosjyrer fra OGC. Dette tolkes som en antydning om tilnærming til privat sektors virkemiddelbruk og ideologi.

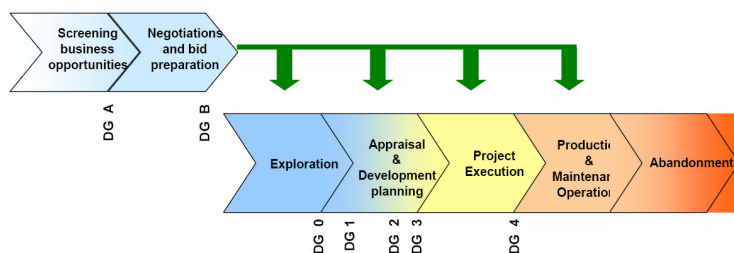
Et annet karakteristisk trekk er den gjennomgående fokuset på måling i alle aspekter av virksomheten. Kvantifisering av nøkkeltall og måling av prestasjoner på alle nivå og i alle sammenhenger er tydelig i alle elementer av systemene og informasjonen om rammeverket. Bedre målinger og nøkkeltall blir fremhevet som en drivende parameter for utviklingen.

Britiske myndigheter har brukt mye ressurser på utvikling av systemelementene i OGC's verktøykasse. Elementer som OGC Gateway<sup>TM</sup> Process og OGC Prince2<sup>TM</sup> er velkjente og i aktiv bruk også internasjonalt og kopieres i dag av andre land, ikke minst de nye søkerlandene til EU.

### **Rammeverk C: Investeringsprosjekt i Statoil**

Norges største oljeselskap, Statoil, har en velutviklet modell, eller rammeverk, for hvordan de kvalitetssikrer og tar sine beslutninger om nye investeringsprosjekt. Den er illustrert i figur 8-3.





Figur 8-3 Illustrasjon av Statoils rammeverk for prosjektbeslutninger

#### Hovedvirkemidlene i modellen:

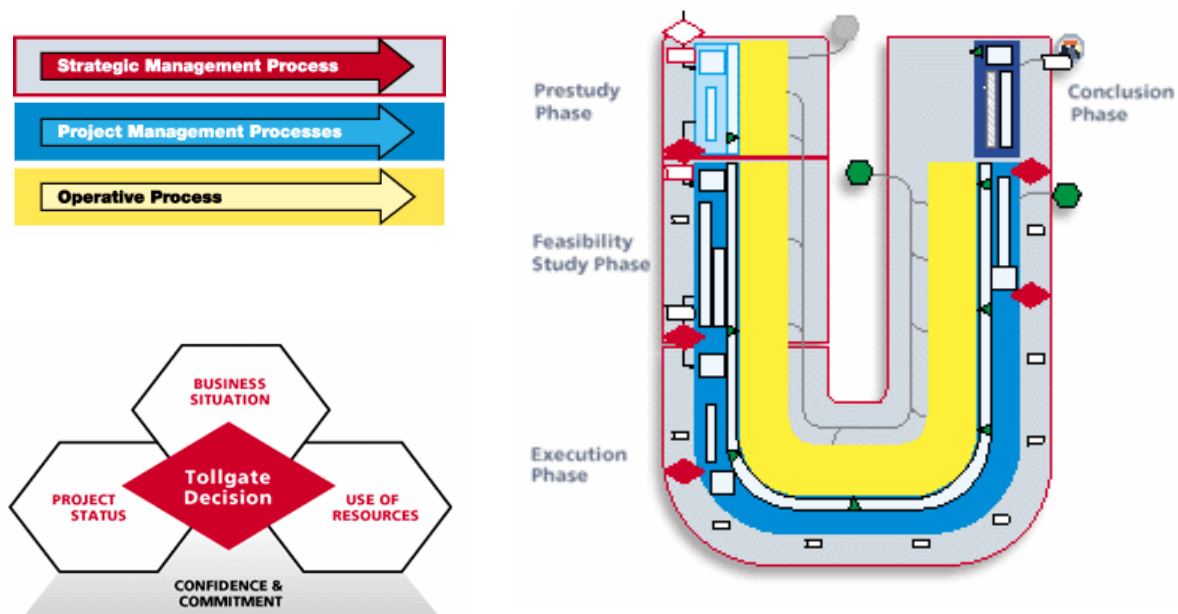
- Klare beslutningspunkt (Decisiongates - DG) med definerte krav til dokumentasjon, gatekeeper og beslutningsprosess.
- Sterk forankring i konsernstrategi og forretningsmessig drift.
- Bruk av arenaer og nettverkstenking internt for å sikre bred belysning av problemstillingene og utnyttelse av samlet erfaring og kunnskap.

Modellen til Statoil fokuserer på beslutninger og det er lagt stor vekt på grundig og tverrfaglig evaluering av beslutningsunderlaget på hvert enkelt steg i prosessen. Arenabegrepet er sentralt. Underlaget for beslutning legges frem på Arena for prosjektutvikling der det gjennomgås grundig og drøftes bredt. Sammen med prosjektforslaget legges fram uavhengig prosjektevaluering utført av andre Statoil-ansatte eller innleide eksperter som ikke har utført planleggingen. Det utføres evalueringer med spesielle fokus (kommersielle eller andre) etter behov.

Utredningsarbeidet og beslutningene blir av Statoil selv karakterisert som en integrert, multidisiplin prosess med fokus på hvor godt prosjektet passer inn i konsernets strategi, på forretningsmessig risiko (økonomisk, kommersiell) og helse, miljø og sikkerhet. Prosjektgjennomføring er naturligvis også et viktig tema. Strategisk viktige initiativ gis spesiell oppmerksomhet fra konsernledelsen (Haugane, 2005).

#### Rammeverk D: Investeringsprosjekt i Ericsson

Et av de mest kjente og anerkjente rammeverk for prosjektgjennomføring er Ericsson sin modell PROPS. Den ble utviklet som en intern prosjektmodell for Ericsson-konsernet, men er i dag markedsført som et eget produkt med verdensomspennende bruk i ulike industrier og bedrifter. PROPS kan trolig regnes som den (eller en av de) første og mest grunnleggende modellen(e) og inspirasjon til en rekke andre rammeverk og modeller i prosjektverdenen. For mer info – se for eksempel internettsiden PROPS Online (2003).



Figur 8-4 Illustrasjoner fra Ericsson sitt rammeverk for prosjektgjennomføring; PROPS. Hovedvirkemidlene i PROPS-modellen:

- Klare beslutningspunkt (Tollgates) med definerte krav til dokumentasjon, gatekeeper og beslutningsprosess.
- Forankring i forretningsmessig vurdering (business situation).
- Bruk av ulike perspektiver (Human perspective, Business perspective, Project organisation perspective and the Single project flow perspective) for helhetsvurdering og balansering av vurderingene.
- Sjekklistene, maler og beskrivelser for alle nødvendige roller, funksjoner og oppgaver i et prosjekt.
- Skalerbar modell, det vil si at den kan tilpasses alle typer og størrelser av prosjekt ved å plukke fra elementene i helheten.

PROPS modellen er en av de mest anerkjent og brukte prosjektmodellene. Den inkluderer det meste som trengs av metodikk og prosjektkunnskap for de fleste situasjoner. Den er en administrativ overbygning til den tekniske utviklingen og gjennomføringen av prosjektet. Som system er den heldekkende for de mer operative sidene ved beslutninger og gjennomføring av prosjekt. Som rammeverk er de mer overordnede aspektene (project governance) i mindre grad dekket, eller i alle fall ikke så godt synliggjort.

### Betydningen av et rammeverk for kvalitetssikring av informasjon

Hva er så fellesnevnerne og vesentlige forskjeller mellom disse rammeverkene? Som beskrivelsen over viser så er det mer som ligner enn som skiller de ulike rammeverkene. Bakgrunnen til de ulike rammeverkene er ulik og prioriteringene varierer en del. Det er likevel en kjerne av felles tenking i bunnen som får litt ulike uttrykk når de kommer til overflaten. Noen hovedmomenter:

Alle rammeverkene inneholder en struktur med definerte beslutningspunkter. Det settes krav til beslutningsunderlaget i disse punktene.

Alle rammeverkene har et innslag av kontroll av dokumentasjon (informasjonen) forut for beslutning. Kontrollen har karakter av tredjepartskontroll og utføres av eksperter enten fra andre deler av virksomheten eller eksternt.

Alle rammeverkene inkluderer klare elementer av systemtenking. Det norske rammeverket for store statlige investeringsprosjekt skiller seg ut med å ha mindre fokus på operative/praktiske forhold og et mer tydelig fokus på det overordnede.

Den norske rammeverket for store statlige investeringsprosjekt skiller seg fra de øvrige med å ta utgangspunkt i det enkelte investeringstiltaket, i stedet for i systemet slik de andre rammeverkene gjør. Det har likevel den samme målsettingen om å forbedre systematikken i virksomheten.

I forhold til temaet i denne rapporten kan virkemidlene i det norske rammeverket karakteriseres slik: Rammeverket definerer prosessen (hva skal kontrolleres, når og av hvem) og hvilke krav som skal stilles til beslutningsunderlaget. Rammeverket må forankres i virksomheten slik at kvalitetssikrerne får mandat til å gjøre jobben sin og hjelp til å forstå hvilke prioriteringer som gjelder, men uten en detaljert henvisning for hvordan den skal gjøres.

### 8.3 Erfaringer fra store offentlige investeringsprosjekt – fallgruver

Denne delen er basert på gjennomgang av kvalitetssikringsrapporter fra ekstern kvalitetssikring av store statlige investeringsprosjekt (KS2) i perioden 2000 – 2006. I disse rapportene er fallgruver identifisert ved gjennomgang av teksten. Fallgruvene er sortert tematisk og nummerert fortløpende for å finne dem igjen i diskusjonene senere i kapitlet.

#### A. Manglende kvalitetssikring

(1) Den største feilen av alle er selvsagt å ikke kvalitetssikre informasjonen i det hele tatt. Dette er åpenbart en farlig feilkilde. Imidlertid er det grunn til å tro at det er et fenomen som forekommer. Hovedårsaken er trolig bekvemmelighetshensyn. Det er mindre arbeidskrevende å bruke den informasjonen som er tilgjengelig enn å skaffe seg bedre informasjon, eller eventuelt dobbeltsjekke at den som er tilgjengelig er riktig. Andre årsaker kan være manglende vilje til å kvalitetssikre eller manglende bevissthet.

(2) Noen vil kanskje også trekke fram at manglende kvalitetssikring av informasjonen er grunnet i god tillit til den som har skaffet den opprinnelige informasjonen. Siden den kommer fra en ”sikker” kilde blir den ikke etterprøvd. Dette er en fallgruve. Det må ikke tolkes som mistillit å forlange informasjonen kvalitetssikret. Alle gjør feil fra tid til annen. Kvalitetssikring av informasjon er et karakteristisk trekk ved kvalitetsbevisste organisasjoner. Dess viktigere oppgave, dess mer selvsagt at informasjonen blir kvalitetssikret. Hva sier det om virksomheten til de som ikke ønsker å prioritere kvalitetssikring av informasjon? Når oppgaven nettopp går ut på å kvalitetssikre det andre har gjort, er det å ikke kvalitetssikre grunnlaget nærmest utilgivelig. Vi skal imidlertid la dette ligge og velger å ikke gi eksempler.

(3) Å velge å ikke bruke en relevant kilde som faktisk er tilgjengelig kan også forekomme. Her eksemplifisert med sitat fra en kvalitetssikringsrapport:

*”[Prosjektet] har som en del av saksframlegget for Fylkesting og Storting utarbeidet en egen usikkerhetsanalyse for [prosjektet]. Det presiseres at vårt arbeid skal være en uavhengig vurdering av beslutningsunderlaget. Dette er avklart i innledende møter. For [KS-rådgiver] har dette bl.a. hatt den konsekvens at vi ikke har benyttet oss av deltakere fra [prosjektet] til å kvalifisere de avdekkede usikkerhetsfaktorer.”*

Det er noe uklart hva formuleringen ”kvalifisere de avdekkede usikkerhetsfaktorer” innebærer, men det kan tolkes som en her har valgt å ikke benytte seg av den kunnskapen som fantes i prosjektet til å underbygge egne selvstendige vurderinger. I den grad dette er en riktig tolkning illustrerer det en fallgruve. Å lytte til andres velkvalifiserte vurderinger truer ikke egen selvstendighet. Egen selvstendighet blir bare truet dersom en ikke kritisk vurderer godheten i den informasjonsbasen vurderingene bygges på, og de konklusjoner som dras. I den tidligste perioden med ekstern kvalitetssikring ble selvstendighet i vurderinger drøftet en del. I noen tilfeller gikk det i retning av å ekskludere eksterne kilder (alternative syn og vinklinger):

*”I samråd med [Etaten] er det besluttet å ikke benytte andre eksterne konsulenter til kvalitetssikring av grunnkalkylen. Dette begrunnes med mulige habilitetsproblemer for konsulentene i tilbudsfasen.”*

Etter hvert falt dette på plass og det ble naturlig å involvere prosjektets fagfolk i diskusjoner for å bygge opp under egen selvstendig vurdering.

(4) En fallgrube beslektet med de over er å ikke ha (ta seg) tid til å kvalitetssikre. Årsaker kan være tidsmangel (korte frister) eller manglende ressurser til å utføre oppgaven (mangel på folk). Grunnleggende er det likevel et spørsmål om prioritering. Den som ønsker å kvalitetssikre planlegger det inn. For å gjøre det effektivt, er tilgangen til data avgjørende.

### **B. Å skaffe god informasjon**

(5) Mange har ikke sikret seg effektiv tilgang på god informasjon/fakta (data). Fordi god informasjon er vanskelig å få fatt i, blir det ofte heller akseptert å benytte den informasjonen som er tilgjengelig, enn å bruke de nødvendige timene til å innhente korrekte, normaliserte<sup>36</sup> data som kan tilpasses oppgaven (behovet). Dette vil vi komme tilbake til senere.

(6) I praksis har det skjedd altfor ofte at planlegging bygger på antakelser og informasjon som ikke er kvalitetssikret. Noen ganger går det bra, noen ganger blir det avslørt, men ofte ikke. Årsakene kan være alle de over nevnte eller kombinasjoner av dem. Det må konstateres at denne fallgruben også har gjort seg gjeldende innenfor kvalitetssikringsordningen. Konkret knytter dette seg til at basiskalkylen i kostnadsanalysene i mange tilfeller ikke har blitt kontrollert eller etterprøvd av kvalitetssikrer. Sågar utviklet dette seg til en normal prosedyre (og akseptert som sådan) i en periode. Av bekvemmelighetshensyn ble det akseptert å bygge på basiskalkylen fordi ”det er urealistisk at kvalitetssikrer kan etablere en like god kalkyle på så kort tid som er tilgjengelig”. Senere har kalkylene fått en langt mer fremtredende posisjon i kvalitetssikringene.

### **C. Manglende evne til å vurdere og behandle data**

(7) Mangel på fagkompetanse er klart en fallgrube i forbindelse med kvalitetssikring av informasjon. I den tidlige perioden av KS-ordningen ble diskusjonen ført inn på spørsmålet om hvilken kompetanse kvalitetssikringsrådgiver skulle ha for å kunne gjøre egne, selvstendige vurderinger. Dette henger også sammen med kvalitetssikring av informasjon i praksis fordi det er et spørsmål om hvilken kompetanse enkeltperson har til å gjøre egne vurderinger av informasjonens godhet, og hvilken tilgang vedkommende har til god informasjon.

I starten fra 2000 og fremover gikk diskusjonen høyt om hvordan selvstendige vurderinger skulle sikres og behovet for fagkompetanse ivaretas. Senere ble det mer og mer stille rundt dette spørsmålet. Dette har trolig sammenheng med en kombinasjon av flere faktorer: På den ene siden ble det mer akseptert at det er mulig å kompensere for egen dyp innsikt i et fagfelt med å innhente kvalifiserte vurderinger fra andre, uavhengige kilder. Nettverk og

---

<sup>36</sup> Med ”normaliserte data” menes at rådata som er hentet fra tidligere tilfeller er korrigert for spesielle karakteristiske dimensjoner og avvikende forutsetninger, slik at de er velegnet for gjenbruk i senere planlegging.

systematikk/metodekompetanse hadde større fokus enn behovet for egen fagkompetanse. I tillegg ble det akseptert at prosjektets fagfolk kunne benyttes aktivt i underbyggingen av egne selvstendige vurderinger. En kan selvsagt ikke se vekk fra at det også hadde med å gjøre at konsulentene ble mer profesjonelle etter hvert som de fikk trening og var trygge i sin rolle, samt at aksepten steg hos de som ble kvalitetssikret. Spørsmålet om egen fagkompetanse er nå redusert til et spørsmål om effektiv kommunikasjon.

(8) Feil valg av forutsetninger og antakelser. Hva gjør vi når vi ikke har svarene? Da velger vi en forutsetning. Planlegging eller analyser får aldri lov å stoppe opp fordi det ikke er tilgang på sikker informasjon. En naturlig konsekvens er selvsagt å gjøre tydelig oppmerksom på at det er gjort usikre forutsetninger, det vil si å tilkjenne egen vurdering av kvaliteten på informasjonsunderlaget for egne vurderinger. Det var planlagt å benytte eksempler fra KS-rapporter her, men overraskende nok er ingen slike eksempler funnet.

En annen måte som er vanlig å forholde seg til denne typen problemstillinger på, er å velge forutsetninger som gjør det mulig å komme videre, selv om ikke alt er godt dokumentert. Å velge forutsetninger er både nødvendig, riktig og legitimt så lenge er tydelig og lar seg overprøve. Det er også en nødvendig teknikk for å gjøre problemstillingen i analysen håndterbar. Dette eksempelet illustrerer hvordan valg av forutsetninger blir benyttet:

*”Prosjektet fremstår som nøkternt og godt planlagt. Usikkerheten er liten. Forhold som angår forse majeure og politiske beslutninger som endrer premisene for prosjektet er ikke tatt hensyn til i kostnadsanalysene.”*

Fallgruven består i å velge urealistiske, ikke konsistente eller feilaktige forutsetninger for å komme videre. Disse blir lett stående i ettertid og har en tendens til å leve videre også etter at riktige informasjon er kommet fram. Til det siste eksempelet; er det realistisk at politiske beslutninger som ikke endrer premisene for prosjektet ikke vil forekomme? Tiden vil vise, men dette er neppe en god forutsetning å legge til grunn for en usikkerhetsanalyse. En (viktig?) kilde til usikkerhet er dermed eliminert uten videre. Forutsetningen sier bare at det ikke er tatt med i analysen, ikke at det ikke vil forekomme. Hvilket budskap satt leserne av rapporten igjen med når de studerte tallene i denne rapporten?

Det har (foreløpig) ikke lyktes å finne en KS-rapport som er eksplisitt kritisk til tilgangen på informasjon bak egne vurderinger. Derimot er det lett å finne eksempler der KS-rådgiver har påpekt tilsvarende fallgruver hos planleggerne (dette er en hovedoppgave i kvalitetssikringen):

*”Styringsdokumentet bør snarest oppdateres.... Det bør tilstrebes størst mulig innbyrdes konsistens mellom de forskjellige strukturene i prosjektet. Dette gjelder først og fremst strukturene knyttet til organisasjon, kontraktsinndeling, framdriftsstyring, estimat og økonomistyring.”*

*” [KS-rådgiver] konstaterer at det er usikkerhet knyttet til den teknologiske utviklingen og konsekvensene for [prosjektet]. Vi regner med at den teknologiske utvikling vil medføre endringer i prosjektet med økonomiske konsekvenser.”*

Overstående eksempel viser mangler ved underlaget. Et annet eksempel viser KS-rådgivers arbeid for å kvalitetssikre data i et tilfelle der vurderingen blir at prosjektets underlag er godt underbygget. Også slike er heldigvis lette å finne:

*"[KS-rådgiver] har i sitt kvalitetssikringsarbeid av prosjektet foretatt en gjennomgang og kontroll av prosjektets grunnkalkyle. Grunnkalkylen er sjekket i forhold til nøkkeltall for enbetspriser. Basert på denne gjennomgangen er [KS-rådgiver]s vurdering at prosjektets grunnkalkyle ligger på et fornuftig nivå. Vi velger derfor å legge prosjektets grunnkalkyle, slik den er presentert i kostnadsgjennomgangen (prosjektets ANSLAG rapport) til grunn for den videre vurderingen (se vedlegg 2)."*

#### D. God og dårlig kommunikasjon

Dårlig kommunikasjon er en anerkjent feilkilde og naturligvis en trussel mot kvaliteten på informasjonen vi benytter i planer, kalkyler og beslutningsunderlag. Å ta lett på kommunikasjonen er en fallgrube. Den erkjennelsen trenger neppe ytterligere underbygging. Det er mer et spørsmål om hvilke fallgruver som finnes som kan gi dårlig kommunikasjon og føre til dårlige data, og omvendt.

Et første spørsmål kan være om det er mulig å eliminere kommunikasjonsproblemet? Svaret er nei. Det kan ikke fjernes, men det kan reduseres. Å fjerne problemet forutsetter at det finnes en kilde som er 100 % dekkende og korrekt der alt er tilgjengelig. Slike kilder finnes ikke i dag, selv om enkelte personer/organisasjoner kan synes som de anser seg selv som allvitende. Å redusere problemet er mer realistisk. Det vil kunne oppnås gjennom å sikre seg tilgang til bedre datagrunnlag (sammenfallende fallgrube med (5) over) eller å sikre god dialog med de kildene som har slike data. Vi konsentrerer her om det siste.

(9) Et viktig spørsmål omhandler språk. Har vi et velegnet felles språk eller begrepsapparat som gjør det mulig å ha en god og effektiv dialog? Dette kan sees slik at en må sikre at det er et minimum av etablert felles språk, enten gjennom felles faglig bakgrunn (utdanning), gjennom vedtatte standarder i organisasjonen, eller omforente felles begreper. I forbindelse med KS-ordningen var mangel på felles språk et problem i starten. Det oppsto "god dag mann økseskaft"-diskusjoner på grunn av at enkelte ord og uttrykk hadde ulik bruk og betydning i ulike fagmiljø.

Det er vanskelig å finne dokumentasjon på slike diskusjoner direkte i etterkant, men en kan lett se variasjon i begrepsbruk i KS-rapporter fra den tidlige fasen av ordningen. Følgende sitater (fra rapporter tilhørende hver av KS-rådgiverne med rammeavtale i første periode) illustrerer forskjell i fremstilling av det fenomenet som senere er definert som "avsetning for usikkerhet". Nr. 3 unngår problemet med å beskrive indirekte. I tillegg vet vi at de ulike etatene også hadde sine egne tilsvarende begreper som heller ikke samsvarte med de som er referert her.

1. *"Prosjektreserve utover kontrollestimatet kan fastsettes ut fra et 85% sikkerhetsnivå." (år 2001)*
2. *"[KS-rådgiver] anbefaler følgende retningslinjer for disponering av reserver og margin:" (år 2002)*
3. *"Usikkerhetsspennet for kontantstrømmen for året 2007 [lav verdi, middels verdi, høy verdi]." (år 2001)*
4. *"Avsetningen på 145 mill. kr (kostnadsramme fratrukket styringsramme) kan ses på som en finansiell beredskap." (år 2003)*
5. *"Avsetning for usikkerhet: 14 mill. NOK" (år 2001)*

En KS-rapport sier det eksplisitt; "Det har også i løpet av prosessen forekommet tilfeller av begrepsforvirring, som det gjennom siste analysesesjonen er ryddet opp i."

Som respons til denne typen diskusjoner ble det etablert et felles begrepsapparat som skulle ordne problemet (Klakegg m fl., 2003). Helt borte er det ikke blitt, men en har i alle fall en

felles referanseramme som kan benyttes til å oversette ulike fremstillinger og sikre mot de verste problemene.

(10) Å benytte feil begreper eller misbruke ord og uttrykk er fallgruver som selvsagt vil kunne gi dårlig informasjon til planlegging og kvalitetssikring.

(11) Å tro at problemet er løst ved at det er utarbeidet og ”vedtatt” et felles begrepsapparat er trolig en like stor fallgruve. Det kreves mer.

(12) Den siste typen fallgruve som trekkes frem, knyttet til kommunikasjon, er de holdninger, kultur og åpenhet som er i situasjonen rundt kommunikasjonen. Det må skapes en tillit mellom partene og holdning som tilsier åpenhet. Det betyr at høy grad av profesjonalitet, lojalitet til oppgaven/rollen, evne til å ta kritikk/læring og evne til å få frem de riktige poengene må demonstreres. En slik kultur og gjensidig tillit (her spesielt i en kvalitetssikringsammenheng) er avgjørende for hvor lett en klarer å håndtere og motvirke de andre fallgruvene. En fallgruve vil være å overse dette.

Et par positive eksempler for å illustrere poenget (uthevingene er mine):

*1. Oppdraget ble innledet med innsamling av dokumentasjon fra prosjektet samt møter / intervjuer med personell i prosjektet. Referansekontroll og vurdering av status i forbindelse med planprosessen er utført med bl.a. intervjuer av prosjektets nøkkelpersonell og andre relevante aktører.*

*Hovedinntrykket gjennom hele denne prosessen har vært at man har hatt en åpen og god dialog med alle parter, noe som har bidratt til raske avklaringer og et godt samarbeid.*

*2. ”Det har vært en positiv og åpen tone på samlinger og møter. Informasjon og presentasjoner fra prosjektet er gitt på en systematisk og ryddig måte. [KS-rådgiver] har ikke hatt kontakt med, eller innhentet informasjon direkte fra aktuelle leverandører.”*

Kontrapunktet til den fallgraven som er representert ved å ikke sikre en god tillit og holdning, vil være en parallell til den over nevnte fallgraven om å la være å kvalitetssikre fordi informasjonen kommer fra ”sikker” kilde. Tonen må ikke bli så jovial at det ikke lenger stilles kritiske spørsmål. Dette er omtalt som fallgruve (2).

Realiteten er ikke til å komme bort fra: Det er begrenset med kilder tilgjengelig som har god informasjon som er relevant som sammenlikningsgrunnlag. Når det kommer til stykket er planlegger/kvalitetssikrer/forsker nødt til å stole på tilgjengelige kilder for svar som ikke finnes andre steder. Det innebærer at når en ikke har alternativer og kilden vurderes som troverdig, da blir den informasjonen som er tilgjengelig fra den aktuelle kilden benyttet. Slik må det i praksis være. Dersom det hefter usikkerhet ved eller brist i troverdigheten bør dette gjøres eksplisitt kjent i beslutningsunderlaget.

## **E. Arbeidsformen**

(13) Ustrukturert arbeidsform er også en fallgruve. God dialog oppnås i praksis ved å tilnærme seg saken på en mer eller mindre strukturert måte i direkte kontakt med en eller flere kilder (personer, grupper). Det er sammenfallende med temaet for hele denne rapporten, og vi skal derfor ikke gå veldig langt inn i dette her. I kvalitetssikringsordningen er eksempler på



både metoder for strukturerte intervjuer (dialog med enkeltpersoner) og gruppeprosesser (dialog med flere samtidig). Normalt benyttes begge deler av alle KS-rådgivere (og forskere) men i ulike former, mengder og med ulike formål. Noen eksempler:

1. *"Det er intervjuet utvalgte nøkkelpersoner som er identifiserte i prosjektets kvalitetsplan og samt nøkkelpersoner som i Håndbok 144 og 151 fremgår som ansvarlige for prosjektet på et prosjektfaglig, så vel som overordnet nivå (Ref. kjerneprosess: "Prosjektering og bygging av veg")." .... Og videre:*

*"For å gjennomføre kvalitetssikringen og intervjuene på en så effektiv og verdifull måte som mulig, ble det utviklet et "Prosessnotat- og spørreskjema" (se vedlegg 2). Dette dokumentet inneholder spørsmål relatert til de ulike temaer som i henhold til rammeavtalen er ønsket behøyst. Spørsmålene er av både objektiv og subjektiv karakter, for til en viss grad å underbygge de subjektive utsagn med objektive "data"."*

2. *"Foruten oppstartsmøtet er det avholdt flere møter med Jernbaneverket og alle relevante interessenter har vært intervjuet. Hovedtemaene har vært:*

- *Nyttkostnadsanalyse*
- *Organisering og styring*
- *Kontraktsstrategi*
- *Nøkkeltallsanalyse*
- *Estimatgjennomgang*
- *Usikkerhetsanalyse"*

3. *"Usikkerhetsanalysen ble gjennomført i henhold til Trinnvisprosessen, som er basert på suksessprinsippet og tilsvarer Statens vegvesen sin egen metode, Anslag.<sup>37</sup>"*

4. *"Med basis i fase 1 gjennomfører [KS-rådgiver] en gruppeprosess sammen med prosjektets nøkkelpersoner. Deltagere i gruppeprosessen skal samlet presentere nødvendig kunnskap og erfaring til at prosessen blir god nok. Sammensetningen er derfor viktig og skjer i utgangspunktet med prosjektets nøkkelpersonell og [KS-rådgiver]s fagpersoner. Dersom det er nødvendig kan ytterligere fagpersoner hentes inn."*

Kommunikasjon og kvaliteten på informasjon er ikke bare et spørsmål om språk og systematikk. Det er også et spørsmål om psykologiske effekter. Her vil vi trekke frem en problemstilling som angår kommunikasjonen og som er tatt opp av flere forfattere: Hvordan vi spør styrer resultatet, dvs. avgjør hva vi får til svar, påvirker altså kvaliteten på den informasjonen vi klarer å innhente. (Teigen, 2006). Det er altså slik at vi kan påvirke svaret med måten vi spør på.

(14) En fallgrube vil da være å spørre på feil måte. En kan risikere å spørre på en måte som ikke gir objektivt riktige svar og dermed farget eller forvrengt informasjon. Brukt bevisst kan dette være en måte å manipulere resultatet av planlegging, analyse eller forskning. Dette har vi ikke funnet dokumentasjon på at det har forekommet. Noe vi derimot er sikker på har

---

<sup>37</sup> Trinnvis-prosessen og Anslag-metoden er to ulike navn på den samme typen strukturerte gruppeprosesser ledet av en prosessleder. Se Klakegg 1993 og Klakegg 2000.

forekommet, er at en ubevisst spør på en måte som ikke er optimal i forhold til å få frem det beste svaret. Dette er en fallgrube som fortjener mer oppmerksomhet enn den har fått hittil. Teigen (Jørgensen og Teigen, 2005) påviser at vi er altfor sikre i våre vurderinger av usikre størrelser. De gir praktiske eksempler fra sin forskning. Basert på disse og lignende funn presenterer de alternative måter å spørre på.

Eksempler:

1. *"Ulike beskrivelser av usikkerhet;*

- *I tall: Det er 50% sjans for regn*
- *Med ord: Det er "en mulighet" for regn*
- *I usikkerhetsintervall: Det vil komme 0-10 mm. nedbør."*

2. *"Konfidensintervaller er (nesten) alltid for trange; 98% min-maks intervaller for arbeidstid ga 60% treff (ref. Conolly & Dean referert i Jørgensen og Teigen, 2005).*

3. *"Graden av konfidens påvirker ikke intervallbredden. Grupper som skulle oppgi 50%, 70% 90% eller 99% intervaller for arbeidstid oppga like store intervaller.*

(15) Tidspunktet påvirker muligheten for å kvalitetssikre informasjon. Problemstillingen posisjonering i forbindelse med konkurranse om oppdrag (likebehandling av leverandører) påvirker muligheten for å sjakke ut informasjon. Flere av de gjennomførte kvalitetssikringene i den første perioden med ekstern kvalitetssikring refererer til liknende problemstillinger som dette, gjengitt fra en kvalitetssikringsrapport:

*"Av hensyn til pågående dialog mellom prosjektet og leverandør har man valgt ikke å gi [KS-rådgiver] anledning til å kontakte leverandør(er) for behyse prosjektet fra dette ståsted. .... [KS-rådgiver] har ikke hatt kontakt med eller innhentet informasjon direkte fra aktuelle leverandører."*

Et annet sitat viser tydelig misnøye med en liknende situasjon under kvalitetssikringen i et annet prosjekt:

*"Tidspunktet for kvalitetssikring kunne neppe vært mindre ideelt enn det var i dette tilfellet. Arbeidet foregikk i parallell med forhandlinger og midt oppi en hektisk beslutningsprosess. Det bør legges stor vekt på å gjennomføre tilsvarende kvalitetssikring tidligere for fremtiden."*

Dette peker på en fallgrube som kanskje ikke er så åpenbar i forhold til kvalitetssikring av informasjon; tidspunkt for kvalitetssikring. Ettersom vurderingen er at kvalitetssikrer i de utvalgte eksemplene ikke kan ha kontakt med leverandør på det aktuelle tidspunktet for å få verifisert informasjon eller supplere med manglende kunnskap, er det ikke et velvalgt tidspunkt. Det er også karakteristisk at dette skjedde i starten av ordningen med ekstern kvalitetssikring. De første prosjektene som ble gjennomgått var egentlig kommet lenger enn det som var utgangspunktet for utforming av ordningen. Dette er et problem som gjelder generelt ved innretning av nye ordinger. Et sted må grensen gå for hvilke tilfeller som skal inngå og hvilke som ikke skal inngå. Greit nok, men tidspunktet for kvalitetssikring av informasjonen er uansett et poeng: Tidspunktet påvirker hvilke kilder som er tilgjengelig. Et annet eksempel som illustrerer dette, er tidspunkt for sluttrapport fra ferdig utførte prosjekt.

Det er viktig for både produktiviteten og kvaliteten på sluttrapport at den blir utarbeidet før de sentrale personene i prosjektet forlater prosjektet og går løs på det neste.

Den som har prøvd, vet at det er mye lettere å få tak i riktig informasjon så lenge nøkkelpersonene i prosjektet er tilgjengelige og i det aktuelle prosjektet, enn det er senere. Tilgjengeligheten av god informasjon i skriftlige kilder og arkiver etter et prosjekt er ofte vanskelig å finne frem i. Det er vanskelig å finne den relevante informasjonen, og ennå vanskeligere å få verifisert at det er siste, korrekte versjon av dokumentene etc., for ikke å snakke om å få bekreftet at det som står er korrekt og i samsvar med virkeligheten.

## **F. Lagring og gjenfinning av informasjon**

Lagring av informasjon er en vanskelig og viktig oppgave. Slike problemer er vanlige men blir aldri omtalt i kvalitetssikringsrapportene eller andre tilsvarende kilder. Nedenfor er benyttet egne erfaringer som kilde. Temaet har ”alltid” hatt stort fokus, så det er egentlig oppsiktsvekkende at problemet ikke er bedre håndtert enn det er. Fortsatt er det kun papirarkiver tilgjengelig i mange ferdige prosjekter. Dette kan selvsagt være et bevisst valg for å gjøre informasjon vanskelig tilgjengelig (hemmelighold etc.). Store statlige investeringsprosjekter er så mangt, og i enkelte tilfeller kan det være relevant med forsiktighet (for eksempel enkelte av Forsvarets materiellkjøp). Vi antar i denne sammenheng at nødvendig sikkerhet kan ivaretas gjennom gode holdninger og systemer hos den som skal benytte informasjonen. Utgangspunkt for videre diskusjon er at det er ønskelig å få tak i korrekt informasjon og at det skal være mulig å få denne informasjonen verifisert/kvalitetssikret.

Det er etter hvert en selvsagt ting at arkiver i bruk hovedsakelig er elektroniske. Dette gjør lagringen mye mer konsentrert og dokumentene mer tilgjengelig for den som har de riktige akkreditiver. Det tillater effektive elektroniske søk, noe som potensielt kan gjøre en forsker, planlegger eller kvalitetssikrers jobb mye enklere. Imidlertid viser praktisk erfaring at det ikke er så enkelt likevel. Søk kan mislykkes av mange årsaker. De viktigste og mest vanlige årsakene er at den informasjonen vi søker ikke blir funnet ved målrettet søk på stikkord eller i fritekstsøk. Vanligvis betyr det enten at

- informasjonen ikke finnes i dokumentet, eller
- det ikke er samsvar mellom søkekriteriene og den teksten en er ute etter.

(16) Det første kan bety at informasjonen ikke finnes i det hele tatt, eller ikke er dokumentert. Det er en åpenbar felle som ikke diskuteres videre her. Det kan sees på som underliggende til fallgruve (4). Alternativt kan det bety at informasjonen er plassert i et annet dokument.

(17) Hvis informasjonen er i feil dokument eller feil begreper er brukt, kan det bety at en ikke har fulgt prosedyrene som bestemmer innhold og struktur eller navngiving av dokumenter, der slike finnes. Å ikke ha noen form for føring på slikt er en fallgruve på systemnivå.

(18) Å ikke følge slike rutiner der de finnes er en fallgruve på individnivå. Hvis det ikke blir treff på elektroniske søk kan det være en indikasjon på at det er benyttet andre begreper og

termer enn det som er vanlig eller faglig norm. Feilen kan være forårsaket av den som har laget dokumentet, den som har valgt søkekriteriet eller begge. Derfor er det god hjelp til kvalitetssikring av informasjon at det finnes standarder for begrepsbruk og termer. Å overse eller neglisjere slike standarder vil være en fallgrube, uansett hvilken rolle en har i dette spillet.

## Oppsummering

Det er mange fallgruver i arbeidet med å innhente og kvalitetssikre informasjon. Vi har funnet indikasjoner som bekrefter det generelle inntrykket at her er det mye som kan og bør forbedres. Vi kan ikke ut fra dette materialet konkludere at det er mye slendrian i arbeidet med innsamling og bearbeiding av informasjon. På den andre siden kan det i alle fall ikke motbevise. Det er veldig lett å finne kritiske bemerkninger om andres informasjonsgrunnlag og bearbeiding av informasjon. Det er bemerkelsesverdig få tilløp til kritikk av grunnlaget for egne vurderinger. En hovedbetragtning denne forfatteren har gjort seg i løpet av dette arbeidet er at en bør se på dette arbeidet med større grad av kritisk sans. Hvis jobben (samle og bearbeide informasjon der grunnlaget er tynt) er verd å gjøre – så er det verd å gjøre den skikkelig.

## 8.4 Kvalitetssikring av beslutningsunderlag – beste praksis

Utgangspunktet for dette avsnittet er å hente frem god systematikk fra ulike kilder. Også her er KS-rapportene sentrale, men ikke de eneste kildene. Det gjelder å utnytte det beste som er tilgjengelig, uansett hvor det kommer fra. Det kunne kanskje hevdes at dette temaet er synonymt med hele rapportens tematikk. Forskjellen er at vi her ser på noen praktiske eksempler og lar teorien ligge.

Som antydte både mellom linjene og eksplisitt i avsnittet om fallgruvene fremfor er det mye som kan gjøres for å forbedre systematikk og kommunikasjon i alle faser og aspekter ved prosjekt. Kapitlene både lenger foran og bak i rapporten omhandler dette. Noen av aspektene som blir omtalt i andre kapitler kunne lett komme inn her også, men det er gjort et forsøk på å unngå dublering av tema.

### Bevisst valg av ressurser for utførelse

En problemstilling som er i slekt med de to foregående, men som nevnes for seg likevel er god praksis for valg av utførende ressurser. Poenget her er ikke å velge noen for å gå gjennom det samme materialet, men å velge de ressursene som er best egnet til å gjøre oppgaven i første omgang. Dette er en viktig detalj i kvalitetssikringen. Spesielt viktig i tidlig fase med tynt informasjonsgrunnlag. Denne situasjonen stiller spesielt store krav til utførende ressurs. Her vises et par eksempler på å velge andre ressurser enn de ordinære til enkeltoppgaver. Eksempelene er hentet fra KS-rapporter (mine uthevinger):

1. Internt: *"Identifikasjon og kvantifikasjon av usikkerhetslementer ble gjennomført ved gruppesamlinger og intervjuer under. [KS-rådgiver] etablerte på bakgrunn av dette analysemodellen og tilhørende resultater. Man har i tillegg gjennomført en verifisering av underlag, forutsetninger og usikkerhetsspenn med programkontoret. [KS-rådgiver] ønsket å gjennomføre analysen med størst mulig grad av objektivitet. Andre ressurser i [KS-rådgiver] fikk derfor ansvar for usikkerhetsanalysen enn det personell som var ansvarlig for å følge forprosjektet."*

2. Eksternt: "[KS-rådgiver] har ikke hatt kontakt med eller innhentet informasjon direkte fra aktuelle

entreprenører. Deler av kostnadsoverslaget baserer seg på tidligere informasjon fra aktuelle leverandører og entreprenører innhentet av [Etaten] og deres interne og eksterne rådgivere. **[KS-rådgiver] har i sin kvalitetssikring valgt å knytte til seg noen av disse eksterne rådgiverne. I tillegg til allerede å ha innsikt i prosjektet, er de blant de ledende tekniske rådgivere innen bru- og vegplanlegging i Norge. Situasjonen i anleggsmarkedet er diskutert med markedsanalytikere.**"

3. I analyse: "Med basis i fase 1 gjennomfører [KS-rådgiver] en gruppeprosess sammen med prosjektets nøkkelpersoner. Deltakere i gruppeprosessen skal samlet representere nødvendig kunnskap og erfaring til at prosessen blir god nok. **Sammensetningen er derfor viktig og skjer med prosjektets nøkkelpersoner og [KS-rådgiver]s fagfolk. Dersom det er nødvendig kan ytterligere fagpersoner hentes inn.**"

### God, strukturert prosess

Har en fått på plass de beste ressursene og dermed skaffet seg et godt startpunkt, er neste skritt mot god kvalitet å sikre en god prosess som er klargjørende og prinsipiell, samtidig som den er enkel og effektiv. Slike prosesser er representert i andre kapitler i denne rapporten. Her tas med noen eksempler (referanser) fra KS-rapporter:

1. "Usikkerhetsanalyse av ytelse, tids- og kostnadsestimater er gjennomført i henhold til [KS-rådgiver] standard prosess for usikkerhetsanalyse. [KS-rådgiver] hadde i forkant av analysen en grundig gjennomgang av det endelige styringsunderlag. Gjennomføring av selve analysen ble planlagt sammen med [Prosjektet] for å sikre at alle nøkkelpersoner avsatte nødvendig tid til å bidra. Identifikasjon og kvantifikasjon av usikkerhetslementer ble gjennomført ved gruppesamlinger og intervjuer under.

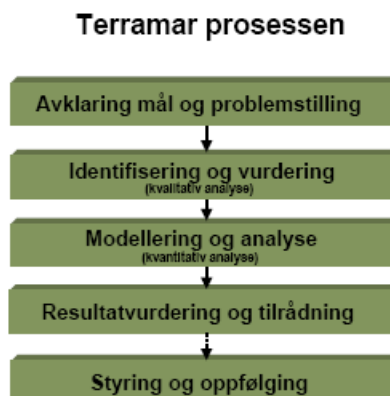
.....

[KS-rådgiver] etablerte på bakgrunn av dette analysemodellen og tilhørende resultater.

En sentral del av usikkerhetsanalysen var samlingene som ble gjennomført med prosjektaktører. Usikkerhetsanalysen ble gjennomført i henhold til Trinnvisprosessen."

Vi velger å ikke gå inn i detaljer på hvordan disse prosessene skal struktureres. Det er et tema som er dekket av en rekke kilder. I Concept-rapport nr 12 Usikkerhetsanalyse; Metoder (Austeng m.fl. 2006) er dette dekket for usikkerhetsanalyser slik som den som refereres i eksempelet. Figur 1 viser et tilfeldig valgt av de aktuelle metodene som benyttes aktivt i ekstern kvalitetssikring. Figuren illustrerer prinsipielle steg i prosessen, enkelhet logisk oppbygging. Som påvist av Austeng m.fl. er prosessene til de ulike KS-rådgiverne nokså like i oppbygging og innhold.

Usikkerhetsanalysen som utføres i forbindelse med kvalitetssikring av store statlige investeringsprosjekter følger Terramar-prosessen. De ulike fasene i prosessen er illustrert i Figur 1. Innholdet i de ulike fasene beskrives i de påfølgende kapitler.



Figur 8-5 Terramar prosessen – Eksempel på strukturert prosess for analyse.

Før vi går inn i vurderingene i denne prosessen, et lite hjertesukk: Husk å la det være tid til noe egenbedømmelse (self assessment) og kritisk ettertanke i prosessen du setter i gang. Ofte har tidspresset en medvirkende årsak til at vi går i fallgruvene fra forrige avsnitt.

### Vurderinger i prosessen

Det er umulig i et kort resymé å komme inn på alle sider ved selve vurderinger. Her tas med en liten antydning om hva som kan være et bidrag til god praksis i forhold til fallgruven om å spørre feil og dermed oppnå at eksperten sier noe annet enn det som gir deg beste mulige data. Det er et utsnitt av artikkelen ”Skjønn og skivebom” av Karl Halvor Teigen (Teigen, 2006):

*”4) Så hva kan gjøres for å motvirke noen av disse feilkildene?*

*(a) Lage intervaller ut fra enkle regler istedenfor skjønn (pluss minus 50% eller lign).*

*(b) Ta et ytre istedenfor et indre perspektiv (hvor store feilmarginer er vanlige?)*

*(c) La personer som ikke er involverte anslå feilmarginene*

*(d) Føre statistikk over treffsikkerheten*

*(e) Spørre om øvre og nedre grenser separat (ikke spør ”mellom”)*

*(f) Spør om usannsynlige og ikke sannsynlige verdier (over/ under bedre enn minimum/ maksimum; grenseverdiene er ekskludert og ikke inkludert i intervallet).*

*(g) Bestem intervallet først og anslå sikkerheten etterpå (ikke bygge inn krav om at man skal være ”90% sikre”) ”*

Noen generelle og allmenngyldige anbefalinger synes å kunne trekkes ut av dette utsnittet:

- Spør på en bevisst, strukturert og tydelig måte. Måten du spør på, påvirker svarene.
- Ta på alvor at den du spør har et forhold til det du spør om. Den som er veldig nært knyttet til utfallet (resultatet) vil se ting annerledes enn en som ser ting mer på avstand. Spør begge to. Se ting fra flere sider.
- Evaluér jobben som blir gjort (vurderingene) i etterkant slik at det er mulig å lære av det.

Hvor mange kan med hånden på hjertet si at de alltid følger disse enkle levereglene? Og til alle de andre: Hvorfor følger vi ikke disse enkle levereglene? Aksepterer vi heller å ha dårlig kvalitet på den informasjonen vi bygger våre viktige beslutninger på for å spare bryet? Eller er kanskje ikke de viktige beslutningene så viktige likevel?

### **Peer review / kollegagjennomgang**

Den enkleste og mest utbredte form for kvalitetssikring av enkeltelementer av informasjon og eventuelt hele beslutningsunderlaget, er å la en kollega (likemann) se over og vurdere materialet før arbeidet avsluttes og sendes til beslutningstaker. Denne enkle men effektive formen for kvalitetssikring bør være obligatorisk i alle viktige beslutningsprosesser. Det anbefales selv om det også skal benyttes andre virkemidler etterpå. Denne måten å arbeide på har mange gode egenskaper som bør settes pris på i enhver organisasjon: En vellykket kvalitetssikring gjør at de store tabbene blir avslørt før de kommer ut av huset. Dessuten får flere personer innsyn i det arbeidet som gjøres, noe som gjør organisasjonen mindre sårbar. Det medfører også læring på tvers og øker den kollegiale samfølelsen i organisasjonen. Men først og fremst er det en effektiv måte å heve kvaliteten på beslutningsunderlaget.

Hva er det som gjør at kollegagjennomgang har den positive effekten på kvaliteten i materialet? Det forutsettes en organisasjonskultur som tillater og underbygger kritisk holdning til materialet som legges fram. Det må ikke bare være spill for galleriet, men en virkelig kritisk vurdering som utføres. Det at en annen person kommer inn med egne vurderinger og et mer ”nøytralt” blikk på saken gjør at flere sider kommer frem og blir vurdert. Kollegaen har også dels andre erfaringer og tilgang til annet nettverk og andre kunnskapskilder. Dette gjør sjansen større for at feil, mangler og inkonsistens blir avslørt. En kollega kan imidlertid ikke sies å være habil, så nøytraliteten er nokså illusorisk.

### **Tredjepartskontroll**

Dersom det kreves en mer reelt nøytral og habil vurdering benyttes en annen fagperson utenfor egen bedrift. Effekten er den samme som beskrevet over, men ekstra styrke i dette virkemidlet ligger i at det er en helt utenforstående som vurderer. En slik oppgave har også aspekt av ansvarliggjøring over seg. En som tar på seg et slikt oppdrag går god for kvaliteten med sitt eget (firma-)navn og eventuell kvalitetssvikt senere vil henge ved også den som gjorde en dårlig tredjepartskontroll.

I noen tilfeller kan det synes som det oppstår et konkurranseforhold mellom den som gjør førstevurderingen og den som gjør tredjepartskontrollen om å være den som vet best. Det kan vise seg ved at de to ikke er enige og ikke kommer frem til en samlende vurdering. Dette kan medføre en polarisering av vurderingene som gjør det vanskeligere å vurdere helheten for den som skal ta beslutningen. Da oppstår en situasjon som er ganske lik det vi har registrert også tilknyttet KS-ordningen.

Forskning på den norske KS-ordningen viser at beslutningene om kostnadsramme i perioden 2000 – 2004 (basert på gjennomgang av 31 prosjekter i Trailbase) oftest – i 58 % av tilfellene - ble basert på KS-rådgivers anbefaling. I en del tilfeller ble det også fra etatens side laget nye kostnadsoverslag basert på resultatene/anbefalingene fra ekstern kvalitetssikrer (Magnussen og Olsson, 2006). KS-rådgiver tilsvarende den eksterne tredjepartskontrollør i forhold til prosjekt/etat. Denne posisjonen veier tydeligvis svært tungt. Vurderingene fra

tredjepartskontrollør blir brukt og Magnussen & Olsson viser også til at prosjekteier (departementene) opplever denne tredjepartsvurderingen som nyttig. Det er imidlertid ikke grunnlag for en klar anbefaling når det gjelder hvem en skal lytte til når beslutningen skal tas. Hvem skal en tro på når to eksperter er uenige?

### **Innhenting av data**

Som vist i andre delen er flere av fallgruvene knyttet til tilgang på data. Dette må altså gjøres skikkelig. Problemstillingen er i høyeste grad aktuell i ekstern kvalitetssikring. Her er et eksempel fra en KS-rapport som omtaler arbeidet med informasjonsinnhenting og håndtering:

*“Evalueringen er metodisk basert på en top-down tilnærming hvor grad av detaljering/ytterligere fokus på enkeltforhold er basert på fortløpende resultater. Arbeidet har i hovedsak bestått i å systematisere en stor mengde informasjon mottatt fra ulike kilder. Datainnhenting har vært omfattende på grunn av antall tema og problemstillingenes kompleksitet. Mye av nødvendig underlagsinformasjon har vært funnet i eksisterende dokumentasjon, men det har vært behov for et betydelig arbeid knyttet til identifisering, oppdatering/normalisering og tilrettelegging av informasjonen. I de tilfellene det har vært motstridende informasjon, har det vært behov for avklaringer med prosjektet og etatene, og deretter vektning av innspillene.”*

Eksempelet går ikke i detalj om de mange problemene en slik oppgave medfører, men av egen erfaring kan det listes opp en rekke utfordringer: Vanskelighet med å identifisere relevante referanseprosjekter og kilder til data. Tilgangen til data er relativt åpen. De fleste norske etater og firma er relativt sjenerøse i sin håndtering av det de har, men de har ofte ikke mye å bidra med, og kvaliteten på det en får er så som så. Ofte er det enkelttall som ikke er med, enkelte data som åpenbart ikke stemmer (verre er det med de som er feil men ikke åpenbart feil – de blir neppe oppdaget). Spesielt vanlig og vanskelig er at forutsetningene ofte mangler. Noen kommersielle kilder er tilgjengelige men for den som gjerne vil dykke ned i dybden i disse tallene blir bildet omtrent det samme som beskrevet over. Rett nok virker systematikken bedre, men tilgangen på forutsetninger og kunnskap om kildene er ofte mangelfull. Å ha god tilgang til data er i dag en klar konkurransefordel i forhold til visse typer oppgaver, og et potensielt produkt i seg selv.

### **Referansesjekk**

Et begrep som ble introdusert relativt tidlig i KS-ordningen er referansesjekk. Allerede i 2001 er det omtalt i en av de aller første KS-rapportene – se eksempel 3 nedenfor. Ettersom det er eneste flyplassprosjekt som er kvalitetssikret har det ingen hensikt å anonymisere dette bidraget. Begrepet står for en egen kontroll der kostnadstall fra det aktuelle prosjektet som blir kvalitetssikret sammenlignes med tilsvarende nøkkeltall fra andre sammenlignbare prosjekt. Hensikten er å fastslå om hele, deler av, eller enkelte kostnadselementer har et rimelig kostnadsnivå i forhold til andre prosjekter som er kjent. Dette er en kontroll på rimeligheten i kostnadsestimatene og bidrar til kritisk tenkning rundt både sammenligningsgrunnlag og forutsetningene for kostnadsestimatene som benyttes.

I dag foregår dette på ulike måter hos de ulike KS-rådgiverne. Det er veldig ulik grad av beskrivelse knyttet til dette i rapportene. Alt fra den korte omtalen i eksempel 1. til den omfattende utgaven som er dokumentert i egen rapport. Det er ikke grunn til å tro at det er direkte sammenheng mellom antall sider dokumentasjon av referansesjekk og omfanget av



arbeidstimer som går med til denne oppgaven hos de enkelte KS-rådgivere. Oppgaven er viktig og krevende, så arbeidsmengden antas betydelig også hos KS-rådgivere som velger å kun vise en kort omtale av det.

Eksempler på ulike former for referansesjekk:

1. *"Spenn og usikkerheter er diskutert med prosjektorganisasjonen og de eksterne rådgiverne. Denne informasjonen er sammenlignet med erfaringsdata fra tilsvarende prosjekter."*

2. *"[KS-rådgiver] har i sitt kvalitetssikringsarbeid foretatt en gjennomgang av grunnkalkyletall utarbeidet av [Prosjektets rådgiver]. Grunnkalkylen er sjekket mot nøkkeltall for enhetspriser.*

.....

*[KS-rådgiver] har gjort følgende endringer i [Prosjektets rådgiver]s kalkyle: .... "*

3. *"I forbindelse med 1. avrop under Rammeavtale inngått med Finansdepartementet har det blitt utført en uavhengig referansesjekk for Stavanger Lufthavn. Følgende flyplassrelaterte referanseprosjekter er gjennomgått:*

- *Trondheim Lufthavn Værnes, nytt terminalbygg, rehabilitering eksisterende terminalbygg og utendørsarbeider flyside.*
- *Harstad/Narvik lufthavn Evenes, nytt ekspedisjonsbygg og utendørsarbeider flyside/ landside.*
- *Tromsø Lufthavn, nytt ekspedisjonsbygg, fly- og landside.*
- *Oslo Lufthavn Gardermoen, nytt ekspedisjonsbygg.*
- *Oslo Lufthavn Gardermoen, Schengen tilpasning ekspedisjonsbygg.*
- *Ny Flytogterminal på Oslo S for Gardermobanen.*

Gjennomgangen bygger på innsamlede forprosjektdokumenter, tegninger og erfaringstall fra utførte utbyggingsprosjekter. Alle tall er oppjustert til 1999 priser for sammenligning med kostnadstall for Stavanger lufthavn, Sola.

Gjennomgangen er utført på et overordnet nivå, basert på den informasjon det har lyktes å skaffe innenfor tidsfristen. Dokumentet setter kostnadene ved utvidelse av Stavanger lufthavn i ulike perspektiver for å vurdere realismen i kostnadsnivået som er lagt til grunn.

*Vi finner det dokumentert at kostnadsnivået i det materialet som er lagt til grunn for analyse og kostnadsoverslag er realistisk i forhold til relevante referanseprosjekt."*

I eksempel nr. 2 benyttes sammenligning av nøkkeltall på elementnivå. På dette nivået er det lett å finne sammenlignbare tall og sammenligningen kan gjøres direkte når forutsetningene er presise. Eksempel 3 viser en sammenligning på overordnet nivå. På dette nivået er forutsetningene atskillig mindre konkret og sammenligningen må gjøres som en helhetsvurdering der en rekke avvikende forutsetninger vurderes under ett.

Eksempel 3 er relativt godt dokumentert og beskrevet. Vi tar derfor med litt flere detaljer her. Denne, og noen av de andre KS-rådgiverne, legger vekt på at referansesjekken skal utføres av andre fagpersoner enn de som har dialogen med prosjektet som kvalitetssikres. Dette har den hensikt å unngå at vurderingene farges av dialogen med prosjektets nøkkelfolk. Referansesjekken har noe av samme funksjon som peer review (likemannskontroll). Følgende sitat forteller både om oppgaven og hvilke utfordringer den representerer:

*"Dette etableres i en egen kostnadsanalyse som dokumenteres separat. Eventuell identifikasjon av store avvik i kostnadsnivå ("skivebom") bør få konsekvenser for det totale usikkerhetsbildet i prosjektet og kan kreve tiltak. Dette skal medføre økning av den totale kvaliteten på kostnadsoverslaget og prosjektets planlegging."*

... og videre:

*"Informasjon om referanseprosjektene er innhentet direkte fra primærkilder, det vil si fra personer og organisasjonsenheter direkte involvert i de enkelte referanseprosjektene. Informasjonen er hentet fra kopier av originaldokumenter, de er ikke bearbeidet/tilpasset til vårt formål av kilden. I dette tilfellet er tilgangen på referanseprosjekter i Norge relativt begrenset. Det ble likevel valgt å ikke gå ut over Norges grenser for å innhente erfaringstall da disse ville være meget vanskelige å vurdere opp mot norske forhold."*

Sammenligning av de ulike prosjekter er vanskelig på grunn av at alle prosjektene stiller opp sine kostnadsoverslag på ulike måter og definerer noen kostnadselementer ulikt i forhold til NS 3453. I tillegg kommer at prosjektene har ulikt omfang, geografisk plassering og kompleksitet. Følgende karakteristika ved prosjektene danner grunnlag for å vurdere erfaringskostnadene:

Byggherreorganisasjon

Omfang/dimensjon (størrelsesfaktor)

Valg av kvalitet og løsninger, estetikk (materialer, overflater, former)

Teknisk kompleksitet (integrasjon av systemer m.m.)

Arealfordelinger på ulike funksjoner (noen funksjoner er dyrere enn andre)

Grad av rehabilitering og provisorier, driftshensyn (konflikt mellom bygging og drift, flytting av funksjoner)

Kontraktstrategi/entrepriseform (type entrepriser, størrelse på kontrakter)

- Lokale prisforhold (marked på kontraheringstidspunktet, konkurranseforhold, lokal kapasitet)

*Erferingstall for de enkelte prosjekt må justeres med hensyn til disse forholdene for å bli direkte sammenlignbare. Vi har valgt å foreta en grov sammenligning uten å kvantifisere disse forholdene."*

En kan i dette eksempelet kjenne igjen flere av fallgruvene fra forrige avsnitt, og hvordan KS-rådgiver i denne sammenhengen har håndtert disse. Fra konklusjonen gjengis en utsnitt:

*”Det synes som prosjektet har benyttet relevante referanser og erfaringsdata i sin planlegging. Nivået er tilstrekkelig sannsynliggjort gjennom den innsamlede dokumentasjonen. Det antas at erfaringer fra de tidligere flyplassprosjektene i Luftfartsverkets regi er samlet opp og brukt i planleggingen av dette prosjektet. Ikke minst siden prosjektleder (og andre nøkkelpersoner) også har deltatt i referanseprosjektene.*

*Vi finner det dokumentert at kostnadsnivået i det materialet som er lagt til grunn for analyse og kostnadsoverslag er realistisk i forhold til relevante referanseprosjekt.*”

Norge er et lite land. Nøkkelpersoner har deltatt i flere av referanseprosjektene og de samme referanseprosjektene er også benyttet som grunnlag for planleggingen av prosjektet som kvalitetssikres. Dette avslører feilkilder som gjør at for eksempel feilaktige forutsetninger kan bli kopiert med fra prosjekt til prosjekt. Likevel fungerer denne referansesjekkningen som en kontroll på rimeligheten i kostnadsoverslaget.

### **Referanseklasse-estimering**

Som et innspill i diskusjonen om beste praksis nevnes i denne sammenhengen Reference class forecasting som en metode som bygger på mye av samme arbeidsmåte som referansesjekkningen beskrevet over, men tar dette et par skritt lenger. Reference class forecasting er utviklet av Daniel Kahneman som fikk Nobelprisen i økonomi i 2002 sammen med Adam Tversky. Bent Flyvbjerg har tatt tak i metoden og startet å utvikle den praktiske siden ved den (Flyvbjerg, 2006), se eget kapittel i denne rapporten. Ennå gjenstår en del, men den er etter hvert tatt i bruk i enkelte prosjekter i UK og får også stor oppmerksomhet i USA og Europa.

Hovedpoenget i Kahnemans tenking er at det er for mange feilkilder i den tradisjonelle måten vi gjennomfører kostnadsestimering på. Ved å fokusere innad mot prosjektet og målsettingene for det ene vi arbeider med, blir vi ikke i stand til å se problemene som vil komme og de mange detaljene i prosjektet gjør oss blind. Vi ser ikke skogen for bare trær. Ved tvert i mot å fokusere utad mot andre tilsvarende prosjekter og spørre hvordan dette aktuelle prosjektet vil plassere seg i forhold til de mange som er utført før, unngår en mange av problemene med tradisjonell kostnadsestimering.

Arbeidsmåten går kort fortalt ut på å samle data på overordnet nivå fra en referanseklasse av tilsvarende prosjekter. Det skal ikke fokuseres på detaljer eller avvik mellom prosjektene. Referanseklassen skal være bred nok til å inneholde de samme utfordringene som det aktuelle prosjektet vil møte, og smal nok til å være relevant for sammenligning. Basert på disse dataene defineres en sannsynlighetsfunksjon/fordelingsfunksjon for sluttkostnaden til gruppen. Tilgang til tilstrekkelig med data er selvsagt en utfordring. Deretter plasseres det aktuelle prosjektet i forhold til fordelingsfunksjonen til gruppen. Dette gir et uttrykk for forventet kostnad for prosjektet.

Metoden gjør altså ingenting for å prøve å identifisere de usikre hendelser og utfall som faktisk vil komme. Den baserer seg på at noe vil alltid skje og at en aksepterer dette. Fokuset settes på den statistiske sannsynligheten for at prosjektet vil utvikle seg innenfor (eller i alle fall i forhold til) det utfallsrommet som referanseklassen beskriver. Dette er mer inngående behandlet i flere artikler, deriblant av Kahneman med flere (Lovallo og Kahneman, 2003).

Referanseklasseestimering er anbefalt av American Planners Association som et tillegg til tradisjonell kostnadsestimering (Flyvbjerg, 2006). Så lenge det mangler tilgang til et tilstrekkelig referansemateriale til å definere gode referanseklasser (tilstrekkelig data til å definere en sannsynlighetsfordeling) er det trolig vanskelig å etablere dette som fast praksis.

Imidlertid vil det nok være et interessant supplement i fremtiden. I mellomtiden kan referansesjekken med sin sammenligning av relevante enkeltcase og enkelte kostnadselementer være et sted å starte.

## Kostnadskontroll

Forsvaret har en prosedyre som skal sikre god kvalitet på informasjonen i tidligfase av sine prosjekter. Den benyttes rett nok ikke alltid, men brukes når det ikke er trygghet for at konkurransen i markedet kan kompensere for mangelen på presis informasjon. Her er et utdrag av kap. 11 i Anskaffelsesreglementet for Forsvaret (Forsvaret 2005):

*”Når det ikke er mulig å gjennomføre anskaffelsen basert på konkurranse, skal det normalt gjennomføres kostnadskontroll av pristilbudet fra leverandøren. Kostnadskontrollen gjennomføres med andre ord for å oppveie for manglende konkurranse. I slike tilfeller skal prisen baseres på at leverandøren får dekket sine relevante kostnader samt en rimelig ankastning på investert kapital for kontraktsgjennomføringen, under forutsetning av at leverandøren har en mest mulig effektiv utnyttelse av sine ressurser (innsatsfaktorer).*

*Gjennom kostnadskontrollen kan Forsvaret forsikre seg om at leverandøren har basert sine kalkyler på korrekte, oppdaterte og fullstendige data, og at leverandøren har fulgt de bestemmelser for kostnadsberegning som er nedfelt i dette regelverket. Kostnadskontrollen gjennomføres før kontrakt er inngått og/eller etter at kontrakten er fullført, avhengig av kontrakttype. Gjennomført kostnadskontroll danner basis for forhandling av en riktig pris for kontraktsgjennomføringen mellom anskaffelsesmyndigheten og leverandøren.*

### 11.1.1.1 Når foretas kostnadskontroll

*Kostnadskontroll foretas som hovedregel ved kjøp etter forhandlinger når konkurransen er utilstrekkelig. Det foretas derfor normalt ikke kostnadskontroll ved kjøp etter anbudskonkurranse, uten i helt spesielle tilfeller som f.eks. ved mistanke om prissamarbeid. Ved gjennomføring av kostnadskontroll må innsynsrett være sikret. Om innsynsrett i*

*tilknytning til kostnadskontroll (se pkt. 23.1 Generelt).*

### 11.2 Kostnadskontrollen

*Kostnadskontroll omfatter oppgavene kostnadsanalyse, kostnadsprøving, prisglidningsanalyse samt vurdering av leverandørens økonomisystem.”*

Det finnes detaljerte forskrifter og rutiner for dette arbeidet. Strukturert arbeidsform og logisk oppbygging er naturligvis sikret. Men i hvor stor grad blir det brukt? Og hvor godt fungerer det? Med en slik systematikk er vi i ytterkanten av hva som kan defineres inn i denne rapporten som skal omhandle situasjoner med tynt informasjonsgrunnlag. Her forutsettes at informasjon faktisk finnes, og at den kan bli tilgjengelig for Forsvaret fra leverandørene. I et ikke-fungerende marked har også kunden en spesiell maktposisjon som i dette tilfellet Forsvaret utnytter til å skaffe seg slik informasjon. Det er tatt med her fordi også dette er en del av det totale bildet. Det offentlige har maktposisjon i sine store investeringsprosjekter til å skaffe seg informasjon. Noen ganger ved å opptre som kontrollør slik det er vist i eksempelet over.

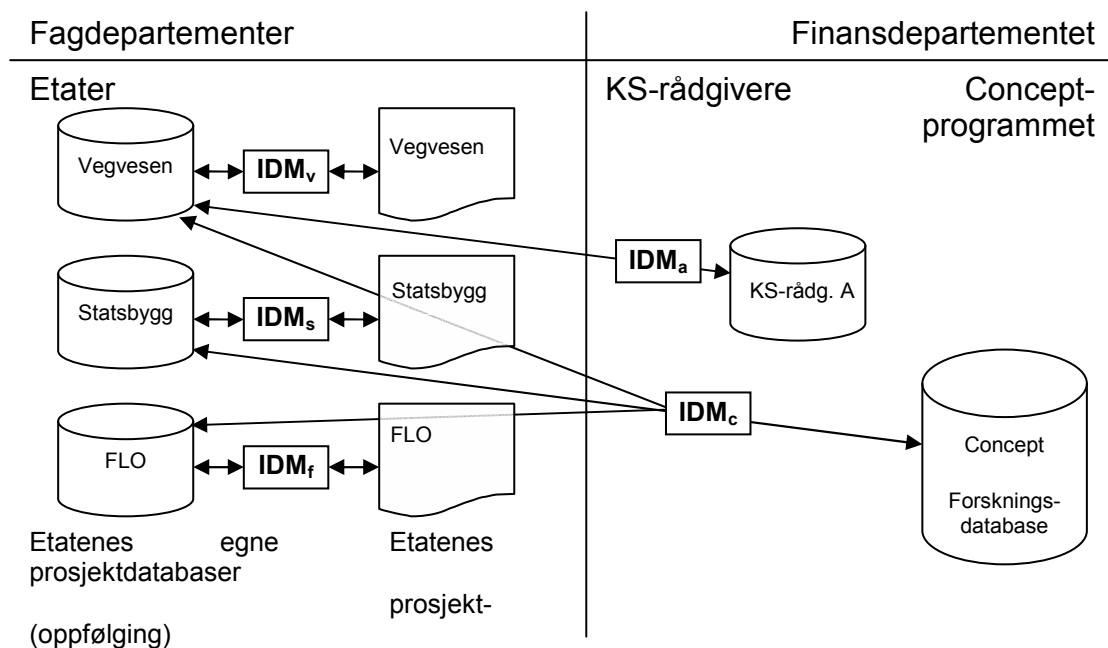
En annen slik mulighet er å benytte sin posisjon i markedet til å innlede samarbeid som åpner for tilflyt av slik informasjon. Det kan være et alternativ for noen. Det kunne være fristende å gå inn på ulike kontrahering og kontraktformer med dette perspektivet her, men det får være til en annen anledning.

### **Oppsummering**

Vi avrunder her med en slags konklusjon; tilgangen til (god) informasjon i mange tilfeller er bedre enn en kanskje skulle tro. Noen ting kan vi ikke vite noe sikkert om. Da finnes det, som andre kapitler i denne rapporten vil vise, metoder og arbeidsformer som kan hjelpe. Når det faktisk finnes noe informasjon å få tak i, så finnes det metoder og arbeidsformer som gjør at det er mulig å skaffe mye av denne informasjonen. I forhold til dagens situasjon er det i alle fall mye å hente på mer bevisst og systematisk arbeid. Men det kommer ikke gratis. Som dette kapitlet har vist må det gjøres en betydelig innsats for å innhente og kvalitetssikre informasjon i praksis. **Vi er nødt til å ta oss den tiden og bruke de ressursene som må til for å sikre oss godt informasjonsgrunnlag.** Neste avsnitt ser litt på om det er mulig å bedre på den situasjonen.

## **8.5 Hva fremtiden vil bringe**

Se for deg følgende visjon; at det riktige og relevante beslutningsunderlaget (den informasjonen som faktisk finnes) er lett tilgjengelig til enhver tid. Er det mulig? Ja, teknologien er tilgjengelig. Men er det realistisk? Ikke umiddelbart kanskje, men innen ikke altfor lang tid. La oss ta en titt på en skisse til løsning. Figur 8-5 illustrerer noen utvalgte sammenhenger og symboliserer et opplegg for automatisert innsamling og tilrettelegging av relevant prosjektinformasjon til flere formål.



Figur 8-6 Prinsippskisse for automatisk innhenting av relevante, oppdaterte prosjektdata til planlegging, kvalitetssikring og forskning.

Figuren skisserer at etaten selv, kvalitetssikrere og eventuelt forskerne kan hente den relevante og oppdaterte informasjonen ved et tastetrykk. En elektronisk hjelper, en IDM (Information Delivery Manual) gjør jobben automatisk når den blir bedt om det (Wix, 2005). Vi skal gå litt nærmere inn på forutsetningene for en slik løsning, uten å gi oss i kast med de tekniske detaljene.

Forutsetninger for en slik løsning:

- Etaten må ha en sentral prosjektdatabase med en bestemt, fast lagringsstruktur og tilhørende koding av informasjonselement.
- Etaten må ha etablert gode rutiner for planlegging og oppfølging av prosjekter, slik at prosjektene faktisk er representert i etatens prosjektdatabase med relevant, historisk og oppdatert informasjon.
- Informasjonsinnholdet i databasen må dekke behovet for både saksmessig og kontekstuell informasjon av god kvalitet. Det forutsetter gode rutiner og kompetanse hos de ansvarlige. Dette er kjent som en utfordring (Langlo et al, 2006).
- Det må etableres tilstrekkelige standarder for informasjonsutveksling (grensesnitt) for å muliggjøre slik utveksling av data og informasjon (Karud, 2006). Slike informasjonsstandarder er utviklet for flere industrier. Eksempel på en generisk standard er STEP (Standard for the Exchange of Product model data), (ISO 10303). NORSOK og POSC/CEASAR er to norske initiativ innen oljeindustrien. I noen

industrier begrenser standardene seg foreløpig til å gjelde innenfor den enkelte verdikjeden, eksempler er bilindustri og flyindustri. I det siste har dette fått stor oppmerksomhet innenfor bygg (IAI<sup>38</sup>/IFC/Building SMART). Nå i 2006 kommer anlegg etter (Aarhus 2006).

- Det må utvikles egne IDM for hver av etatene. De må utvikles til den enkelte etats behov. Tilsvarende for andre brukere og formål. Hvor mye som er felles og hvor mye som er etatsspesifikke løsninger avhenger av hvor mye felles det legges i informasjonsstandardene. Det er disse IDM som kommuniserer med prosjektdatabasen, innhenter oppdaterte data, behandler dem til relevant informasjon og presenterer den slik at den er hensiktsmessig for formålet (slik beslutningstaker har definert sitt behov).
- Brukerne må tillate sin prosjektdatabase å avgi data til den aktuelle IDM (IDM går ikke inn i databasen selv).

Denne typen informasjonsstruktur gir mulighet for å bygge helhetlige datastrukturer som samler all informasjon om et prosjekt, knyttet til en elektronisk prosjektmodell. Etatenes nytte av en slik løsning vil være stor. Redusert tidsforbruk til informasjonsinnhenting og rapportering og bedre kvalitet på informasjonen som benyttes i planlegging, vil i seg selv gi stor forbedring for hver etats prosjektarbeid. I tillegg kan etatens strategiske planlegging bli tilsvarende bedre med egne IDM til dette formålet. Ved å sikre en fornuftig og automatisk logging av historisk informasjon på viktige tidspunkt i tidsforløpet, sikres tilgang til all den relevante informasjonen som trengs til analyser, etterprøvinger og evalueringer. I dag er det et reelt problem som virksomhetene/etatene må ta tak i. Vegvesenet er i gang med å sikre tilgang til historiske data for etterprøving av investeringsprosjekt (Statens vegvesen, 2006 s. 10).

I USA har U.S. Department of Commerce Technology Administration fått utført en studie som dokumenterer at en bare innen eiendom (privat og offentlig) tapte \$15,8 billion (96,2 milliarder kroner) i 2002. 2/3 av dette beløpet taper eier/drifter av eiendommene. Det resterende taper BAE-bransjen og andre utførende (NIST, 2004). Det økonomiske potensialet er altså enormt.

---

<sup>38</sup> IAI – International Association for Interoperability, er en organisasjon som jobber for utvikling av internasjonale standarder for utveksling av intelligente data i en bygningsinformasjonsmodell BIM – Building Information Model. Standardene heter IFC – Industry Foundation Classes. De benyttes til å knytte alle slags informasjon til beskrivelsen av det fysiske objektet. Dette tillater at alle data kan utveksles mellom partene og bearbeides av alle som er autorisert til det. Lagringen foretas i en dertil egnet database. Norge er langt fremme i denne utviklingen når det gjelder bygg og anlegg. Statsbygg og Vegvesenet har signalisert at de vil kreve slik datalagring i sine prosjekter i fremtiden (Aarhus 2006) (Christoffersen 2005). Statens bygningstekniske etat er en kunnskapskilde for slik informasjonshåndtering i Norge (Rooth, 2005).

Det er påvist at en mer systematisk definisjon og innsamling av data fra den norske BA-bransjen (trolig overførbart til andre bransjer) vil gi grunnlag for forbedringstiltak ved mer systematisk benchmarking og nøkkeltallsoppfølging på nasjonalt nivå (Karud, 2006). Karud påpeker at det er betydelig potensiale for å utnytte eksisterende kanaler for innsamling av statistikk til nasjonale databaser og bransjestatistikk. Han påpeker (i del 7 av sin doktoravhandling) at det naturligvis vil medføre tilpasningskostnader, men at det også gir nytteverdi både lokalt og sentralt. Automatikk fordrer strenge rutiner for sikkerhet og autentisering. Det kreves god kvalitet på informasjonsstrømmen og systemer som hindrer misbruk. Elektronisk og automatisk rapportering er et sentralt virkemiddel for å redusere byrden på næringslivet (Karud, 2006 s. 199). Det samme gjelder internt i statens virksomheter.

For staten er det mye å hente, også indirekte: Ved å gi KS-rådgivere og forskningsprogram som Concept tilsvarende tilgang kan tidsbruk til informasjonshenting reduseres kraftig. Effekten vil være en bedre og mer effektiv ekstern kvalitetssikringsordning og forskning på store statlige investeringsprosjekt.

Følgende fallgruver fra andre del av kapittelet vil bli eliminert eller sterkt forebygget med en slik løsning:

Tilgangen til relevant og oppdatert informasjon blir meget god til enhver tid. Fallgruver nr. (4), (5), (6) og (15).

Tidsbruken til innhenting av slik informasjon blir meget liten. Problemet med for liten tid er eliminert. Fallgruve nr. (4).

Ingen kilder er lenger unntatt for kvalitetssikring, det eliminerer både manglende kvalitetssikring og faren for å akseptere dårlig informasjon ukritisk. Fallgruver nr. (1), (2), (3), (5), (6) og (17). Obs! også denne nye måten kan bli oppfattet som hevet over feil, noe som kan skape en ny og like stor fallgruve (11).

Å satse på en løsning som dette vil være å ta kommunikasjonen på alvor, noe som eliminerer fallgruve nr. (9). Også her gjelder at å tro at dette er den kommunikasjonsformen som løser alle problemer vil skape en ny farlig fallgruve (11).

Problemene med manglende informasjonsstandarder og mismatch i definisjoner og termer blir eliminert i informasjonsinnhenting for disse problemene må løses som en del av utviklingen av informasjonsstandarder og IDM. Hver etat kan for så vidt ha sin egen terminologi fordi IDM ikke vil ha problemer med å oversette til en felles norm. Det presiseres at problemet ikke elimineres i forhold til kommunikasjon innad i de enkelte fagmiljø og prosjekter i etatene og mellom ulike grupper. Mellommenneskelige relasjoner vil være like utfordrende som før. Her er det bare utvikling av felles definisjoner og felles forståelse av begreper som hjelper. Fallgruver nr. (9), (12), (16) og (17).

Problemet med lagring og gjenfinning av informasjon elimineres. Det er en forutsetning som følger av at det kreves en de facto standard for strukturering av informasjon i databasen. Fallgruver nr. (10) og (17).

Det er mulig å eliminere mange av problemene knyttet til forutsetningene og konsistens fordi det med utgangspunkt i den skisserte løsningen for lagring og behandling av data kan



defineres en rekke automatiske kontroller og konsistens-sjekker. Så lenge det finnes en norm å sammenligne med kan kontroll lett automatiseres. Fallgruve nr. (8).

Gjenstående fallgruver handler om mennesker og forholdene mellom dem. En teknisk løsning slik som beskrevet her vil ikke påvirke disse direkte. Det gjelder fallgruvene nr. (7), (12), (13), (14) og (18). Fordi dette blir en mer nøytral (teknisk basert) måte å sikre god kommunikasjon på, vil også forholdet mellom mennesker påvirkes indirekte fordi klimaet for kommunikasjonen endres. Det forventes positiv effekt fordi kommunikasjonen blir basert på mer korrekt informasjon. Det gjør prosessen mindre utsatt for mistanken om at mennesker med skjult agenda kan ha påvirket tallene. Dette eliminerer en del av årsakene til mistillit mellom partene (12). Derfor vil en slik form for kvalitetssikring styrke tilliten og evnen til samarbeid. På den andre siden kan denne tekniske formen for løsning (IDM) fremstå som en "black box" som er vanskelig å forstå. Dette kan slå ut negativt for tilliten hos enkelte.

Fallgruven mangel på fagkompetanse (7) påvirkes ikke av denne formen for kommunikasjon. Det kreves fortsatt like mye fagkompetanse å bruke informasjonen rett. Tilgangen på fagkompetanse blir ikke påvirket av denne løsningen. Det kommer et nytt indirekte moment inn; tilgangen på spesialkompetanse som kan utvikle denne nye formen for IKT-basert kommunikasjon kan være begrenset.

Alle "teknikalitetene" knyttet til innhenting av god informasjon som i dag koster mye ekstraarbeid og skaper problemer i form av redusert effektivitet og dårlig kvalitet i planleggingen. Disse problemene elimineres eller reduseres kraftig med denne løsningen. Ved å benytte IDM teknologien for å innhente relevant informasjon fra flere kilder kan en effektivt oppnå vesentlig bedre informasjonsgrunnlag for planlegging og beslutningstaking enn det som er mulig i dag. Når underlaget for beslutning er dårlig er det avgjørende å gjøre det beste ut av det som finnes.

Ulempen er selvsagt at det vil koste en god del utviklingsarbeid å komme til det punktet at denne formen for kvalitetssikring er klar for implementering i full skala. Det vil ta noe tid å få på plass forutsetningene for slik informasjonsutveksling og kvalitetssikring, men teknologiske løsninger finnes i dag og er implementert i enkelte industrier. Det er trolig en vesentlig del av fremtidens løsning.

## 8.6 Oppsummering

Vi har sett i første del av kapittelet at det er utviklet en rekke gode rammeverk for planlegging og kvalitetssikring av informasjon i prosjekter. Både offentlige og private virksomheter har utviklet rammeverk som bidrar aktivt til at det er mulig å sikre god kvalitet når beslutningsunderlag utarbeides. I kapittels andre del er vist at det finnes en rekke fallgruver en kan gå i når informasjon skal innhentes og kvalitetssikres i praksis. Et utvalg av viktige fallgruver er

- Manglende bevissthet og kompetanse til kvalitetssikring, dårlig struktur på prosessen.
- Kvalitetssikring blir ikke utført, eller for dårlig utført pga. dårlig tid, dårlig tilgang etc.
- Dårlig kommunikasjon, feil begrepsbruk, mangel på felles språk, standarder etc.
- Valg av urealistiske eller feilaktige forutsetninger, manglende kritisk sans.

Vi har videre konstatert i del tre at det finnes en del gode arbeidsmåter og hjelpemidler som kan bidra, men at det viktigste bidraget fortsatt er vår egen kritiske sans og gode innsats. Gode praksiser kan finnes i både privat og offentlig sektor og er tilgjengelig for den som er villig til å sette seg inn i det. Selv enkle tiltak kan ha stor effekt på kvaliteten av informasjonen som benyttes i planer og beslutningsunderlag. Den siste delen av kapittelet tegner konturene av en framtid der både fallgruvene fra andre del og ressursbruken omtalt i tredje del påvirkes positivt. Det er mulig å lage tekniske løsninger som effektivt skaffer tilgang til bedre informasjon, eliminerer mange av fallgruvene og gir tilgang til denne bedre informasjonen mye raskere og med mye mindre ressursbruk. Men det krever vilje til en innledende investering.

## 8.7 Referanser

- Aarhus, Christian. 2006. *BA-nettverket setter standarden*. Byggeindustrien nr 8, s. 50.
- Austeng, Torp, Midtbø, Helland og Jordanger. 2006. *Usikkerhetsanalyser - Metoder*, NTNU, Concept-programmet. Rapport nr 12.
- Berg, Peder m.fl. 1999. *Styring av statlige investeringer*. Sluttrapport fra styringsgruppen Prosjektet for styring av statlige investeringer. Finans- og tolldepartementet 1999. Ikke offentliggjort.
- Christoffersen, Øyvind 2005. Statsbygg. *Investors and BIM – Requirements from public builders*. Foredrag ved den internasjonale konferansen Building SMART i Oslo.
- Concept, 2006. [www.concept.ntnu.no](http://www.concept.ntnu.no). Internettssidene til forskningsprogrammet Concept.
- Flyvbjerg, Bent. 2006. *From Nobel Prize to Planning Practice. Reference Class Forecasting as Antidote to Optimism Bias and Strategic Misrepresentation*. Universitetet i Aalborg.
- Forsvarsdepartementet. 2004. FOR 2004-09-01 nr 1270: *Anskaffelsesregelverk for Forsvaret (ARF)*.
- Haugane, Øyvind. 2005. *Slik finner vi de rette prosjektene i lønnsomhetsperspektiv*. Foredrag ved NSP Temadag 3. mai 2005; Tidlige beslutninger – hvordan sikre at rett prosjektalternativ blir valgt? Statoil Arena for prosjektutvikling.
- Holte, Eilif, Haanes, Sverre og Larsen, Stein Vegar. 2004. *Beslutningsunderlag og beslutninger i store statlige investeringsprosjekter*. Concept rapport nr 3.
- ISO 10303. *Industrial automation systems and integration - Product data representation and exchange*. ISO technical committee TC 184, Technical Industrial automation systems and integration, *sub-committee SC4* Industrial data.
- Jørgensen, Magne og Teigen, Karl Halvor. 2005. Kan vi unngå å si at ”så å si helt sikkert” bare betyr ”60% sikkert”? *Prosjektledelse* nr 2, 2005.
- Karud, Ole Jørgen. 2006. *Datafangst til prestasjonsmålinger, benchmarking og analyser i norsk bygge- og anleggsnæring*. Doktoravhandling ved NTNU, 2006:73.
- Klakegg, Ole Jonny. 1993. *Trinnvis-prosessen*. Institutt for bygg- og anleggsteknikk, NTNU.

- Klakegg, Ole Jonny. 2000. *Håndbok 217 Kvalitetssikring av kostnadsoverslag Anslag-metoden*. Statens vegvesen.
- Klakegg m. fl. 2003. *Felles begrepsapparat KS2*. Finansdepartementet, Concept-programmet.
- Langlo, Jan Alexander, Andersen, Bjørn og Fagerhaug, Tom. 2006 *Process Analysis and Process Improvement as Tools for Analysing and Improving Interorganizational Processes in Offshore Value Networks*. Paper ved ProMAC konferansen 2006, Sidney, Australia.
- Lovullo, Dan, og Kahneman, Daniel. 2003. Delusions of Success; How optimism Undermines Executives' Decisions. *Harvard Business Review*, july edition.
- Magnussen, Ole Morten, and Nils O.E. Olsson. 2006. Comparative Analysis of Cost Estimates of Major Public Investment Projects. *International Journal of Project Management* 24 (4):281-288.
- NIST, 2004. *Cost Analysis of Inadequate Interoperability in the U.S. Capital Facilities Industry*. NIST GCR 04-867. U.S. Department of Commerce Technology Administration.
- OGC, 2006. [www.ogc.gov.uk](http://www.ogc.gov.uk). Internettssidene til Office of Government Commerce.
- PROPS Online 2003. [http://www.semcon.se/spm/eng/model/props\\_intro\\_en/intro.props.asp](http://www.semcon.se/spm/eng/model/props_intro_en/intro.props.asp) Ericsson.
- Rooth, Øyvind. 2005. *Presentation of e-Government in Singapore and Norway – eSubmission and code checking – ePermit - GIS – State of the art and future development*. Foredrag på konferansen Building SMART. Statens bygningstekniske etat.
- Statens vegvesen. 2006. *Veileder for etterprøving av prissatte konsekvenser av store vegprosjekt*. UTB2006/02.
- Teigen, Karl Halvor. 2006. Universitetet i Oslo. *Skjønn eller skivebom: Hvordan vi bedømmer usikkerhet*. Foredrag på NSP Temadag i Oslo 30. mars.
- Wix, Jeffrey. 2005. *Information Delivery Manual; Enabling Information Exchange in AEC/FM Business Processes*. Jeffrey Wix Consulting Ltd.

## 9 SKJEVHETER I FAGFOLKS SKJØNN: HVORDAN KAN BESLUTNINGSTAKING FORBEDRES?

GEIR KIRKEBØEN\*

”Expert judgments have been worse than those of the simplest statistical models in virtually all domains that have been studied”

(Camerer & Johnson, 1991, 203)

“Professionals are not stupid, but they are human”

(Bazerman, 2006)

### 9.1 Innledning

Folks skjønn er ikke perfekt. Til og med fagfolk gjør feil. Og mer interessant, feilene i fagfolks skjønnsmessige bedømminger og beslutninger er ikke tilfeldige. Fagfolks skjønn avviker nemlig *systematisk* fra normative, ”rasjonelle” standarder. Systematiske skjevheter gir pr definisjon redusert beslutningskvalitet og kan følgelig ha alvorlige konsekvenser. Men det at skjevhetene er systematiske gir også håp. Kan vi identifisere skjevhetene og deres årsaker, bør det være mulig å motvirke dem.

Skjevhetene i folks beslutningstaking, hvordan vi avviker fra normative standarder og hvorfor, er godt kartlagt i faget beslutningspsykologi. I dette kapitlet gir jeg et innblikk i beslutningspsykologiens forståelse av mennesket som beslutningstaker, forklarer ulike typer skjevheter som systematisk reduserer beslutningers kvalitet, samt at jeg gir noen eksempler på effektive strategier for å forbedre fagfolks beslutningstaking. Jeg starter med å gi noen glimt fra beslutningspsykologiens (for)historie.

### 9.2 Fra L’HOMME CLAIRE til ’begrenset rasjonalitet’

Hvor mange dukater vil du betale for å delta i følgende spill: Du skal kaste kron og mynt inntil du får mynt. Får du mynt på første kast, så får du utbetalt 2 dukater. Mynt på andre kast gir deg 4 dukater. Mynt på tredje kast gir 8 dukater. osv.?

I 1713 stilte matematikeren Nicolas Bernoulli dette spørsmålet til både lek og lærd i St Petersburg. Ingen ville betale mer enn 10 dukater for å få spille Bernoullis spill. Dette brøt

---

\* Geir Kirkebøen er professor ved Psykologisk institutt, Universitetet i Oslo

dramatisk med en akseptert regneregul for å bestemme ”forventet utbetaling” (*expected value*), dvs. den gjennomsnittlige utbetalingen pr. spill som du kan forvente dersom du spiller spillet utallige ganger, fra et sjansespill. Denne regelen, som i 1713 hadde vært kjent og brukt av gamblere i århundrer, sier at vi finner den forventede utbetalingen ved, for hvert mulig utfall av spillet, å multiplisere utfallets utbetaling med utfallets sannsynlighet, for så å summere over alle mulige utfall. Hvilken forventet utbetaling gir Bernoullis spill ut fra denne regelen?

Jo, mulige utfall av spillet er: mynt på første kast, mynt på andre kast, mynt på tredje kast osv. Det er altså et uendelig antall mulige utfall av spillet. For hvert eneste utfall vil utbetalingen utfallet gir multiplisert med sannsynligheten for utfallet, være:  $2^N$  (Utbetaling) \*  $\frac{1}{2^N}$  (Sannsynlighet) = 1 dukat. (N angir hvilket kastnummer som først gir mynt.) Ut fra regneregelen ovenfor er da den forventede utbetalingen av Bernoullis spill:  $1 + 1 + 1 + \dots$  i det uendelige, altså forventet utbetaling er uendelig stor. Likevel, ingen var altså villig til å betale mer enn 10 dukater for å få delta i spillet. Dette er siden omtalt som ”St Petersburg”-paradokset.

Dette var nemlig et paradoks. Det var et paradoks i forhold til en antagelse om rasjonalitet som hadde vært rådende siden antikken. Ut fra denne antagelsen ble *l’homme clair* - mennesket på sitt beste - sett på som normen på rasjonell eller optimal tenkning. Man antok at lovene som de ”klareste hoder” resonnererte etter var identiske med matematikken og logikkens lover, at de siste var å betrakte som avledet fra de første. Men selv ”klare hoder” avvek altså dramatisk fra en akseptert normativ regel på Bernoullis spørsmål. Dette er én grunn til at St.Petersburg-paradokset er historisk viktig. Det utfordret den klassiske antagelsen om rasjonalitet, og åpnet slik opp for det nå bredt aksepterte skillet mellom normativ og deskriptiv beslutningsteori, altså skillet mellom hvordan beslutninger ideelt sett bør tas og hvordan folk, inkludert fagfolk, faktisk tar dem.

St.Petersburg-paradokset inspirerte dessuten begrepet *utility* (”nytte” eller ”personlig verdi”), et sentralt begrep i all senere beslutningsteori. Daniel Bernoulli, nevøen til Nicolas, lanserte i 1738 dette begrepet i et forsøk på å forklare St. Petersburg-paradokset. Han hevdet at det ikke er forventet utbetaling (objektiv verdi) som er avgjørende for hvilket spill eller alternativ man velger, men derimot den forventede *subjektive* verdi som spillet gir. To personer med ulike ønsker og ulik økonomi vil derfor ikke nødvendigvis vurdere identiske spill likt. Bernoulli mente altså at et rasjonelt eller optimalt valg er optimalt relativt til en beslutningstaker. Han antok videre at rasjonelle beslutningstakere velger det alternativet som maksimerer forventet nytte, altså den forventede subjektive verdi, av valget. Med disse to antagelsene la Daniel Bernoulli grunnlaget for moderne beslutningsteori. Hans antagelser utgjør første versjon av såkalte ”forventet nytte”-teorier.

Vi tar et langt sprang fram til 1947. Da publiserte matematikeren Von Neumann og økonomen Morgenstern boka *Theory of games and economic behavior*. En sentral antagelse i økonomisk teori hadde lenge vært at økonomiske aktører opptrer rasjonelt. Men hva betyr det? Von Neumann og Morgenstern ga i sin bok et presist svar på dette spørsmålet. De spesifiserer i form av en matematisk teori hva det vil si å velge rasjonelt. Kjernen i deres *Expected Utility* (EU) -teori er en liten mengde av aksiomer. Von Neumann og Morgenstern viste at dersom man tar valg i overensstemmelse med disse aksiomene, så kan man alltid knytte en bestemt nytte (*utility*) til hvert mulig valgutfall på en slik måte at et alternativ foretrekkes framfor et annet, hvis og bare hvis alternativets forventede nytte er større enn andre alternativets forventede nytte.

EU-teorien forutsetter at sannsynligheten for alle mulige valgutfall er objektivt gitt. Savage (1954) modifiserte teorien slik at sannsynlighetene for utfall også kan være subjektivt anslått. Resultatet, Savage såkalte *Subjective Expected Utility* (SEU)-teori, kan vi si generaliserer EU-teorien fra en teori om valg under objektiv usikkerhet (risiko) til en teori om valg under usikkerhet generelt.

(S)EU-teorien presenterer altså en kompleks teori om rasjonelle valg i form av et lite sett av aksiomer, hver for seg enkle og avgrensede. Disse aksiomene kunne gjøres til gjenstand både for normative og deskriptive undersøkelser: Er aksiomet rimelig som et normativt prinsipp? Tar folk sine valg i overensstemmelse med prinsippet? Eller mer overordnet, tar folk beslutninger slik SEU-teorien som helhet antar?

Hva krever så SEU-teorien av et rasjonelt valg? Jo, man må kartlegge alle valgalternativer, bestemme alle mulige utfall av hvert alternativ, vurdere sannsynligheten for hvert av disse utfallene, klargjøre konsekvensene av hvert av dem, anslå hvilken nytte konsekvensene av hvert utfall har relativt til hverandre, og så avslutte med å velge det alternativet som gir høyest forventet nytte.

Tar folk beslutninger i overensstemmelse med SEU-teorien? Herbert Simon, en senere nobelprisvinner i økonomi, var på 50-tallet den mest markante kritikeren av antagelsen om at SEU-teorien gir en relevant beskrivelse av menneskelig rasjonalitet. I artikkelen *A behavioral model of rational choice*, lanserer Simon (1955) sitt begrep om ”begrenset rasjonalitet” (*bounded rationality*). Han argumenterer her for at ”the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed [by humans]” (99).

Simon (1990) mente SEU-modellen var totalt urealistisk som en modell på menneskelig rasjonalitet: ”the SEU-model is a beautiful object deserving a prominent place in Plato’s heaven of ideas (...) But vast difficulties make it impossible to employ it in any literal way in making actual human decisions” (194). I følge Simon har vi ofte ikke tilgang til all den informasjonen som vi trenger for å ta valg slik SEU-teorien foreskriver. Vår evne til å behandle informasjon er dessuten så begrenset at vi bare unntaksvis kan følge den oppskriften på et rasjonelt valg som SEU-modellen gir.

Simon (1990) lanserte en alternativ måte å tenke rundt beslutningstaking på. Han hevdet at ”human rational behavior (...) is shaped by a scissors whose two blades are the structure of task environment and the computational capabilities of the actor” (7) Rasjonell atferd må forstås ut fra begge ”saksens blader”, altså både ut fra vår begrensede kognitive kapasitet og ut fra (informasjonsstrukturen i) omgivelsen der valget tas. Det ”kognitive bladet” krever at modeller av hvordan vi tar beslutninger må baseres på realistiske antagelser om vår kognitive kapasitet: “However adaptive the behavior of organisms in learning and choice situations, this adaptiveness falls far short of the ideal of ‘maximizing’ in economic theory. Evidently, organisms adapt well enough to *satisfice*; they do not, in general, *optimize*” (Simon, 1956, 129; *min uthevelse*).

Simon avviser altså den klassiske antagelsen om *l’Homme Claire* som normen på rasjonalitet. Hans ”begrenset rasjonalitet”-perspektiv på beslutningstaking setter agendaen for mye av senere beslutningspsykologi, nemlig å avdekke hvilke *satisficing*-strategier vi bruker i ulike situasjoner og hvordan vi avviker fra normative modeller som SEU-teorien.

### 9.3 Beslutningspsykologi: Noen sentrale innsikter

Tversky og Kahneman (1981) stilte én gruppe folk ovenfor følgende valg:

Tenk deg at USA forbereder seg på utbruddet av en uvanlig asiatisk sykdom som forventes å drepe 600 mennesker. To alternative programmer for å bekjempe sykdommen er foreslått. Anta at det eksakte vitenskapelige estimatet av konsekvensene av de to programmene, er som følger:

Program A: Sikkert at 200 mennesker vil reddes

Program B: 1/3 sjans for at 600 vil reddes og 2/3 sjans for at ingen vil reddes.

Hvilke av de to programmene ville du foretrukket?

De fleste (72%) foretrakk program A. En annen gruppe ga Tversky og Kahneman valget;

Program A': Sikkert at 400 mennesker vil omkomme

Program B': 1/3 sjans for at ingen vil omkomme og 2/3 sjans for at 600 vil omkomme.

Da foretrakk de fleste (78%) program B'.

Ved litt ettertanke ser man at de to valgene er identiske. Program A har samme effekt som program A', og B har samme forventede effekt som B'. Likevel, foretrekker altså det store flertallet program A i det første valget og program B' i det andre. I studien var forsøkspersonene psykologistudenter. En liknende studie der fagfolk (leger) var forsøkspersoner, ga tilsvarende resultat (McNeil et al. 1982). Hva er forklaringen?

Før "asiasyke"-studien hadde psykologisk beslutningsforskning i stor grad blitt neglisjert av økonomer. Simons "begrenset rasjonalitet"-perspektiv hadde i liten grad påvirket økonomenes grunnantagelse om at den enkelte aktør opptrer rasjonelt. Beslutningsatferd som Simon så som uttrykk for at vår rasjonalitet er begrenset, ble gjerne forklart (bort) som en rasjonell tilpasning til kostnadene ved å søke etter informasjon.

Funnene fra "asiasyke"-studien var det verre å (bort)forklare. Det er nemlig sentralt i SEU-teorien at preferanser ikke skal påvirkes av irrelevante forhold som måten identiske valgalternativer beskrives på. Dette var (og er) også et essensielt aspekt ved det rasjonalitetsbegrepet man lenge hadde operert med i økonomisk teori. Tversky og Kahnemans "asiasyke"-studie viste tydelig at denne antagelsen ikke holder. Positive og negative formuleringer av samme saksforhold får her dramatisk innvirkning på beslutningstaking. De hadde dessuten selv et par år tidligere publisert en helhetlig psykologisk teori om beslutningstaking, den såkalte prospektteorien (Kahneman og Tversky, 1979), som predikerte funnene i "asiasyke"-studien.

#### Prospektteorien

Kahneman og Tverskys prospektteori spesifiserer hvordan beslutningstakere systematisk avviker fra (S)EU-teorien. Mange av enkeltelementene i teorien hadde vært kjent i lang tid før teorien ble publisert i 1979. Kahneman og Tverskys store bidrag var at de maktet å formulere

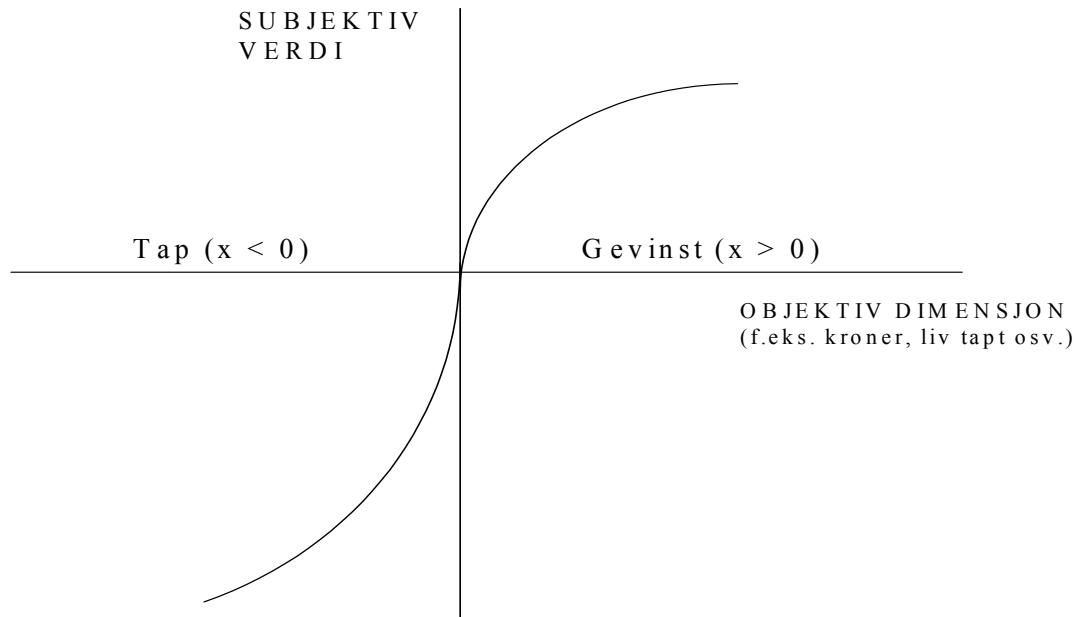
en presis matematisk teori på samme form som SEU-teorien, én teori som forklarte svært mange av disse avvikene. I ettertid beskriver de strategien de fulgte i utformingen av teorien slik:

“The theory that we constructed was as conservative as possible (...) We did not challenge the philosophical analysis of choices in terms of beliefs and desires that underlies utility theory, nor did we question the normative [status of] models of rational choice (...) The goal (...) was to assemble the minimal set of modifications of expected utility theory that would provide a descriptive account” (Kahneman & Tversky, 2000, x)

Hovedgrunnen til at de var ”så konservative som mulig” var selvfølgelig at de så på økonomer som en viktig målgruppe. Det var viktig at økonomene kunne kjenne seg igjen i teorien. Og det gjorde de. Artikkelen der prospektteorien først ble publisert, er den nest mest siterte artikkelen i økonomisk litteratur (Wu et al, 2004). Og Kahneman fikk i 2002 nobelprisen i økonomi, i stor grad på grunn av prospektteorien.

De tre sentrale komponentene i prospektteorien er en *verdifunksjon*  $v(x)$ , en *beslutningsvektfunksjon*  $\pi(p)$  og idéen om et *referansenivå* som angir hva beslutningstakeren (i valgsituasjoner) opplever som skillet mellom tap og gevinst.





$$v(x) = \begin{cases} x^\alpha & \text{if } x > 0 \\ -\lambda (-x^\lambda) & \text{if } x < 0 \end{cases} \quad (\text{typisk } \alpha = 0.88 \text{ og } \lambda = 2.25)$$

Prospektteoriens verdifunksjon angir hvordan beslutningstakere subjektivt verdsetter objektive verdier, eller mer korrekt, endringer i objektive verdier. Verdifunksjonen har fire karakteristiske trekk. Den inkorporerer for det første, Bernoullis idé om at vi ”mettes av rikdom”, altså at den subjektive verdien som vi tillegger en økning i objektiv verdi avtar med hvor mye vi har fra før. Det andre karakteristiske trekket ved prospektteoriens verdifunksjon er at den antar en tilsvarende ”metning” ved opplevd tap. For det tredje viser funksjonen at ubehaget ved å tape en bestemt verdi (for eksempel et pengebeløp) oppleves som drøyt to ganger større enn behaget ved å vinne den tilsvarende verdien (beløpet). For det fjerde antar prospektteorien at når vi tar beslutninger så opplever vi tap og gevinst ut i fra et subjektivt referansenivå, det vil si det nivået som skiller mellom hva beslutningstakeren opplever som henholdsvis tap og vinst.

David Bernoullis (1738) feil var i følge Kahneman og Tversky at han antok at størrelsen på forventet nytte alene var avgjørende for hvordan folk velger. Økonomer og beslutningsteoretikere hadde gjentatt denne feilen helt fram til de publiserte sin prospektteori i 1979. Prospektteorien avviser at det er størrelsen på den forventede nytte av ulike valgalternativ som er avgjørende for folks preferanser. Det er *endringen* i forventet nytte som bestemmer folks valg. Teorien antar nemlig som nevnt at preferanser er relative til et *referansenivå*, altså til det nivået som beslutningstakeren opplever som skillet mellom tap og vinst i valgsituasjonen.

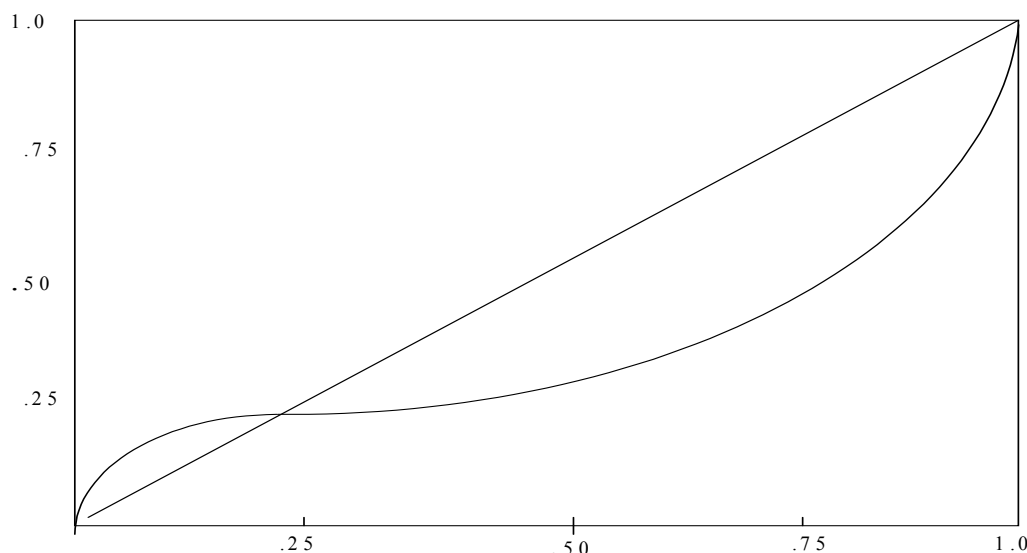
Funnet fra ”asiasyke”-studien ovenfor forklarer prospektteorien med at beskrivelsene av alternativene i de to valgsituasjonene påvirker folks referansenivå, altså hvordan de forstår og

strukturerer (*framer*) valget i termer av tap og gevinst. Det første valget strukturerer folk slik at referansenivået er at 200 liv reddes. Følgelig opplever man at man risikerer å tape disse 200 dersom man velger program B. I det andre valget derimot, forstår folk typisk valget slik at referansenivået er at ingen omkommer. De opplever da at de kan unngå å tape 400 liv ved å satse på alternativ B'. I og med at verdifunksjonen sier at tap oppleves mer enn to ganger så negativt som vinst oppleves positivt, ser man direkte av verdifunksjonen hvorfor folk typisk svarer som de gjør i "asiasyke"-studien.

Thalers (1980) påvisning av den såkalte "gaveeffekten" (*endowment effect*) var et annet overbevisende tidlig eksempel på at vi systematisk avviker fra SEU-teorien på en måte som predikeres av prospektteorien. Thaler fant at det maksimumsbeløpet folk er villig til å betale for å skaffe seg et gode vanligvis er mye mindre enn det minimumsbeløpet de krever for å skille seg av med det samme godet. Salgsprisen overgår kjøpsprisen med en faktor litt større enn to, ganske nøyaktig hva prospektteoriens verdifunksjon predikerer.

En utbredt markedsføringsstrategi i varehandelen er å selge varer med full returrett. Elkjøp tilbyr for eksempel hele 30 dagers full returrett. Kundene opplever dette som gunstig, men de benytter seg svært sjelden av returretten. Prospektteorien og gaveeffekten forklarer hvorfor. Når vi mottar en gave eller kjøper en vare, så skifter vi referansenivå til å inkludere varen. Det å bytte varen inn igjen, oppleves da som et tap. Det er for øvrig en rekke studier og eksempler som viser relevansen av prospektteorien i den virkelige verden (se f.eks. Camerer, 2001).

Prospektteorien er altså som SEU-teorien en teori om valg under usikkerhet. SEU-teorien skiller imidlertid ikke mellom estimerte sannsynligheter og hvordan sannsynligheter subjektivt oppleves. Prospektteoriens beslutningsvektfunksjon gjør nettopp det. Den spesifiserer de systematiske "skjevhetene" i folks (subjektive) oppfatning av (objektiv) usikkerhet.



$$\Pi(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}$$

$\gamma = 0.69$  for negative vekter og  $0.61$  for positive (I SEU er  $\Pi(p) = p$ )

Figuren som altså visualiserer prospektteoriens beslutningsvektfunksjon, viser for det første at beslutningstakere gjerne tillegger små sannsynligheter en uforholdsmessig stor vekt (den såkalte "mulighetseffekten"). For sannsynligheter rundt 0.2 derimot, er det vanligvis godt samsvar mellom faktisk usikkerhet og hvordan denne oppleves. Fra sannsynligheter mellom 0.2 til 0.8 er det en tendens til subjektivt å undervurdere usikkerhet. Men så, når vi beveger oss fra at noe er litt usikkert til at det blir helt sikkert, så tillegger vi gjerne små endringer i (objektiv) sannsynlighet en mye større subjektiv vekt (den såkalte "sikkerhetseffekten").

Prospektteoriens beslutningsvektfunksjon forklarer en rekke beslutningsavvik fra SEU-teorien. Et eksempel er et fenomen som lenge hadde vært et paradoks sett ut fra SEU-teorien og andre normative teorier, nemlig folks store forkjærlighet for å delta i lotterier. Dette forklares av "mulighets"-effekten, altså den relativt store subjektive vekten som vi gjerne tillegger det at en hendelse er mulig, selv om den er ekstremt usannsynlig.

### Heuristikker og skjevheter

Blant kvinner på 40 år har 1% brystkreft. Dersom en kvinne har brystkreft, så er sannsynligheten 80% for at hun vil teste positivt på mammografi. Dersom en kvinne ikke har brystkreft, så er sannsynligheten 10% for at hun likevel vil teste positivt. Hva er sannsynligheten for at en kvinne på 40 år som tester positivt faktisk har brystkreft?

David Eddy (1982) stilte tidlig på 80-tallet dette spørsmålet til leger. Det store flertallet av legene (95%) svarte at sannsynligheten for at kvinnen har brystkreft er mellom 70 og 80 %. Det riktige svaret er 7.5 %. Hvordan kunne velutdannede leger svare så fryktelig galt på en type spørsmål som de jevnlig stilles overfor i sin praksis?

Eddy fant svaret i Tversky og Kahnemans (1974) tidlige studier av resonnering med usikkerhet. Sannsynlighetsregningens regler angir hvordan vi optimalt sett bør resonnerere med usikkerhet. Men å følge disse reglene blir raskt for komplisert. Vi tvinges til, i Simons terminologi, å *satisfice*. Kahneman og Tversky påviste at folks resonnering med usikkerhet styres av noen få "satisficing"-strategier eller hva de kalte heuristikker.

Legene ovenfor skulle vurdere sannsynligheten for om en kvinne som tester positivt også har brystkreft. Kahneman og Tversky fant at når vi foretar denne type sannsynlighetsvurderinger, altså vurderinger av om en person (hendelse eller objekt) hører hjemme i en bestemt kategori eller klasse, så vurderer vi ofte dette ved å bruke såkalt *representativitetsheuristikk*, dvs. at vi svarer ut fra hvor typisk eller representativ vi vurderer at personen er for kategorien. Legene i Eddys studie svarte altså at det er 70-80% sjanse for at kvinner som tester positivt har brystkreft. Det tyder på at legene forvekslet hvor sannsynlig det er at kvinnene har kreft, med hvor typisk eller representativt det er at kvinner som har kreft tester positivt (som altså 80% av kvinner med kreft gjør).

Kahneman og Tversky understreker at representativitetsheuristikk både forenkler sannsynlighetsvurderinger og ofte også fungerer bra. Bruk av representativitetsheuristikk fører imidlertid til at folk i noen tilfeller avviker dramatisk fra sannsynlighetsregningens regler. Normativt sett burde legene resonneret i overensstemmelse med Bayes' formel. Formelen sier at når man skal vurdere sannsynlighet for brystkreft gitt positiv mammografi, så må man også ta hensyn til forholdet mellom baseratene, altså forholdet mellom hyppigheten av hhv. brystkreft og positiv mammografitest blant kvinner på 40 år. Dette forholdet er altså litt mindre enn 1/10, det vil si at det er ti ganger flere kvinner som tester positivt enn det er

kvinner med kreft. Det er nettopp når forholdet mellom baseratene er skjevt, at bruk av representativitetsheuristikk i vurderinger av kategoritilhørighet gir systematiske avvik fra sannsynlighetsregningens regler.

Hvor stor prosentdel av husarbeidet gjør du hjemme? Eller hvor stor prosentdel av arbeidsmengden i det siste prosjektet du var med i, sto du for? Når man stiller ektefeller og (andre) prosjektdeltagere hver for seg denne type spørsmål, så har prosentsummen av utført arbeid en tendens til å bli langt større enn 100. Hvorfor?

Bruk av tilgjengelighetsheuristikk (*availability heuristics*), en annen hovedstrategi Tversky og Kahneman (1974) mener styrer vår resonnering med usikkerhet, kan forklare dette. Med tilgjengelighetsheuristikk menes tendensen til at vurderinger av hvor hyppig noe (X) forekommer (i forhold til noe annet) påvirkes av hvor lett tilgjengelig X er, dvs. hvor lett X er å legge merke til, huske eller forestille seg.

Det er klart at du registrerer og husker langt lettere arbeid som du selv har utført enn det arbeidet som kona eller kollegaer har gjort. Ditt eget arbeid er av mange grunner langt mer "tilgjengelig" for deg når du i ettertid tenker over hvem som gjorde hva. Da blir også dine anslag av hvor stor del av arbeidet du selv har utført lett overestimert, og det altså uten at du selv er "skyld" i denne feilvurderingen.

Tilgjengelighetsheuristikk er ofte forklaringen på *How we know what isn't so*. I boka med denne fantastiske tittelen, gir Gilovich (1991) en rekke eksempler på hvordan blant annet bruk av tilgjengelighetsheuristikk gjør at både lek og lærd kan være overbeviste om sammenhenger som slett ikke fins. Det er for eksempel en utbredt myte at par som tror de ikke kan få barn, lettere får egne barn etter å ha adoptert. En vanlig forklaring er at dette skyldes at paret etter adopsjon stresser mindre. Klinisk forskning har imidlertid vist at fruktbarheten slett ikke forandrer seg. Hvorfor er det likevel så mange som tror det? Bruk av tilgjengelighetsheuristikk kan forklare dette: Vi legger veldig godt merke til at "ufruktbare" par får egne barn etter at de har adoptert. Informasjonen som viser at par som ikke adopterer får barn like hyppig etter å ha forsøkt like lenge, er derimot sjelden tilgjengelig for oss.

### **Følelser og valg: Affektheuristikk**

Hvordan følelser påvirker valg har ingen (eksplisitt) plass i SEU-teorien. Heller ikke Simons "begrensede rasjonalitet"-perspektiv vektla følelsers betydning. Vi så dessuten ovenfor at Kahneman og Tversky bestrebet seg på å gi prospektteorien samme form som SEU-teorien. Dette bidro til at hvordan følelser påvirker beslutningstaking heller ikke fikk noen plass i prospektteorien. De siste to tiåra har imidlertid måten følelser påvirker beslutningstaking blitt et sentralt og stort tema i beslutningspsykologi (f.eks. Rottenstreich & Shu, 2004). Jeg skal her avgrense meg til se på en grunnleggende form for affektiv påvirkning av beslutningstaking, nemlig bruk av såkalt "affektheuristikk" (Slovic et al., 2002).

En ny person er ansatt i firmaet. Sjefen sier gjerne at han etter en grundig vurdering har valgt den beste søkeren. Psykologisk forskning gir grunn til å tro at den som ansettes, heller er søkeren som sjefen umiddelbart likte best. Ansettelse, som andre valg, påvirkes nemlig i stor grad av affektfølelser.

Mye forskning tyder på at vår første reaksjon på stimuli er affektiv (Zajonc, 1980). Når du treffer en person, ser et hus, begynner å lese et bokkapittel osv., så danner du deg umiddelbart

en oppfatning om dette er noe du liker eller misliker. Reaksjonen er gjerne ubevisst, du har ofte problemer med å begrunne følelsen. Men den er der. Og den styrer hvordan du videre oppfatter og vurderer personen, huset eller kapittelet.

Kahneman (2003) framhever (i en artikkel basert på hans nobelforedrag i 2002) affektheuristikk som den tredje hovedheuristikken, altså sammen med representativitets- og tilgjenglighetsheuristikk, som styrer våre vurderinger og beslutningstaking. Han presenterer for øvrig i denne artikkelen en ny forståelse av hva heuristisk tenkning generelt består i, nemlig at man (uten å være klar over det) ganske enkelt erstatter et (mer) komplisert spørsmål med et enklere spørsmål. For eksempel, når vi blir spurt hva sannsynligheten er for at X hører hjemme i kategorien Y (X kan for eksempel være en kvinne som har testet positivt på mammografi, og Y kategorien brystkreft) så svarer vi i stedet på spørsmålet: Hvor typisk eller representativ er X (positiv test) for Y (brystkreft)? Da bruker vi representativitetsheuristikk. Tilsvarende innebærer bruk av affektheuristikk at vi (ubevisst) erstatter kompliserte spørsmål som: ”Hvem er best egnet?”, med det langt enklere spørsmålet: ”Hvem liker jeg best?”

### **System1 vs. system 2: Intuisjon vs. analyse**

Svar raskt på denne oppgaven:

En ball og et balltre koster 110 kroner.

Balltreet koster 100 kroner mer enn ballen.

Hvor mye koster ballen?

Heuristikkene omtalt ovenfor ble lenge sett på som enkeltstående kognitive strategier. Nå er det bredt akseptert å betrakte bruk av slike heuristikker som karakteristisk for én av to hovedmåter vi resonnerer på, nemlig som såkalt ”System 1”-tenkning. Denne type tenkning er assosiativ, affektiv, i stor grad automatisert, implisitt (utilgjengelig for introspeksjon), rask (parallell informasjonsbehandling), og den krever små kognitive anstrengelser.

”System 2”-tenkning, motsetningen til ”system 1”-tenkning, er affektiv nøytral, analytisk, veloverveid, langsom (seriell informasjonsbehandling), kontrollert, regelstyrt, fleksibel og krever stor kognitiv anstrengelse.

Hva var forresten prisen på ballen ovenfor? De fleste har en umiddelbar tendens til å svare 10 kroner. Grunnen er at 110 naturlig kan deles opp i 100 og 10 (og selvfølgelig også at det er naturlig at ballen koster noe i nærheten av 10 kroner.) Svarte du feil, så var det fordi du ikke ”sjekket” svaret. Du ga det svaret som ”system 1” umiddelbart foreslo, uten å koble inn ”system 2”. Dette ”systemet” antas nemlig å ha som hovedfunksjon å overvåke og kontrollere raske ”system 1”-responser.

Dersom du svarte feil, kan du trøste deg med at du er i godt selskap. Over halvparten av studentene ved Princeton svarte også feil. Det at så vidt mange svarer galt på denne enkle oppgaven, indikerer hvor slapt responsene fra ”system 1” (oftest) overvåkes av ”system 2”: Vi unngår å tenke analytisk og veloverveid. Det er for anstrengende. I stedet er vi ofte tilbøyelig til å stole på de tilsynelatende plausible svarene som ”system1” raskt gir. Ut fra et slikt ”system 1”/”system 2”-perspektiv kan vi si at intuitive vurderinger svarer til vurderinger som ikke er modifisert av ”system 2.”

I eksemplene så langt er intuisjon koblet til dårlig utførelse, men intuitiv tenkning kan også være kraftfull og nøyaktig. Gitt visse forutsetninger (jf. betingelsene for å lære av erfaring nedenfor), så kan høyt ferdighetsnivå erverves gjennom praksis, og utførelsen kan være hurtig og uanstrengt. Stormestere i sjakk er ett eksempel på dette. Kvaliteten på deres spill i simultansjakk og lynsjakk er høy selv om spillingen i stor grad er automatisert. Og studier av enkelte typer profesjonsutøvere kan tyde på at erfarne beslutningstakere gjerne tar bedre beslutninger når de stoler på sine intuisjoner og ikke går inn i detaljert analyse (Klein, 1998).

## 9.4 Beslutningsskjevheter – en oversikt

En beslutningsskjevhet (*decision bias*) kan vi definere som en tenkemåte som gjør at vi systematisk avviker fra hva vi ovenfor omtalte som rasjonell beslutningstaking. Slike tenkemåter gir altså per definisjon redusert beslutningskvalitet. Mengden påviste beslutningsskjevheter er mange, og raskt voksende. Det er ikke mulig her å gi en fullstendig oversikt. Vi skal nøye oss med å se på noen vanlige slike skjevheter, klassifisere disse i noen hovedtyper, samt få fram de grunnleggende årsakene til dem.

### Skjevheter knyttet til prosessering av informasjon

Vi har ovenfor sett eksempler på noen former for beslutningsskjevheter. I Eddys studie så vi at legene systematisk ikke tok hensyn til relevante baserater når de skulle anslå sannsynlighet for brystkreft på grunnlag av mammografiresultat. *Insensitivitet til baserater* er én av flere systematiske skjevheter i sannsynlighetsvurderinger (av kategoritilhørighet) som skyldes bruk av representativitetsheuristikk. En annen skjevhet som typisk følger slik ”representativ tenkning” er den såkalte ”konjunksjonsfeilen” (*conjunction fallacy*) dvs. at vi vurderer hendelsene eller egenskapene A & B som mer sannsynlige enn bare hendelsen eller egenskapen A. Ett eksempel, Tversky & Kahneman (1983) ga folk en beskrivelse av en person kalt Linda, og ba folk så rangere hva som ut fra beskrivelsen var mest sannsynlig av blant annet utsagnene ”Linda er bankfunksjonær” og ”Linda er bankfunksjonær og aktiv i kvinnebevegelsen.” Flertallet svarte at det siste utsagnet var mer sannsynlig enn det første, trolig fordi beskrivelsen av Linda var representativ for feminister.

Bruk av tilgjengelighetsheuristikk gir også opphav til predikerbare skjevheter. Vi så at bruk av denne heuristikken kan forklare hvorfor folk feilaktig tror adopsjon øker hyppigheten av fødsler blant par som har hatt problemer med å få barn. Et annet eksempel. Når amerikanere ble spurt om hva de vurderer som den hyppigste dødsårsaken i USA av narkotikabruk, våpenbruk, trafikkulykker, dårlig diett /fysisk inaktivitet og tobakk, ga svært få korrekte svar. Det riktige svaret er den motsatte rekkefølgen som dødsårsakene ble listet opp i ovenfor. I begge disse eksemplene overvurderer folk hyppigheten av fenomenene vi (og media) lettest legger merke til. Hvor lett/vanskelig et fenomen er å forestille seg eller å huske gir tilsvarende systematiske skjevheter i våre hyppighetsvurderinger som følge av at vi bruker tilgjengelighetsheuristikk.

Tilgjengelighets- og representativitetsheuristikk er strategier som forenkler prosessering av informasjon når vi vurderer sannsynligheter. Den tredje forenklingsstrategien som Tversky og Kahneman (1974) framhever er ”forankring og justering” (*anchoring and adjustment*). Når vi estimerer så gjør vi dette ofte i to trinn. Vi tar utgangspunkt i en mer eller mindre begrunnet verdi (ankeret) ut fra informasjon som er umiddelbart tilgjengelig for oss, for deretter å justere oss fram til et plausibelt estimat fra denne ”ankerverdien.” Det har vist seg at selv helt

vilkårlige ”ankere” kan ha stor innvirkning på hvilket endelig estimat vi lander på. I en klassisk studie som illustrerer dette, delte Tversky og Kahneman (1974) folk i to grupper. Begge gruppene skulle anslå hvor mange prosent av FNs medlemsland som er afrikanske. Men før folk kom med sine anslag, ble de bedt om å sende av gårde en rullettkule. Rulletten var manipulert slik at for folk i den ene gruppa så stoppet kula på 65, for folk i den andre stoppet den på 10. Snittsvaret på hvor mange av FNs medlemsland som er afrikanske var 45% i gruppa der kula stoppa på 65, og 25% i gruppa der kula stoppet på 10. Selv om ”ankeret” her altså er helt uten relevans for det som skal estimeres, har det en markant innflytelse på folks estimater. Slik ”forankring” er et meget robust fenomen som også påvirker fagfolks estimater (se f.eks. Chapman og Johnson, 2002).

En beslutningsskjevhet som ofte er medvirkende årsak til at prosjekter gjerne ikke blir ferdig etter planen, er den generelle tendensen vi har til å overvurdere sannsynligheten for ”konjunktive” hendelser, altså sannsynligheten for at både hendelsene A og B og C vil inntreffe, når vi kjenner sannsynlighetene for at A, B og C hver for seg vil inntreffe.

Tilsvarende har vi en tendens til å undervurdere sannsynligheten for disjunktive hendelser, altså at A *eller* B *eller* C vil finne sted. Dette er ofte en medvirkende årsak til at vi undervurderer sannsynligheten for at komplekse systemer, som menneskekroppen eller kjernekraftanlegg, vil bryte sammen. Selv om sjansen for at hver enkelt essensiell komponent i systemet skal bryte sammen er ørliten, så kan sannsynligheten for at hele systemet bryter sammen likevel være stor dersom mange essensielle komponenter er involvert. Og denne sannsynligheten for at en eller annen komponent bryter sammen, har vi altså en tendens til å undervurdere.

### **Preferansereversering og skjevheter knytta til presentasjon av informasjon**

Vi har ovenfor ikke bare sett eksempler på skjevheter knyttet til hvordan informasjon prosesseres, men også til hvordan informasjon *presenteres*. Vi så i ”asiasyke”-studien ovenfor at folks beslutningstaking dramatisk ble påvirket av om valgalternativene ble beskrevet i taps- eller vinsttermer. Dette er ett eksempel på såkalt *preferansereversering*, nemlig at man foretrekker alternativ A framfor B når alternativer presenteres på én måte, men skifter til å fortrekke B framfor A når samme alternativer presenteres på en annen, men normativt sett ekvivalent, måte.

Tversky et al. (1988) påviste en annen type årsak til preferansereversering. De ga én gruppe av folk følgende spørsmål: ”Tenk deg at 600 omkommer årlig i trafikkulykker. To tiltak er aktuelle for å gjøre noe med problemet, hhv. å iverksette program X eller program Y. Anslagene over kostnad og effekt av de to programmene er:

Program	Forventa omkomne	Kostnader
X	500	\$55 M
Y	570	\$12 M

Hvilket program av X eller Y mener du bør velges?”

De fleste (67%) valgte program X, altså det programmet som redder flest liv men til en høyere pris pr. liv enn hva det alternative programmet Y forventes å gjøre.

En annen gruppe ble presentert for tabellen ovenfor bortsett fra at kostnaden av program X ikke var ført inn:

Program	Forventa omkomne	Kostnader
X	500	?
Y	570	\$12 M

De fikk så spørsmålet om hvor mye kostnaden av program X måtte være for at de ville vurdere de to programmene som likeverdige. Bare 4 % av de spurte svarte da \$55 millioner eller mer., Det vil altså si at langt de fleste (96%) her mente at program Y var et bedre alternativ enn X, dersom program X hadde kostet 55 millioner dollar.

Tversky et al. (1988) såkalte "prominens"-hypotese forklarer denne form for preferansereversering med at når vi skal *velge* mellom alternativer, så har vi en tendens til å legge relativt større vekt på det attributtet vi opplever som det viktigste, her altså antall liv som vil berges. Dette igjen forklares med at når vi velger så tar vi gjerne høyde for at vi må begrunne valget (for oss selv eller andre) i ettertid, og i begrunnelser av valg er det gjerne det viktigste attributtet som står i fokus. Det er for øvrig påvist en rekke andre typer årsaker til preferansereversering (se f.eks. Hsee et al., 2004, for en oversikt)

### Motivasjonelle årsaker til beslutningsskjevheter

Vi har sett at våre kognitive begrensninger tvinger oss til å bruke ikke-optimale *satisficing*-strategier. Det er én hovedgrunn til at vi avviker systematisk fra rasjonell tenkning. Det faktum at vi har en tendens til å attribuere mer av arbeidet utført i et prosjekt til oss selv enn det vi faktisk har bidratt med, forklarte vi ovenfor rent "kognitivt" med bruk av tilgjenglighetsheuristikk. Det er lettere for oss å legge merke til og huske hva vi selv gjør enn hva kollegaer gjør. Men nå er det ikke bare slik at vi har en tendens til å ta en uforholdsmessig stor del av æren for kollektiv suksess. Vi tar typisk også for lite ansvar for kollektive feil. Det er det verre å forklare rent kognitivt.

En helt annen type årsak til at vi avviker fra rasjonell tenkning er vår tilbøyelighet til hva som med en samlebetegnelse omtales som "motivert tenkning" (*motivated reasoning*), nemlig vår tendens til å trekke våre konklusjoner i en retning som følelsesmessig tiltaler oss. "Selv-tjenende"-resonnering (*self-serving reasoning*) er en form for motivert tenkning som forklarer vår tendens til både å ta en for stor del av æren for kollektiv suksess og for lite ansvar for kollektive feil. Slik resonnering gjør gjerne også at vi trekker våre vurderinger av hva som er riktig og galt i en retning som tjener oss selv. Messick og Sentis (1979) delte folk i to grupper. Folk i den ene gruppa ble bedt om å forestille seg at de selv hadde arbeidet sju timer i en jobb, mens en annen person hadde arbeidet ti timer i samme jobben. For folk i den andre gruppa, ble timetallene reversert. Man fikk så vite at personen som arbeidet sju timer ble betalt med 25\$. Deretter ble folk spurt hvor mye de som jobba i ti timer burde få. "Sju-timers gruppa" mente i snitt at de som jobba ti timer burde få 30.29\$, mens "ti timers gruppa" i snitt mente de fortjente 35.24\$. Differansen mellom gruppensnittene, altså 4.95\$, tallfester i dette tilfellet den "selv-tjenende" skjevheten i vurderingen av hva som er rettferdig lønn for ti timers arbeid.

Tendensen vi har til å trekke vurderinger i en retning som tjener oss selv, er eksempel på en form for såkalte "positive illusjoner" (*positive illusions*). En annen slik illusjon er tendensen



vi har til å ha et urealistisk positivt syn på oss selv. Dette gir seg blant utslag i den såkalte ”over-snittet”-effekten. Ærlighet, samarbeidsevne, rasjonalitet, kjøreferdighet, helse, intelligens, nesten uansett hvilke positive ferdigheter eller egenskaper det er snakk om, så mener langt over halvparten av oss at vi selv hører med til den bedre halvdel.

At vi har en tendens til å vektlegge og vurdere informasjon slik at det passer med våre egeninteresser, er neppe særlig overraskende. Mer overraskende og enda mer problematisk for vår beslutningstaking, er det at vi har en gjennomgående tendens til å ignorere en særdeles viktig type informasjon, også uten at det er motivasjonelle årsaker involvert.

### **Bekreftelsesfella (Confirmation trap)**

Følgende er et eksempel på en sekvens av tre tall som passer med en enkel regel som jeg tenker på: ”2-4-6.” Regelen sier noe om relasjonen mellom de tre tallene. Din oppgave er å oppdage denne regelen ved å skrive ned sekvenser med tre tall. (...) Etter at du har skrevet ned hver sekvens, vil du få vite om sekvensen passer med min regel eller ikke. (...) Du skal prøve å bestemme regelen på grunnlag av færrest mulig tilbakemeldinger (...) Når du føler deg sikker på at du har funnet fram til regelen, og ikke før, skal du skrive den ned.

Hvordan vil du gå fram? Hva tror du kan være regelen? Hvilke sekvenser av tre tall vil du teste ut? Tenk litt over dette før du leser videre.

Wason (1960) ga instruksjonen ovenfor til 29 studenter. Regelen han tenkte på var: ”Tre tall i stigende rekkefølge.” Bare 6 av studentene fant fram til denne regelen uten først å ha kommet med et galt forslag. De fleste foreslo regler som var langt mer komplekse enn den korrekte. Det viste seg at studentene hadde en sterk tendens til bare å teste ”positive” sekvenser/eksempler, altså sekvenser som kunne bekrefte deres egen antagelse om riktig regel.

For å finne fram til Wasons regel, er du nødt til å prøve å avkrefte den regelen som du selv i øyeblikket tror på. Wason (1960) fant at få hadde det som skulle til for å lykkes, nemlig ”a willingness to attempt to falsify hypotheses, and thus to test those intuitive ideas that so often carry the feeling of certitude” (p. 139). Forsøkte du selv ovenfor å falsifisere din egen hypotese om hva som kunne være regelen?

Wasons studie synliggjør en markant tendens i vår tenkning, nemlig at vi gjerne ser etter informasjon som støtter det vi til enhver tid tror mest på. Når du først har tenkt tanken at du kanskje skal kjøpe en bestemt bil, ansette en bestemt person osv., så har du en sterk tendens til å lete opp informasjon som styrker deg i troen. Men det å søke etter informasjon som eventuelt kan avkrefte det du tror på, er ofte den mest effektive måten å teste ut både svakhetene og styrkene i dine egne oppfatninger. Derfor vil du også normalt ha større utbytte av å bruke et konsulentfirma som gir deg de beste grunnene til å droppe prosjektideene dine, framfor et firma som forteller deg hvor gode dine ideer er.

### **Overkonfidens**

Operahuset i Sydney skulle etter planene i 1957 koste sju millioner dollar og stå ferdig i 1963. En moderert utgave av det planlagte operahuset ble ferdig i 1973 og kostet 102 millioner dollar. Dette trekkes gjerne fram som kroneksempelen på den såkalte ”planleggings”-feilen

(*planning fallacy*), som betegner tendensen både enkeltmennesker og organisasjoner har til å være for optimistiske med hensyn på når planlagte prosjekter vil bli ferdige (se f.eks. Buehler et al., 2002). Et mer allment eksempel, Buehler og kolleger spurte studenter i avslutningsfasen av psykologistudiet når de ”realistisk sett” regnet med å levere inn sine hovedoppgaver, og også når de regnet med å levere oppgavene dersom ”alt gikk så dårlig som tenkes kunne.” Studentenes ”realistiske” anslag var alt for optimistiske. Bare 30% av dem hadde fullført innen sine egne ”realistiske” anslag. I snitt brukte de 55 dager på å fullføre oppgaven, 22 dager lengre enn de hadde forutsett, og sju dager lengre enn hva de i ”verste fall” hadde forestilt seg.

Planleggingsfeilen er et eksempel på såkalt ”overkonfidens” (*overconfidence*), vår generelle tendens til å være sikrere på våre vurderinger og slutninger enn hva det faktisk er grunnlag for. Grad av overkonfidens varierer noe med oppgavens vanskelighetsgrad. Vi er gjerne mer overkonfidente på moderat til vanskelige spørsmål enn på lette. Vi har også en tendens til å være mindre overkonfidente på områder vi er fortrolige med. Positive illusjoner er en (åpenbar) hovedårsak til overkonfidens

Overkonfidens er en beslutningskjevhet som kan gi store negative konsekvenser. Historikere har for eksempel lenge framhevet overkonfidens som en vesentlig årsak til krig. I kriger siden år 1500 har den angripende part tapt krigen i 50 til 75% av tilfellene (litt avhengig av hvordan man regner.) Det er dessuten en markert tendens til at den parten som vant krigen, opplevde seieren som mer dyrkjøpt enn forventet. Og summerer man hvordan partene i krigen vurderte sjansen (i prosent) for seier før krigen brøt ut, blir tallet vanligvis langt over hundre prosent (Johnson, 2004).

## Etterpåklokskap

Berlinmurens fall i 1989. Flyene som styrtet inn i World Trade Center 11/9 2001. Dette er hendelser som du helt sikkert har fått med deg. Men tenk deg tilbake til tida før disse hendelsene fant sted. Hvor sannsynlig mente du da at det var at Muren skulle falle og terrorister fly inn i WTC? Kan du overse det du nå vet når du svarer?

Neppe. Etter at en hendelse har funnet sted, har vi en sterk tendens til å overvurdere i hvilken grad vi på forhånd kunne ha forutsett at den ville finne sted. Denne skjevheten omtales som ”etterpåklokskap” (*hindsight bias*) (Fischhoff, 1975) og forklares gjerne som en forankrings-effekt. Kunnskap om en hendelse blir lett et ”anker” for hvordan vi mener at vi vurderte sannsynligheten for hendelsen i forkant. Justeringer i forhold til slike ”ankere” har en tendens til ikke å være tilstrekkelige. Kunnskap i etterkant av en hendelse bidrar altså til hvor sterkt vi tror at vi forutså hendelsen i forkant.

En uheldig konsekvens av ”etterpåklokskap” er at denne beslutningsskjevheten reduserer muligheten til å lære av våre feilvurderinger. Skjevheten bidrar til å skjule feilene vi har gjort. Slik leder også ”etterpåklokskap” til overkonfidens i våre prediksjoner.

Nært relatert til etterpåklokskap er skjevheten som omtales som ”kunnskapens forbannelse” (*curse of knowledge*). Når vi selv besitter en bestemt kunnskap, så har vi vondt for å forestille oss at andre ikke har den samme kunnskapen. Denne ”forbannelsen” forklarer vanskeligheten fagfolk har med å tilpasse sine redegjørelser til lekfolk. Den forklarer også hvorfor markedseksperter generelt er dårligere til å predikere oppfatninger, verdier og smak hos vanlige forbrukere, enn hva forbrukere, som selv ikke er eksperter, er i stand til (Hoch, 1988).

## Hvorfor tenker vi på måter som gir systematiske skjevheter?

Så langt har fokus vært på skjevhetene som systematisk følger blant annet bruk av kognitive forenklingsstrategier. Men ut fra et ”begrenset rasjonalitets”-perspektiv så bidrar også slike strategier til beslutningseffektivitet, til at vi ofte i det hele tatt kan ta beslutninger i en kompleks virkelighet. De ulike ”satisficing”-strategiene vi typisk bruker bidrar antageligvis langt oftere til adekvate enn inadekvate beslutninger, og tap av beslutningskvalitet vil oftest oppveies av spart tid. For eksempel er bruk av representativitetsheuristikk en rask og god strategi når vi ikke har tilgang til baserate-informasjon, som vi jo ofte ikke har.

Tilsvarende fungerer tilgjengelighetsheuristikk bra så lenge tilgjengelighet og hyppighet samvarierer. Trolig fungerte denne heuristikken aldeles utmerket i en før-moderne verden. Da var stort sett det som var viktig for å holde seg i live, lett både å legge merke til og huske. I vår moderne verden derimot, er statistisk informasjon ofte langt mer pålitelig enn slående, lett tilgjengelige egne erfaringer. Tilgjengelighetsheuristikk gjør at vi - fagfolk inkludert – likevel har en tendens til å legge overdreven vekt på det siste. I stedet for å resonnerer statistisk er vi tilbøyelige til å tenke ”dramatisk”.

Ut fra et ”begrenset rasjonalitets”-perspektiv er altså tenkemåter som systematisk leder til beslutningsskjevheter samtidig tenkemåter som leder til beslutningseffektivitet. Hovedfokuset i beslutningspsykologien har imidlertid vært på skjevhetene og ikke på effektiviteten. Et unntak er Gigerenzer et al. (1999) som har vist gjennom simuleringer at en rekke svært enkle

heuristikker (*satisficing*-strategier) under bestemte betingelser (stor usikkerhet, lite kunnskap etc.) ofte gir tilnærmet like gode resultater som optimale metoder.

De øvrige tenkemåtene nevnt ovenfor som gir systematiske beslutningsskjevheter har også positive funksjoner. Vår markerte tendens til å søke bekreftelse, bidrar for eksempel til å styrke og stabilisere vår tro og våre meninger. Og tvil er jo slett ikke alltid av det gode. I mange sammenhenger vil det også være praktisk umulig, eller i alle fall lite funksjonelt, å gå etter informasjon som vi *ikke* forventer ut fra de hypotesene vi har. Dessuten er mye av det vi tror på av en slik karakter at det vanskelig lar seg avkrefte.

Positive illusjoner og urealistisk optimisme vet vi fremmer både helse, kreativitet og prestasjoner på både fysiske og mentale oppgaver (Taylor, 1989). Overkonfidens som positive illusjoner bidrar til, kan åpenbart også ha positive funksjoner. En lege som er overkonfident med hensyn på hvordan behandlingen hun gir vil virke på pasientene, får (på grunn av placebo-effekter etc.) gjerne bedre behandlingsresultater enn en mer realistisk lege. Og det er gjerne ”overkonfidente” redaktører og bidragsytere som får ut bøker som denne. Hadde vi hatt et realistisk bilde av våre tidsplaner når vi ble bedt om å bidra, så hadde antagelig flere av oss takket nei ...

## 9.5 Hvordan beslutningstaking kan forbedres – Noen effektive metoder

Det er nærliggende å tenke seg at erfaring automatisk leder til bedre skjønn og beslutningstaking. Men det er slett ikke tilfelle:

“In nearly every study of experts carried out within the judgment and decision-making approach, experience has been shown to be unrelated to the empirical accuracy of expert judgments” (Hammond, 1996, 278).

### **Hvorfor hjelper (ofte) ikke erfaring?**

De to generelle betingelsene for (effektiv) erfaringslæring er ofte ikke er til stede i praktisk profesjonsutøvelse. Den ene betingelsen er at vi får umiddelbar, utvetydig og konsistent tilbakemelding *når* vi tar/gjør feil. Den andre, at den tilbakemeldingen vi får gir oss en klar forståelse av *hva* som var feilen.

En hovedgrunn til at disse læringsbetingelsene oftest ikke er til stede, er at profesjonsutøvelse finner sted i hva vi kan kalle *probabilistiske* omgivelser. Samme type bedømming/valg kan i slike omgivelser gi forskjellig (type) tilbakemelding fra gang til gang. (I deterministiske omgivelser derimot, gir samme (type) valg/bedømming samme (type) tilbakemelding.) Dette igjen betyr at vi kan ta en riktig beslutning, uten at resultatet blir godt. Og motsatt, vi kan ta en dårlig beslutning med gunstig resultat. Vi kan altså få positiv tilbakemelding når beslutningene våre er dårlige og omvendt. Da er det åpenbart at betingelsene for erfaringslæring ovenfor ikke er til stede.

En annen grunn til at beslutninger ikke blir bedre med erfaring er at vi ofte ikke får noen tilbakemelding på beslutningene isolert sett. Vi får gjerne tilbakemeldingen på beslutningen pluss de handlingene vi gjør i forlengelsen av den. For eksempel, tenk deg en psykiater som beslutter å tvangsinnlegge en pasient. Dersom denne pasienten etter en uke har roet seg kraftig ned, hvordan vil da psykiateren evaluere sin beslutning uka før? Og hva om pasienten etter en uke er blitt hakkende gal? Antagelig vil psykiateren i begge tilfeller tolke tilbakemeldingen på sin beslutning om tvangsinnleggelse som positiv.

Så det er altså slett ikke slik at erfaring alene gjør vårt skjønn og beslutningstaking bedre. Hvordan skal vi da eliminere skjevheter i skjønn og beslutningstaking? Svaret avhenger av hvordan vi forstår disse skjevhetene.

### **Beslutningsskjevhetene er systematiske og robuste**

Enkelte filosofer og økonomer har avvist at feil og skjevheter i fagfolks beslutninger er systematiske. Filosofen Cohen (1981) hevder for eksempel at det er meningsløst å legge kriteriene for rasjonalitet utenfor mennesket selv. Han forfekter et slags *l'Homme Claire*-syn: Folk gjør feil, men feilene er ikke systematiske. Det er følgelig ikke noe "gap" å lukke mellom normativ og deskriptiv beslutningstaking. I dag har Cohens syn liten støtte i beslutningsforskningen. Evidensen for *systematiske* beslutningsskjevheter er nå overveldende og bredt akseptert.

Fischhoff (1982) la tidlig grunnlaget for denne aksepten. Han påviste at mange av de beslutningsskjevhetene som er beskrevet ovenfor er overraskende robuste mot forsøk på å korrigere dem. Det er for eksempel sjelden tilstrekkelig å eliminere beslutningsskjevheter utelukkende ved å gjøre beslutningstakere oppmerksomme på dem. Fischhoff viste at selv når for eksempel "etterpåklokskap" eksplisitt beskrives for folk og de instrueres om hvordan de kan unngå denne skjevheten, så opprettholdes den. Det samme var tilfellet med overkonfidens. Fischhoff fant at intensiv tilbakemelding på folks bedømminger ga en viss reduksjon av overkonfidens, men effekten var kortvarig.

Det er altså liten grunn til å tro at det å lese et kapittel som dette om beslutningsskjevheter alene bidrar til forbedret beslutningstaking. Det skal mer til.

## **Motvirking av beslutningsskjevheter – noen hovedstrategier**

Det er en rekke ulike metoder som er foreslått for å forbedre beslutningstaking. Larrick (2004) skiller mellom tre hovedstrategier: motivasjonelle, kognitive og teknologiske.

*Motivasjonelle strategier* er basert på en tro på at folk har potensial til å nærme seg normative idealer. De må bare motiveres. Det er to hovedmåter å motivere på: ved å gi *insentiver*, altså belønne beslutningskvalitet, og gjennom *ansvarliggjøring*, som innebærer å pålegge beslutningstakeren å begrunne sine beslutninger overfor andre.

Det har vist seg at insentiver får folk til å anstrenge seg mer og bruke mer tid på beslutningene. Det er nærliggende å tenke seg at dette også bedrer beslutningskvaliteten. Det er imidlertid knapt noen empirisk støtte for dette. En oversiktsstudie av effekten av insentiver konkluderte for eksempel slik: "there is no replicated study in which a theory of rational choice was rejected at low stakes in favor of a well-specified behavioral alternative, and accepted at high stakes" (Camerer & Hogarth, 1999, 33).

Grunnen til at insentiver ofte ikke har effekt er at folk rett og slett ikke har "den kognitive kapital" som skal til for å ta bedre beslutninger. Dersom gode beslutninger for eksempel krever en viss kjennskap til bayesisk resonnering, så hjelper det lite med insentiver dersom du ikke har hørt om Bayes' formel. Det har for eksempel vist seg at insentiver (på prediksjonsoppgaver) får folk til å legge mindre vekt på baserate-informasjon og mer på andre tilfeldige og lite relevante opplysninger som de gjerne også bruker inkonsistent. Insentiver bidrar da til å gjøre folks prediksjoner dårligere. De får hva Larrick (2004) omtaler som en "lost pilot"-effekt: "Jeg vet ikke hvor jeg skal, men det går i hvert fall radig unna."

Ansvarliggjøring bidrar også til at vi anstrenger oss mer, hvilket som nevnt slett ikke nødvendigvis leder til bedre beslutninger. Ansvarliggjøring bidrar dessuten vanligvis til et sterkt behov for å se konsistent ut i andres øyne. Dette kan slå både godt og dårlig ut. Vet man hva de man må begrunne beslutningene ovenfor mener, så kan ansvarliggjøring ofte lede til skjevheter ved at man tilpasser beslutningene til disse. Et annet problem er at man gjerne legger overdreven vekt på slående eller lett begrunnbare aspekter ved beslutningen når man vet at man i ettertid må begrunne den.

*Kognitive strategier* for å forbedre beslutningstaking omfatter å gjøre folk oppmerksomme på beslutningsskjevheter (som ved å lese kapitler som dette), opplæring/trening i (normative) beslutningsregler (som Bayes' formel) eller måter å omformulere beslutningsproblemer på, samt en rekke mer spesielle strategier. Nedenfor gir jeg eksempler på to spesielle kognitive strategier som har vist seg spesielt effektive.

Det å gjøres oppmerksom på typiske skjevheter har som nevnt vist seg å ha liten effekt. Opplæring i normative regler som logikk og sannsynlighetsregning har en viss effekt. Det å re-formulere beslutningsproblemer kan i enkelte tilfeller ha stor effekt. For eksempel har det vist seg at dersom leger får Eddys spørsmål ovenfor formulert i form av frekvenser, altså: 10 av 1000 kvinner på 40 år har brystkreft, 8 av 10 kvinner med kreft tester positivt, 1 av 10 kvinner uten kreft tester også positivt. Så svarer over halvparten av legene tilnærmet korrekt når de får spørsmål om hva som er sannsynligheten for at en som tester positivt har brystkreft (Gigerenzer, 1996). Men en slik reformulering av problemet (på egenhånd) krever en god del av beslutningstakeren.

I motsetning til motivasjonelle og kognitive strategier, så innebærer *teknologiske strategier* (slik Larrick, 2004, bruker begrepet) å forbedre beslutningstaking ved å gå ut over den enkelte beslutningstaker, enten ved å organisere beslutningstakingen som en gruppeprosess eller ved å bruke ulike former for teknologiske hjelpemidler. Det er utviklet og utprøvd en rekke metoder både for å støtte beslutningstakere i ulike aspekter av beslutningstakingen og for å erstatte beslutningstakere i deler av beslutningsprosessen. Nedenfor skal vi se nærmere på to ulike ”teknologiske” strategier. Men før vi gjør det, skal vi se på et par enkle men effektive kognitive strategier.

### **Vurder det motsatte!**

En enkel kognitiv strategi har vist seg å være spesielt effektiv, nemlig å tenke seg grunner til at du selv tar feil. Strategien består ganske enkelt i å stille seg spørsmålet: ”Hvilke grunner taler for at det motsatte av min vurdering/beslutning er riktig?” Denne strategien har vist seg å redusere både overkonfidens, etterpåklokskap og forankringseffekter (f.eks. Soll & Klayman, 2004).

Strategien er effektiv fordi den motvirker den sterke tendensen vi alle har til å søke informasjon som bekrefter egne oppfatninger (jf. ovenfor). Når vi tar andre alternativer i betraktning, så retter vi samtidig oppmerksomheten mot en annen type informasjon enn den som bekrefter våre oppfatninger. Vi utvider informasjonsgrunnlaget for vår beslutningstaking og gjør det mer representativt. Slik motvirker vi det grunnleggende problemet ved assosiativ ”system 1”-tenkning, nemlig at vi gjerne tar et alt for snevert utvalg av informasjon i betraktning.

### **Ta et ”outside”-perspektiv**

Når vi starter opp et nytt prosjekt så er vi opptatt av hvor lang tid det vil ta og hvor mye det vil koste. Vi så ovenfor at våre anslag gjerne er alt for optimistiske. Kahneman og Lovallo (1993) mener en grunn til dette er at vi typisk ser prosjektet i et ”innside”-perspektiv. Det vil si at vi gjør våre estimater om tid og kostnader ut fra dette nye prosjektet isolert betraktet. De gir følgende eksempel: Kahneman var en gang med i et større prosjekt der de skulle lage et pensumverk for et nytt fagområde. Gruppen kom til at det ville ta fra 18 til 30 måneder å fullføre prosjektet. Kahneman spurte da et annet av gruppemedlemmene: ”Vi er sikkert ikke det eneste teamet som har gått i gang med å lage en slik type nytt bokverk der man ikke hadde noe fra før å bygge på. Prøv å finn fram til så mange andre slike tilfeller som du kan. Tenk deg at disse tilfellene er på samme stadium som vi nå er i vårt prosjekt. Hvor lang tid tok det dem å bli ferdig?” Etter å ha undersøkt litt, kom dette gruppemedlemmet til at 40% av slike prosjekter aldri ble ferdig, og ikke noe tilsvarende prosjekt han hadde kommet over hadde blitt fullført på mindre enn syv år. Det viste seg at teamet endte opp med å bruke åtte år på bokprosjektet.

Et ”innside”-perspektiv er altså karakterisert ved at vi ser på tilfellet vi skal vurdere som unikt. Da avskjærer vi oss samtidig fra å trekke lærdom av lignende tilfeller. ”Utside”-perspektivet på den annen side, gjør det nærliggende å se likheter og trekke lærdommer av lignende prosjekter.

Vi veksler ofte mellom et ”outside”- og et ”innside”-perspektiv. Vi kan både tro at våre estimater av tid og pris på eget prosjekt er rimelig nøyaktige (”innside”-perspektiv), men

samtidig være oppmerksom på at de fleste lignende prosjekter tar lengre tid ("outside"-perspektiv): "Det blir alltid dyrere enn man tror." Det er for eksempel ikke uvanlig at folk som skal bygge eget hus, både har en bestemt oppfatning av hva prisen på det ferdig huset vil bli, men samtidig gir uttrykk for at anslaget antagelig er for lavt. De baserer seg da på kunnskap om andre husbyggere som har undervurdert kostnadene på sitt ferdige hus.

### Gruppebeslutningstaking

Multi-vitenskapsmannen Francis Galton, en (halv)fetter av Darwin, besøkte i 1906 en kvegutstilling. På utstillingen var det stilt ut en stor okse. Folk ble invitert til å anslå hvor mye oksene veide. Det var satt opp en pris til beste tips. Etter utstillingen samlet Galton inn alle avgitte gjetninger på oksens vekt. Han fant at snittippet var 1197 pund. Oksene veide 1198. Ingen av ekspertene på kveg som gjettet på oksens vekt var i nærheten av presisjonen i folks snittips. Dette er et tidlig eksempel på en nå anerkjent strategi på å forbedre bedømminger, nemlig å kombinere de estimatene som medlemmer av en gruppe har kommet fram til hver for seg.

Et annet eksempel på at dette kan være effektivt viste en analyse av kvaliteten på de rådene som deltagerne i TV-programmet "Vil du bli millionær", fikk. I dette programmet har deltagerne mulighet til å be om råd. De kan ringe en venn som de tror vet mye om emnet, eller de kan spørre hele auditoriet. Det viste seg at rådene deltagerne fikk fra sine "ekspertvenner" var riktige i 65% av tilfellene. Rådene de fikk fra auditoriet, altså det svaralternativet som flertallet av de som satt i auditoriet stemte på, viste seg å være riktig i hele 93% av tilfellene, selv om det ikke er noen grunn til å tro at de som satt i auditoriet hver for seg hadde spesiell greie på det de ble spurt om (Surowiecki, 2004)

I begge disse eksemplene er vurderingene et resultat av at man kombinerer vurderingen til hvert enkelt medlem av en gruppe. Når grupper diskuterer seg fram til et svar, er kvaliteten på vurderinger og beslutninger langt fra like overbevisende. Sosiale faktorer (andres nærvær, gruppepress etc.) virker ofte uheldig inn når en gruppe sammen diskuterer seg fram til en beslutning/estimat (f.eks. Janis, 1982).

Surowiecki (2004) sammenfatter de ideelle kravene som "kloke grupper" må tilfredsstillere. Det må være en spredning i oppfatninger blant gruppemedlemmene. Hvert medlem må få mulighet til å danne seg en oppfatning uten at den påvirkes av de øvrige medlemmene. Folk i gruppa må basere sine oppfatninger på ulike typer kunnskap. Man må ha en "mekanisme" som kan sammenfatte (aggregere) hvert enkelt gruppemedlems vurdering til en felles konklusjon.

En lenge brukt metode som tilfredsstillere disse kravene, er den såkalte *delphi-metoden*. Denne metoden innebærer at man innhenter estimer/vurderinger fra hver enkelt ekspert i en liten gruppe i minst to runder. Innspillene fra ekspertene er anonymisert for å unngå gruppepress. Etter hver runde blir ekspertenes estimer oppsummert og rapportert tilbake til ekspertene. Ekspertene blir dessuten informert om begrunnelsene for de ulike vurderingene. Man holder på inntil man er kommet fram til en vurdering som alle gruppemedlemmene kan akseptere.

Delphi-metoden ble utviklet på RAND på 1950-tallet og er blitt brukt mye i blant annet næringslivet. I en nylig evaluering av metoden, fant Rowe og Wright (2001) at delphi-metoden ga mer nøyaktige estimer enn tradisjonelle grupper i fem studier, dårligere i ett og den kom likt ut i to. Over alle sammenligningene de gjorde fant Rowe og Wright at delphi-metoden forbedret nøyaktigheten i 71% av tilfellene og svekket den i 12 %.



Ideen om at snittet av mange uavhengige estimater er bedre enn selv de fremste eksperters enkeltestimater, ligger også til grunn for såkalte prediksjonsmarkeder. Dette er ”meningsmarkeder” som fungerer som en børs. Man legger ut for salg påstander som at ”Bush vil vinne valget”, påstander som gir en pengesum i premie om den slår til. Folk kan så handle med slike kontrakter. Markedsprisen slike meninger oppnår har vist seg å være en svært god predikator på en kandidats sjanse til å vinne valget. Prediksjonsmarkeder predikerer for eksempel utfall av valg gjerne bedre enn hva (seriøst gjennomførte) valgdagsmålinger gjør (f.eks. Wolfers og Zitzewitz, 2004).

Store selskaper som Simens har brukt slike prediksjonsmarkeder til å forutsi hvordan et nytt vareprodukt vil klare seg i markedet. Man kan tenke seg at mange personer innen Simens hver for seg har noe informasjon om et slikt spørsmål. Ved å kombinere hver enkelt ansatts uavhengige estimater, så får man ”trukket ut” og nyttiggjort seg den ”sanne” komponenten i hvert enkelt estimat.

### **Erstatt/automatiser skjønn: lineære modeller**

Automatisering av fagfolks skjønn er et alternativ til metoder for å forbedre skjønnets. Jeg innledet dette kapitlet med følgende påstand:

”Expert judgments have been worse than those of the simplest statistical models in virtually all domains that have been studied” (Camerer & Johnson, 1991, 203)

Dette sitatet sammenfatter det mest spektakulære funnet i forskningen på ekspertskjønn, nemlig at helt enkle formler integrerer informasjon systematisk bedre enn hva intelligente og erfarne eksperter er i stand til. Paul Meehl (1954) var pioneren bak denne oppdagelsen. Han skilte mellom innhenting av informasjon og vurdering av hva den innhentede informasjonen samlet indikerer. I 1954 gikk Meehl gjennom de studier som til da var utført som sammenlignet klinikerens skjønnsmessige vurdering av informasjon med enkle, lineære regresjonsmodeller. I de 20 studiene Meehl så på kom modellene stort sett alltid bedre ut enn fagfolkene, og de kom aldri dårligere ut.

Det overraskende med dette funnet var at lineære modeller ikke ”fanger opp” det klinikerne hevdet var essensielt i deres eget faglige skjønn, nemlig å vurdere informasjon helhetlig. Klinikerne hevdet nemlig at vekten de la på ulike data, avhang av hvilken verdi andre data hadde. Slike ”konfigurale” sammenhenger var ikke representert i Meehls lineære modeller. Likevel kom altså modellene systematisk bedre ut enn klinikerne.

Modellene Meehl sammenlignet med fagfolks skjønn var ”ekte” eller aktuariske. Det vil si at vektene var ”trukket ut” fra empirisk registrerte sammenhenger ved hjelp av statistiske metoder. I Meehls sammenligninger hadde dessuten modell og ekspert tilgang til samme informasjon. Det er siden 1954 gjennomført hundrevis av studier der man under disse betingelsene har sammenlignet lineære modeller og ulike typer eksperters skjønn. Det er knapt et eksempel på at ekspertene kommer bedre ut.

Man har dessuten sammenlignet ekspertskjønn mot ”uekte” lineære modeller.

Dette er lineære modeller der vektene ikke er bestemt aktuarisk, men ved hjelp av ulike ikke-optimale metoder. Såkalte ”bootstrapping”-modeller er ett eksempel på en type ”uekte” modell. Dette er modeller av eksperters egne skjønnsmessige vurderinger. Man lar eksperten

på grunnlag av en fast mengde av variabelverdier vurdere N antall tilfeller. Deretter foretar man en regresjonsanalyse av ekspertens vurderinger. Det gir en lineær modell av ekspertens skjønn, dvs. en modell der vektene i modellen svarer til hvordan eksperten typisk vektlegger variablene i sine vurderinger. Det har vist seg at også slike modeller av eksperten "vurderer" nye tilfeller systematisk bedre enn hva eksperten selv gjør. Grunnen er, kort sagt, at modellen i motsetning til eksperten integrerer data konsistent.

Det har dessuten ofte vist seg at helt enkle modeller ofte utkonkurrerer fagfolks skjønn selv om fagfolkene har tilgang til langt mer informasjon enn modellen. Dawes (1971) gir et mye referert eksempel på dette. I USA er det vanlig at en komité av erfarne skolefolk er involvert i opptaket av "high school"-studenter til college. Disse komiteene hadde i Dawes' studie tilgang til elevenes karakterer, et anbefalingsbrev fra skolen de kom fra og informasjon om kvaliteten på denne skolen. Komiteen gjennomførte dessuten et lengre intervju med hver elev.

Basert på tidligere registrerte sammenhenger, utarbeidet Dawes en modell som predikerte hvordan studentene ville gjøre det på college utelukkende på grunnlag av standpunktkarakterer, eksamenskarakterer og kvaliteten på high school. Dawes brukte modellen på 384 søkere. Han fant at modellen kunne eliminere de 55% svakeste elevene uten å eliminere noen av dem som komiteene aksepterte. I tillegg viste det seg at modellen predikerte bedre enn komiteen hvordan de elevene som ble tatt opp, kom til å gjøre det på college. (Ja, Dawes fant at bare elevenes standpunktkarakterer predikerte dette bedre enn komiteene.) Dawes anslo i 1971 at dersom man hadde basert opptaket til college på en slik enkel modell i stedet for å bruke opptakskomiteer, så ville man både få et bedre opptak og dessuten spare 18 millioner dollar i USA årlig. Omgjort til dagens pengeverdi og studenttall, så er det innsparte beløpet anslått til 500 millioner dollar årlig (Bazerman, 2006).

Det er siden 1954 gjennomført hundrevis av studier på en rekke fagområder der fagfolks skjønnsmessige vurderinger av gitte data er sammenlignet med enkle lineære modeller. Konklusjonen er entydig. Vet du hva slags opplysninger som er særlig relevante for en type vurdering som du stadig gjør, så utarbeid en lineær modell og gjør vurderingene ved hjelp av denne. Intelligente og erfarne fagfolks skjønn utkonkurreres nemlig så godt som alltid av slike modeller. Det gjelder altså ofte selv om fagfolkene har tilgang til flere opplysninger enn dem som inngår i modellene (f.eks. Dawes, Faust og Meehl, 1989).

## 9.6 Konklusjon

Fagfolk anvender tenkemåter som gir systematiske beslutningsskjevheter, dvs. avvik fra rasjonell beslutningstaking. Våre kognitive begrensninger er en grunn til at vi anvender slike tenkemåter. En annen grunn er at mange av disse tenkemåtene også har positive funksjoner. Et hovedproblem er at vi ikke selv er oppmerksomme på hvilke tenkemåter vi anvender og hvordan disse negativt kan påvirke vår beslutningstaking. Derfor skiller vi heller ikke mellom situasjoner der tenkemåtene er til fordel for oss og situasjoner der de er potensielt skadelige. En bedre forståelse av dette skillet er en nøkkel til forbedret beslutningstaking. Spesifikke beslutningsskjevheter kan ofte effektivt motvirkes av enkle kognitive strategier. Funn i beslutningspsykologi skulle for øvrig tilsi en langt større bruk av (teknologisk) beslutningsstøtte på mange profesjonsområder enn hva som faktisk er tilfelle.

## 9.7 Referanser

- Arkes, H. R. (1991) Costs and benefits of judgment errors - implications for debiasing. *Psychological bulletin* 110, 486-498
- Arkes, H. R. (2003) The nonuse of psychological research at two federal agencies. *Psychological science* 14 (1): 1-6
- Bazerman, M. H. (2006) *Judgment in managerial decision making*. Wiley
- Bernoulli, D. (1738/1954) Exposition of a new theory on the measurement of risk. *Econometrica*, 22, 23-36
- Buehler, R, Griffin, D., og Ross, M. (2002) Inside the planning fallacy: The causes and consequences of optimistic time predictions. I Gilovich, T., Griffin, D. og Kahneman, D. (eds.) (2002) *Heuristics and biases. The psychology of intuitive judgment*. Cambridge: Cambridge University Press, ss. 250-70.
- Camerer, C. F. (2001) Prospect theory in the wild: Evidence from the field," (pp 288-300) in D. Kahneman and A. Tversky (Eds.), *Choices, Values, and Frames*, 2001. Cambridge: Cambridge University Press
- Camerer, C. F. og Hogarth, R.M. (1999) The effects of financial incentives in experiments: A review and capital-labor-production framework. *Journal of risk and uncertainty* 19 (1-3): 7-42
- Camerer, C. F. og Johnson, E. J. (1991) The process-performance paradox in expert judgment: How can experts know so much and predict so badly? I: K. Anders Ericsson og Jacqui Smith, red, (1991) *Toward a general theory of expertise*. Cambridge: Cambridge University Press, ss. 195-217
- Chapman, G. B og Johnson, E. J. (2002) Incorporating the irrelevant: Anchors in judgments of belief and value. I Gilovich, T., Griffin, D. og Kahneman, D. (eds.) (2002) *Heuristics and biases. The psychology of intuitive judgment*. Cambridge: Cambridge University Press, ss. 120-38
- Cohen, L. J. (1981) Can human irrationality be experimentally demonstrated? *Behavioral and Brain Sciences*, 4, 317-31
- Dawes R.M., Faust D., og Meehl P.E. (1989) Clinical versus Actuarial judgment. *Science*, 243, ss. 1668-1673
- Eddy, D. (1982) Probabilistic reasoning in clinical medicine: Problems and opportunities. I D. Kahneman, P. Slovic, og A. Tversky, eds., *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press
- Fischhoff, B. (1975) Hindsight foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception og Performance* 1, 288-99
- Fischhoff, B. (1982) Debiasing. I D.Kahneman, P. Slovic, og A. Tversky, Eds., *Judgment under uncertainty: Heuristics and biases*. Cambridge, England: Cambridge University Press., 422-444
- Gigerenzer, G. (1996) The psychology of good judgment: Frequency formats and simple algorithms. *Medical decision making* 16, 273-280
- Gigerenzer, G, Todd, P. M. and the ABC Research Group (1999) *Simple heuristics that make us smart*. New York: Oxford university Press

- Gilovich, T. (1991) *How we know what isn't so. The fallibility of human reason in everyday life*. New York: Free press
- Hammond, K. R. (1996) *Human judgment and social policy. Irreducible uncertainty, inevitable error, unavoidable injustice*. New York: Oxford University Press
- Hoch, S. J. (1988) Who do we know: Predicting the interests and opinions of the American consumer. *Journal of Consumer Research*, 15 (3), 315-24
- Hsee, C.K., Zhang, J., og Chen, J. (2004) Internal and substantive inconsistencies in decision making. I D. J. Koehler og N. Harvey (eds.) *Blackwell handbook of judgment and decision making*. Malden, MA: Blackwell, ss. 360-78
- Janis, I. L. (1982) *Groupthink: Psychological studies of policy decisions and fiascos*. Boston: Houghton Mifflin
- Johnson, D.D.P. (2004) *Overconfidence and War: the Havoc and Glory of Positive Illusions*. Cambridge, MA: Harvard University Press
- Kahneman, D. (2003) A perspective on judgment and choice. Mapping bounded rationality. *American psychologist*, 58, 697-720
- Kahneman, D. og Tversky, A. (1979) Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263-91
- Kahneman, D. og Tversky, A. (2000) *Choices, Values, and Frames*. New York: Cambridge University Press
- Kahneman, D. og Lovallo, D. (1993) Timid choices and bold forecasts - a cognitive perspective on risk-taking. *Management science* 39 (1): 17-31
- Klein, G. (1998) *Sources of power. How people make decisions*. Cambridge, MASS.: MIT-press
- Larrick, R. P. (2004) Debiasing. I D. J. Koehler og N. Harvey (eds.) *Blackwell handbook of judgment and decision making*. Malden, MA: Blackwell, Ch 16, pp. 316-337
- Lerner, J. S. og Tetlock PE (1999) Accounting for the effects of accountability. *Psychological bulletin* 125 (2): 255-275
- McNeil, B.J., Pauker, S.G., Sox, H.C., jr. og Tversky, A. (1982) On the elicitation of preferences for alternative. *New England Journal of Medicine*, 306, 1259-1562
- Rottenstreich, Y. og Shu S. (2004) The connections between affect and decision making: Nine resulting phenomena. I D. J. Koehler og N. Harvey (eds.) *Blackwell handbook of judgment and decision making*. Malden, MA: Blackwell, Ch 22, pp. 444 - 463
- Rowe, G. og Wright, G. (2001) Expert opinions in forecasting: the role of the Delphi technique. I J.S. Armstrong, ed., *Principles of forecasting*. Boston: Kluwer, ss. 125-144
- Savage, L. J. (1954) *The foundations of statistics*. New York: Wiley
- Simon, H. A. (1955) A behavioral model of rational choice. *Quarterly Journal of Economics*, 69, 99-118
- Simon H.A. (1956) Rational choice and the structure of the environment. *Psychological Review*, 63, 129-138
- Simon, H. A. (1990) Invariants of human behaviour. *Annual review of psychology*, 41, 1 – 19

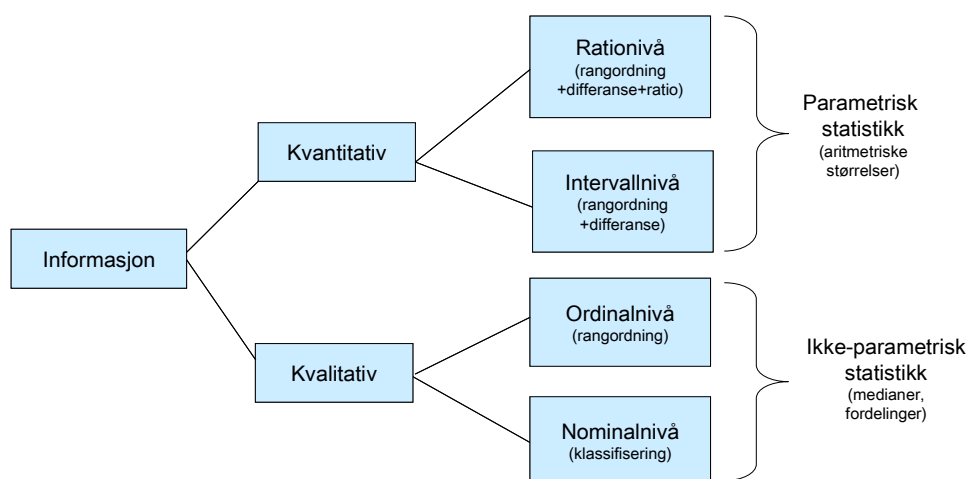
- Slovic, P., Finucane, M., Peters, E. og MacGregor, D.G. (2002) The affect heuristic. I Gilovich, T., Griffin, D. og Kahneman, D. (eds.) (2002) *Heuristics and biases. The psychology of intuitive judgment*. Cambridge: Cambridge University Press, ss. 397-420
- Soll, J.B. og Klayman, J. (2004) Overconfidence in interval estimates. *Journal of experimental psychology-learning memory and cognition* 30 (2): 299-314
- Surowiecki, J. (2004) *The wisdom of crowds*. New York: Doubleday
- Thaler, R. H. (1980) Toward a positive theory of consumer choice. *Journal of Economic Behavior and organization*, 1, 39-60.
- Tversky, A. og Kahneman, D. (1974) Judgment under Uncertainty: Heuristics and Biases. *Science*, 185, 1124-1131
- Tversky, A. og Kahneman, D. (1981) The framing of decisions and the psychology of choice. *Science*, 211, 453-63
- Tversky A., Sattath S., og Slovic P. (1988) Contingent weighting in judgment and choice. *Psychological review* 95 (3): 371-384
- von Neumann, J., og Morgenstern, O. (1947) *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press
- Wolfers, J. og Zitzewitz, E. (2004) Prediction markets. *Journal of economic perspectives* 18 (2): 107-126
- Wu G., Zhang, J. og Gonzalez, R. (2004) Decision under risk. I D. J. Koehler og N. Harvey (eds.) *Blackwell handbook of judgment and decision making*. Malden, MA: Blackwell, Ch 20, pp. 400-23
- Zajonc, R.B (1980) Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151-175

## 10 TILTAK FOR Å BEDRE KVALITETEN PÅ KVALITATIV INFORMASJON – SEMANTIKK OG PRESISJON

KNUT SAMSET\*

### 10.1 Kvalitativ versus kvantitativ informasjon

IT-revolusjonen og en vedvarende og voldsom utvikling av tilgjengelig datakraft gjennom flere tiår har gjort verden stadig mer finmasket. Det har muliggjort en grad av detaljering og presisjon som var utenkelig tidligere. Min mobiltelefon, for eksempel, er koplet mot det globale posisjoningsnettet og kan bestemme hvor jeg befinner meg til enhver tid med 2,3 meters nøyaktighet. Denne utviklingen gjør at det generelt stilles store krav til presisjon og etterprøvbarehet. Mye av utdanningen som former folk intellektuelt legger derfor i økende grad vekt på at informasjon skal være tallbasert. Ikke minst gjelder dette innen tekniske og økonomiske fag. Selvsagt er tallbasert informasjon i større grad enn tekstlig informasjon egnet til systematisk analyse som kan gi grunnlag for generalisering eller påvisning av samsvar med stor grad av sannsynlighet. For at slik statistisk bearbeidning skal være mulig er det en forutsetning at informasjonen kvantitativ, og på intervall- eller helst ratio-nivå (intervalltall med nullpunkt), se figur 10-1.



Figur 10-1 Kvalitativ og kvantitativ informasjon karakterisert etter målenivå (Olsson og Sörensen, 2003)

Økende krav til dokumentasjon som grunnlag for beslutninger gjør at vi ikke alltid ser potensialet i systematisk bruk av tekstlig informasjon. Hovedproblemene er at presisjonen blir lavere og muligheten for etterprøvbarehet ofte er begrenset. I tillegg kommer problemet med at

\* Knut Samset er professor ved NTNU, Institutt for bygg, anlegg og transport.

kvalitativ informasjon ofte preges av og/eller avhenger av individers fortolkning og at det derfor kan oppstå troverdighetsproblemer og misforståelser. Statistisk bearbeiding og fremstilling av slik informasjon er begrenset bl.a. til medianer, kvartiler og fordelinger (ikke-parametrisk statistikk). Imidlertid kan det i et gitt tilfelle være mulig å klassifisere informasjon med relativt høy oppløsning, som illustrert i figur 10-2, selv om dette ofte vil gå på troverdigheten løs ettersom vi har å gjøre med subjektive vurderinger der det lett kan reises spørsmål om fortolkning og kategorisering.



Figur 10-2 Ulike kvalitative uttrykk for subjektiv sannsynlighet. Utgangspunkt for klassifisering av informasjon på nominalnivå. (Teigen, 2006)

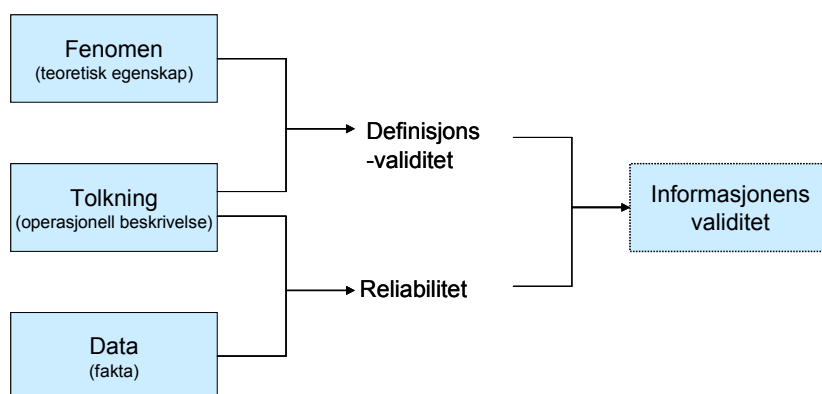
Vår ofte overdrevne trang til å vektlegge verdien av tallbasert informasjon gjør at det er lett å overse de viktigste fordelene ved å bruke tekstlig informasjon, nemlig at slik informasjon kan genereres raskt, at den er vårt viktigste grunnlag for kommunikasjon, men også at den er grunnleggende nødvendig for å gi en helhetsbeskrivelse av komplekse forhold der det inngår mange påvirkningsfaktorer.

Det er også slik at mye informasjon ganske enkelt ikke kan kvantifiseres. Videre at dersom den tallbaserte informasjonen ikke er pålitelig kan det være mer hensiktsmessig å benytte kvalitative uttrykk inntil pålitelig informasjon kan skaffes. Endelig, er det et kjent fenomen at et stort tallbasert datamateriale eller data på høyt aggregeringsnivå ofte kan bidra til å tilslore heller enn å klargjøre et forhold.

Selvsagt er det i praksis ikke et spørsmål om enten eller, men både og. Kvalitative vurderinger hjelper oss til å beskrive helheten mens kvantitativ informasjon gir beskrivelsen presisjon. Det er grunnen til at vi fortrinnsvis bruker kvalitativ informasjon når vi kommuniserer. Meningsinnholdet i det vi kommuniserer kan ses som deler av informasjonshierarkier der den underliggende, og uttalte informasjonen ofte er kvantitativ. Det er vissheten om at det finnes en underliggende informasjon og at denne kan hentes frem om nødvendig som først og fremst gir troverdighet til det som kommuniseres.

## 10.2 Informasjon og validitet

I forskning bruker en begrepet validitet for å karakterisere informasjonens godhet. Validitet betyr gyldighet og innebærer definisjonsmessig at det foreligger samsvar mellom virkelighet og tolkning. En generell modell som beskriver dette er vist i figur 10-3. Informasjon er valid eller gyldig dersom to kriterier er oppfylt: For det første må en sikre definisjonsvaliditet, det vil si at det er samsvar mellom tolkning og det fenomenet en ønsker beskrive, (for eksempel kroppstemperatur som mål på sykdom). I tillegg må en sikre reliabilitet, det vil si at uttrykket er pålitelig, (i dette eksempelet ved at måletallet stemmer med kroppstemperaturen).



Figur 10-3 Validitet som uttrykk for informasjonens godhet (Hellevik, 1991)

Eksemplet ovenfor gjelder kvantitativ informasjon. Når det gjelder kvalitative vurderinger vil en ha mer begrensede muligheter for å sikre presisjonen, og i noen tilfeller, for eksempel i analyse som i stor grad er basert på antakelser, vil det være den definisjonsmessige validiteten som vil være avgjørende for godheten av vurderingene. Utfordringen blir da å sikre at det begrepet eller informasjonselementet som brukes gir et godt uttrykk for hva vi ønsker å beskrive. Om det er tilfellet kan vi i alle fall være forholdsvis sikre på at informasjonen er relevant i forhold til de vi skal analysere, selv om det kanskje skorter på pålitelighet og presisjon.

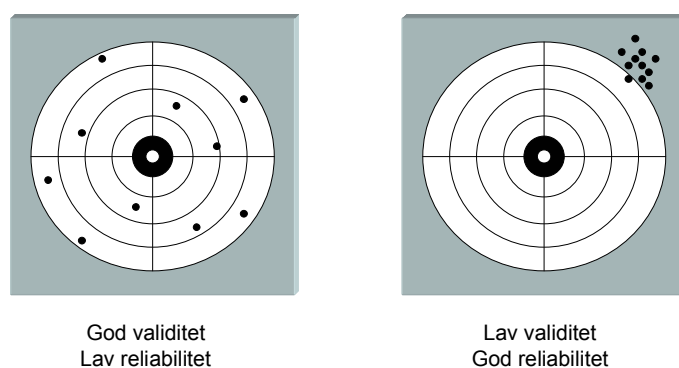
Validitetsproblemet oppstår når det er dårlig samsvar mellom det fenomenet en ønske å beskrive og det uttrykket eller utsagnet en velger å benytte. Når et samferdselsprosjekt har som mål å ”bedre sikkerheten i trafikken” oppstår det et validitetsproblem om en velger å benytte ”lokal sysselsetting” som en indikasjon på at en når målet. Intuitivt kan en anta at det ikke er noe klart samsvar mellom trafikksikkerhet og sysselsetting. Mer direkte indikatorer med god validitet vil for eksempel være *hyppighet av ulike typer ulykker, antall skader eller antall dødsfall* i trafikken.

Reliabilitetsproblemet oppstår når presisjonsnivået i utsagnet eller uttrykket en bruker er dårlig. I målformuleringen ovenfor har en et reliabilitetsproblem først og fremst ved at en bruker adjektivet *bedre*, som gir stort rom for tolkning - og substantivet *sikkerheten*, uten at det fremgår hva en mener med det. For å øke presisjonsnivået ville en måtte være eksplisitt på



*hvem* som omfattes av målgruppen, *graden* av forventet endring, *tidsrommet* før forventet endring er oppnådd, etc.

I den tidligste idefasen av et prosjekt vil en som nevnt tidligere i stor grad være henvist til å basere sine vurderinger og konklusjoner på kvalitativ informasjon. Dette fordi kvantitativ informasjon enten ikke finnes, at en ikke har ressurser til å skaffe slik informasjon, eller at dette ikke er hensiktsmessig, for eksempel i det tilfellet at realisering ligger mange år frem i tid og det er åpenbart at gjeldende verdier vil endres vesentlig og uforutsigbart over tid.



Figur 10-4 Validitet og reliabilitet – skivebom versus presisjon

I tillegg til dette kommer at det ofte er *hensiktsmessig* i en tidlig fase å gi avkall på presisjonen som tallbasert informasjon kan gi og heller nøye seg med adjektiver som ”god”, ”i høy grad”, ”vesentlig”, ”betydelig” for å beskrive omfang, mengde eller nivå. Det kan gi større frihet i arbeidet med å vurdere ulike modeller eller konsepter og plassere disse inn i en større kontekst, samt vurdere de prinsipielle konsekvensene ved endringer i disse. En kan derfor saktens redusere kravet til presisjon midlertidig – men det vil ikke under noen omstendighet være akseptabelt å lempe på kravet til definisjonsmessig validitet. Sagt med andre ord: Vi kan akseptere en viss usikkerhet og spredning med hensyn til hvor godt vi treffer, men et bomskudd er fullstendig uakseptabelt selv om presisjonen er god, som illustrert i figur 10.4.

### 10.3 Skjønnsmessig, sannsynlighetsbasert vurdering av informasjonens godhet

*Reliabilitet*, eller hvorvidt informasjonen er pålitelig, kan i prinsippet testes eller etterprøves. Påliteligheten sikres ved at indikatorene er entydige eller at måleinstrumentet ikke gir systematiske feil. Testen på pålitelighet vil være at dersom flere personer bruker samme indikator uavhengig av hverandre på samme problem så skal de få samme resultat. Kilden og måten informasjonen fremskaffes på er avgjørende for påliteligheten. Når det gjelder generering av kvalitativ informasjon, for eksempel i en spørreundersøkelse, kan dette illustreres ved at informanten, som er kilden, har dårlig informasjon eller mer eller mindre bevisst gir feilaktig informasjon. I tillegg er det selvsagt viktig for svaret hvordan spørsmålet er formulert, i hvilken sammenheng utspørringen skjer, og hvilke tolkningsfeil og misforståelser hos begge parter som måtte oppstå i kommunikasjonen.

Når det gjelder informasjonens *validitet*, er ikke dette noe som kan etterprøves men som i prinsippet må bases på skjønn. Her er det valget av indikator som er det avgjørende. For å sikre at informasjonen er gyldig i forhold til det fenomenet en ønsker å beskrive er det to hensyn en må ta: (1) å velge indikatorer som gir et mest mulig *direkte* mål, og i tillegg (2) å benytte *flere indikatorer* som til sammen gir god indikasjon på det fenomenet som skal beskrives.

Dersom fenomenet som skal beskrives for eksempel er ”*utdanningen ved et universitet*”, så vil en direkte indikator være ”*antall uteksaminerte*”. Men dette gir ikke et tilstrekkelig bilde av det som er oppnådd. I tillegg vil det være naturlig å supplere med informasjon om karakterer eller ”*prestasjonsnivå*”, samt hvilken ”*type utdanning*” det gjelder. Det vil også være naturlig å karakterisere ”*utdanningens kvalitet*”, for eksempel gradert i forhold til enkelte andre universiteter. I tillegg kan det være aktuelt å se på ”*gjennomføringstid*” i forhold til normert tid, og ”*frafallet*”, det vil si andel studenter som ikke gjennomførte utdanningen. Til sammen får en da et mer dekkende bilde av utdanningen ved universitetet. Eksemplet illustrerer informasjonens hierarkiske karakter. Det overordnede fenomenet som studeres er utdanning. Dette beskrives ved hjelp av underordnede indikatorer eller informasjonselementer som til sammen skal gi en akseptabel, valid beskrivelse av fenomenet. Noen elementer er viktigere for å gi en gyldig beskrivelse og har derfor høy validitet, mens andre kanskje er unødvendige eller direkte feil og har lav validitet.

Tabell 10-1 Et utvalg indikatorer for å vurdere brukertilfredshet i et kontorbyggprosjekt

Validitet	Indikator
Høy	Brukernes oppfatning av byggets funksjonalitet
Høy	Gjennomtrekk av leietakere
Høy	Etterspørsel etter kontorene i markedet
Middels	Brukernes oppfatning av trivsel på jobben
Middels	Vedlikehold av bygningen
Lav	Hvor lenge folk sitter på jobben utover kvelden
Lav	Firmaenes lønnsomhet
Lav	Prisnivået på kontorene

Når en opererer hovedsakelig med kvalitativ informasjon og så å si gir avkall på presisjonen er det ikke mindre viktig å sikre at informasjonen er valid. Dette må som nevnt gjøres skjønnsmessig, for eksempel ved å benytte sannsynlighetsbaserte vurderinger. Ta evalueringen av et kontorbyggprosjekt som eksempel: En ønsker å få et godt bilde etter en viss tid av ”*i hvilken grad brukerne er tilfreds med bygget*”. I tabell 10.1 er det listet åtte mulige indikatorer som kan brukes for å samle informasjon om dette. Informasjonens validitet kan testes indirekte ved å anslå i hvilken grad hver av disse indikatorene identifiserer typer informasjon

vil gi et vesentlig bidrag til å beskrive fenomenet. Oppfatning av ”byggets funksjonalitet”, for eksempel, er en direkte indikator, og derfor et godt uttrykk for hva en ønsker å måle. Hvorvidt folk ”trives på jobben”, derimot, påvirkes av flere andre forhold enn opplevelsen av bygget og er derfor mindre valid. Hvorvidt folk ”sitter på jobben utover kvelden” har antakelig lite med byggets utforming å gjøre og har derfor lav validitet. I tabellen er indikatorene rangert, og det blir relativt enkelt å se hvilken informasjon som er nyttig og hva som bør kuttes ut. Det er åpenbart slik av valid informasjon er nyttig, mens informasjon med lav validitet ikke bare er unødvendig, den bidrar til å tilsløre og kanskje forvrengte konklusjonen. Det er derfor gode grunner til å forholde seg stringent til kravet om validitet, for å fokusere, øke verdien av vurderingene og økonomisere med ressurser.

## 10.4 Presisjon - klarhet og entydighet

Som nevnt ovenfor er entydighet en forutsetning for pålitelig informasjon. Dette gjelder så vel indikatoren som angir hvilken informasjon en ønsker, som den informasjon som fremskaffes og som er grunnlag for fortolkning. I kvalitative utsagn kan språkbruken i seg selv skape problemer. Her som på mange andre områder står en ovenfor de utfordringer som er beskrevet i det fagfeltet som kalles semantikk. Ordene en bruker bør være veldefinerte både i leksikal betydning og stipulativ betydning.

Leksikale definisjoner av ord er gjerne slike som en finner i ordbøker og leksika. De beskriver ordet slik det brukes i dagligtale, vanligvis ved å presentere en beskrivelse med tilnærmet lik betydning, der ordet ikke inngår. Slike definisjoner er ofte for generelle til å anvendes i en gitt faglig sammenheng.

For å bøte på dette må en bruke en stipulativ definisjon, eller velge spesielle begreper som mer presist betegner det en vil uttrykke. En leksikalsk definisjon av ordet ”motstand”, for eksempel, er ”å forsvare” eller ”å ta til motverge”. Imidlertid har begrepet forskjellig betydning innen fysikk, medisin, finansanalyse og politikk. En stipulativ definisjon som anvendes innenfor fysikken er at motstand = forholdet mellom spenningen som legges over en strømkrets og strømmen som går gjennom denne. For å unngå uklarheter ved å bruke ordet kan man knytte en presisering til dette, det vil si ”elektrisk motstand”, som avklarer at det i denne sammenhengen refererer til en stipulativ definisjon. Enda klarere blir det om en unngår problemet helt, ved å innføre et helt nytt ord, i dette tilfellet ”resistans”. Det er dette forskerne gjør for å lette kommunikasjonen innenfor sine fagområder. Samtidig oppstår det da ofte problemer med kommunikasjonen på tvers av fagene, og i forhold til allmennheten som forholder seg mest til den leksikalske forståelsen av begrepene.

Ikke alle problemer knyttet til presisjon kan løses med mer presise definisjoner. Særlig aktuelt er det problemet at uklarheter kan oppstå når en bruker adjektiver som uttrykk for eksempel for mengde eller omfang. Uttrykk som ”godt”, ”høyt” og ”vesentlig”, er kategorier innenfor klassifiseringer på nominalnivå, som illustrert i figur 10.2, og dermed åpne for feiltolkning. I tillegg kommer at mange ord kan tolkes med ulikt meningsinnhold avhengig av ståsted. Ordet ”normalt” anvendt i omtale av et samferdselsprosjekt kan for eksempel for en part oppfattes som (1) innenfor grensene av det som er vanlig, men for en annen part (2) innenfor grensene av det som er akseptabelt, som kanskje ligger vesentlig lavere. Mange ord kan ha flere meninger. Hva betyr for eksempel at innholdet av kvikksølv i et drikkevann er normalt? Er det i forhold til innholdet i andre sjøer, i forhold til sjøer i uberørt natur eller i forhold til grenseverdier for skadelig påvirkning for mennesker? (Hansson, 2003).

Valg av ord og begreper og spesielt ord som angir verdi har betydning både ved innhenting og formidling av informasjon. Det er ikke uvanlig at forskjellige interessenter bruker begreper med forskjellig tolkning om samme fenomen, eller legger ulik betydning i samme begrep, for eksempel i forhold til planleggingen av et kontroversielt prosjekt. Ikke minst gjelder dette begreper som ”behov” og ”nytte” der det ofte er stor uenighet hos forskjellige parter både om fortolkning av begrepene og av informasjoninnholdet som formidles.

## 10.5 Komplekst meningsinnhold – koplinger

Et av de forholdene som skaper størst problemer i kvalitativ analyse er at en ofte benytter sammensatte utsagn. Det vil si at en setning inneholder en sammenkopling av flere meningsbærende elementer. Det gjør det vanskelig å forholde seg til utsagnet, ettersom en må vurdere hvert element for seg før en kan vurdere helheten. Dette problemet blir ofte svært tydelig i spørreundersøkelser. Ta for eksempel spørsmålet:

*”Har du hatt eller blitt behandlet for influensa eller forkjølelse det siste året?”*

I formuleringen har en tatt inn to sykdommer og to behandlingsgrader i samme utsagn. Ettersom spørreundersøkelsen her bare tillater to reelle svaralternativer (ja/nei) blir det ikke mulig å svare på spørsmålet. Det ville kreve fire separate spørsmål. Skal en svare på et slikt sammensatt spørsmål må dette gjøres tekstlig med et svar som tar for seg delene separat.

Det samme problemet oppstår om et sammensatt utsagn skal brukes som utgangspunkt for kvalitativ analyse eller beslutning, som tilfellet er i målformuleringen nedenfor. Eksemplet gjelder det såkalte ”kampflyprosjektet”. Det strategiske målet som skal begrunne anskaffelsen av nye kampfly i Norge er formulert på følgende måte:

*”Kampflykapasitet for å bidra til nasjonal trygghet, suverenitet og handlefrihet”*



Formuleringen inneholder ikke mindre enn fire mål, som samlet kanskje gir en viss mening, men som isolert sett består av fire helt forskjellige mål på ulike ambisjonsnivå sett i forhold til tiltaket, dvs. anskaffelsen av et antall jagerfly. Om en ønsker å analysere prosjektets strategi vil en måtte ta for seg hvert element separat. Det er faktisk slik at strategien ikke kan analyseres uten å splitte opp de separate meningsbærende elementene i slike komplekse utsagn. Dette gjelder ikke bare strategianalyse, men også for eksempel om en ønsker å beskrive og teste ut forskjellige scenarier, vurdere risiko etc. Målformuleringer er ofte svært sammensatt, dels fordi en ikke har en klar samlende begrunnelse for prosjektet, dels fordi en ønsker å få med alle gode hensikter uten at en ser implikasjonene av dette. I slike tilfeller er det nødvendig å studere kausalitet, det vil si sammenhengen mellom målene. Vi skal se nærmere på kampflyeksemplet nedenfor.

## 10.6 Rasjonelle valg, kausalitet og sannsynlighet

Det som betegnes som logisk, i vid forstand, er det som er i samsvar med fornuften, eller allment aksepterte prinsipper for rasjonell tenkning og handling. Det som ikke er i samsvar er ulogisk. En feilslutning er en samlebetegnelse for argumenter som har en logisk form som ikke er gyldig. Logikk som vitenskap arbeider med prinsippene for det rasjonelle og anvendelsen av disse. Dette gjelder ikke minst innenfor språkområdet: Hvilke ord vi bruker, hvordan vi kombinerer disse og hvilken mening det gir. Når vi planlegger en handling eller et prosjekt er det vanlig å gjøre dette på et rasjonelt grunnlag. To viktige prinsipper som legges til grunn for analysen og vurderingen av rasjonalitet er *kausalitet* og *sannsynlighet*.

*Kausalitet*, eller årsak-virkning-sammenheng hjelper oss til å avgjøre hvilken handling som må iverksettes for å oppnå en ønsket effekt. En kan tenke seg at ulike handlinger kan gi samme effekt, Valget av hvert av de handlingsalternativene som fører frem til ønsket effekt vil da definisjonsmessig være rasjonelle. Hvis de alternative strategiene er ulike for eksempel med hensyn til hva som vil medgå av tid og ressurser, så vil den strategien med minst forbruk av ressurser være det rasjonelle valg. En kan også tenke seg at en bestemt handling kan gi ulike effekter i tillegg til den ønskete effekten. Dette kompliserer vurderingen ettersom det nå er andre årsak-virknings-forhold som også må trekkes inn betraktningen. Noen av sideeffektene kan være uønsket og i enkelte tilfeller uakseptable. Et rasjonelt valg må ta med en avveining som tar hensyn til betydningen av uønskete effekter og kanskje eliminere strategier som gir uakseptable sideeffekter.

Denne typen rasjonalistisk tenkning lar seg lett applisere innenfor fysiske systemer, men er langt mer problematisk innenfor sosiale systemer. Årsaken er at det som skjer innenfor et fysisk system følger fysiske lover og derfor i prinsippet er forutsigbart. Det som skjer innenfor sosiale systemer, det vil si samfunnet, er i prinsippet ikke forutsigbart. Årsaken til dette er at enhetene kan ta egne beslutninger. Forsøk på storstilt rasjonalistisk planlegging vil derfor i større eller mindre grad mislykkes. Forholdet mellom årsak og virkning er altså mer problematisk enn beskrevet ovenfor. Det som skjer i et samfunn følger ikke nødvendigvis endimensjonale årsak-virknings-kjeder eller todimensjonale handlingstrær. Det kan bedre beskrives i form av dynamiske systemer med innbyrdes påvirkning mellom elementene i systemet. Slike systemer kan beskrives matematisk og til en viss grad simuleres, men erfaring tilsier at resultatet ofte er nedslående, nettopp fordi systemet i prinsippet er uforutsigbart.

En enkel tilnærming til dette problemet, som opplagt kan anvendes på et tidlig tidspunkt i en planleggingsprosess, er å holde fast ved forsøket på å kartlegge enkle årsak-virkningsforhold, men trekke inn sannsynlighetvurderinger for å ta hensyn til den usikkerheten som påvirker forholdet. I så fall vil en strategi være rasjonell dersom sannsynligheten for å lykkes er minst like høy som sannsynligheten for at noen av de alternative, likeverdige strategiene skal lykkes.

Utfordringene i denne situasjonen ligger da i å identifisere de viktigste årsak-virkningsforholdene, og gi en mest mulig realistisk vurdering av sannsynlighet. Erfaringer viser at utfordringene kanskje er større enn en tror. Samtidig later det til, om en ser på vanlig praksis i planlegging av prosjekter i dag, at terskelen for forbedring ligger meget lavt og at mulighetene for betydelig forbedring er store. For eksempel var dette konklusjonen i en studie av 30 store internasjonale prosjekter som analyserte årsak-virkningskjedene som formelt lå til grunn for styringen av prosjektene, slik de var formulert i styringsdokumentene (Samset, 2006). Det innebar at målene var godkjent både av finansierende og gjennomførende part. Prosjektene var

et spesielt utvalg av hva en må anta var ”best-case” prosjekter, ettersom de alle var utformet etter samme metode i et ”log-frame”-format, se figur 4.2 og omtale i kapittel 4.

Undersøkelsen viste at alle prosjektene hadde vesentlige feil i årsak-virkningskjedene. Målene var gjennomgående beskrevet i sammensatte uttrykk, og det var sammenblanding og forveksling av mål på ulike ambisjonsnivå i hvert eneste prosjekt. Enkelte manglet beskrivelse av forventet effekt, eller hadde strategiske mål som var langt mer ambisiøse enn realistisk. Når ”økt levestandard” er ført opp som forventet effekt av et marginalt veiprosjekt og ”økonomisk vekst” som det en venter skal komme ut som resultat av et planteskoleprosjekt forstår en intuitivt at avstanden mellom årsak og virkning er for stor og at målene er for ambisiøse i forhold til tiltaket. Feilene i utformingen av disse prosjektene var så omfattende at en konkluderte med at ingen av styringsdokumentene egnert seg for styring og overordnet beslutning. Det som var oppsiktsvekkende var at de fleste feilene var trivielle og burde vært unngått, ettersom en i utformingen av alle prosjektene hadde benyttet en velprøvd metode som tok sikte på nettopp å unngå slike feil.

## 10.7 Tre eksempler på målformuleringer

Prinsippene som er beskrevet er anskueliggjort nedenfor ved hjelp av informasjon i tidligfasen fra tre meget store offentlige prosjekter: (1) Anskaffelsen av nye kampfly, (2) Bygging av Stad skipstunnel, og (3) samferdselsprosjektet Lofast 2 i Lofoten.

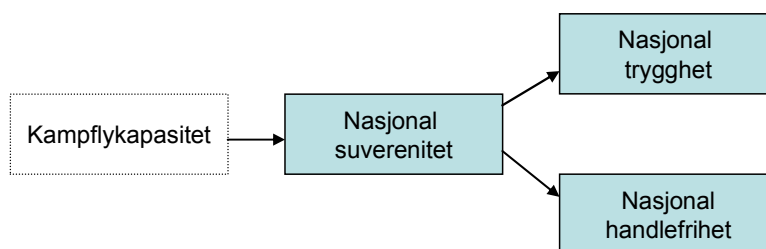
### Anskaffelse av nye kampfly

Se nærmere på mål-formuleringen i kampflyeksemplet:

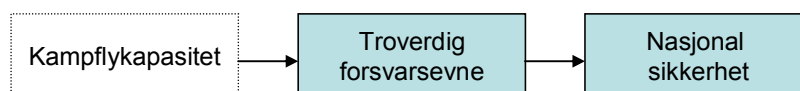
*”Kampflykapasitet for å bidra til nasjonal trygghet, suverenitet og handlefrihet”*

Som nevnt ovenfor inneholder denne fire mål. Disse er koplet sammen antakelig på en tilfeldig måte. Uttrykket ”å bidra til ” fastslår imidlertid kausalitet. Det vil si at ”kampflykapasitet” er årsak og det som står i resten av setningen angir forventet effekt. Nå er både nasjonal trygghet, suverenitet og handlefrihet noe vi allerede har, slik at uttrykket ”bidra til” blir feil, her bør det stå ”oppretholde”. Formuleringen blir da: *Kampflykapasitet for å opprettholde nasjonal trygghet, suverenitet og handlefrihet.* Når det gjelder målet *kampflykapasitet* så forteller det oss ikke noe annet enn det som er uttrykt som prosjektets resultatmål, nemlig at et antall jagerfly er levert og tatt i bruk. Dette målet er realisert ved overlevering og derfor ikke et strategisk mål. Det skal fjernes i målformuleringen.

De gjenværende tre målene er skilt med komma, som ikke gir noen indikasjon på kausalitet. Spørsmålet er da om de skal oppfattes som sideordnet eller parallelle mål. For å svare på det må en se om det kan være en årsak-virkning-sammenheng mellom disse. En kan tenke seg at suverenitet, ettersom det er noe vi allerede har er det som først og fremst gir oss nasjonal trygghet, men også handlefrihet. Et forslag vil da være å beskrive årsak-virkningskjeden som nedenfor. Andre vil vurdere sammenheng annerledes.



Valg av strategisk mål står da mellom de tre gjenværende, og spørsmålet blir på samme måte som i diskusjonen om validitet en sannsynlighetsvurdering: Gir Kampflykapasitet et valid bidrag til realiseringen av henholdsvis nasjonal handlefrihet, nasjonal suverenitet eller nasjonal trygghet. Eller sagt på en annen måte: Gitt tilstrekkelig Kampflykapasitet, hvor stor er sannsynligheten for at hvert av de tre målene realiseres? Her er det kanskje *nasjonal trygghet*, hva en nå måtte mene med det, som er det målet som ligger nærmest. Imidlertid er nasjonal trygghet antakelig mer omfattende enn trygghet mot angrep fra militær makt utenfra, og omfatter også for eksempel spørsmål om økonomi, økologi, matsikkerhet, selvforsyning etc. Kan hende er *nasjonal sikkerhet*, eller kanskje *troverdige forsvarsevne* i betydningen avskrekking mot militær intervensjon fra andre land, et mål en kan samles om ettersom dette ut fra samme betraktning som ovenfor er på et lavere og mer realistisk ambisjonsnivå, sett i relasjon til tiltaket. I så fall vil det være rimelig å benytte dette som strategisk mål til erstatning for de andre, som mer er å oppfatte som honnørord som har sneket seg inn stortingsproposisjonen i bestrebelsene for å begrunne en offentlig investering - noe som ikke er uvanlig. Årsak-virkningskjeden blir da:



Kampflykapasiteten blir en av flere elementer i det som skal fremstå for andre nasjoner som en troverdige forsvarsevne, hvorav vår deltakelse i en større allianse er et vesentlig element. Den overordnede begrunnelsen for dette er uttrykt ved det strategiske målet, nemlig å opprettholde vår nasjonale sikkerhet.

### Stad skipstunnel

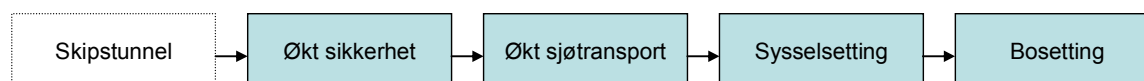
Et annet prosjekt med ennå mer sammensatt målformulering er "Stad skipstunnel" der en tar sikte på å bygge en tunnel for mindre fartøyer mellom to fjordarmer på strekningen mellom Ulsteinvik og Måløy. Prosjektets strategiske mål er uttrykt på denne måten:



*”Styrke rammeforutsetningen for sjøtransport langs kysten ved å øke tilgjengelighet og sikkerhet for denne type transport i regionen, samt positivt legge til rette for lokal sysselsetting og bosetting”*

Formuleringen inneholder i alt fem meningsbærende elementer, det vil si: (1) rammeforutsetningen for sjøtransport, (2) økt sjøtransport, (3) bedret sikkerhet, (4) lokal sysselsetting, og (5) lokal bosetting. Uttrykket ”ved å” peker tilbake på det første målet, i motsetning til forrige eksempel der en benyttet uttrykket ”for å” som peker fremover. Slik det står betyr dette altså at rammeforutsetningen for sjøtransport blir virkning mens det resterende blir årsak. Dette er åpenbart feil. Det er ”Styrket rammeforutsetning for sjøtransporten”, hva en nå måtte mene med det, som skal føre til at en får økt transport, bedre sikkerhet, sysselsetting og bosetting, ikke omvendt.

En nøktern fortolkning av begrepet ”rammeforutsetning” i dette tilfellet vil være at det betyr selve skipstunnelen. Et velkjent og godt råd anbefaler er at en skal kalle en spade en spade, slik at en da få følgende årsak-virkningskjede:

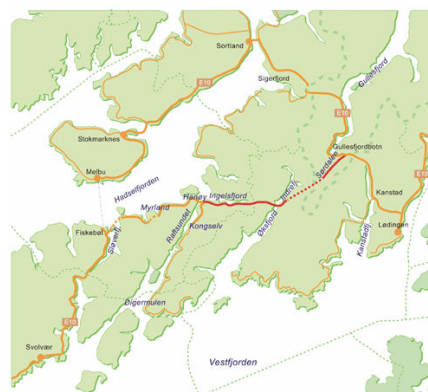


På samme måte som i forrige eksempel kan skipstunnelen (eller her rammeforutsetningen for sjøtransporten) strykes ettersom dette ikke er noe annet enn prosjektets resultatmål. *Bedret sikkerhet og økt sjøtransport* på strekningen er det rimelig å tro må oppfattes som første ordens effekt av prosjektet. En står da igjen med som strategiske mål å opprettholde *sysselsetting*, noe som er en forutsetning for *bosetting*.

Dette er den logiske strukturen, som etterlater et spørsmål som må avklares kanskje mer på et politisk grunnlag enn logisk: Hva er det som begrunner dette prosjektet? Er det bosettingsproblemet, og i så fall, er da skipstunnelen et tiltak som med tilstrekkelig sannsynlighet vil sikre at denne opprettholdes? Eller er det sysselsettingen som er det sentrale, og i så fall, kan prosjektet rettferdiggjøres ut fra en større samfunnsøkonomisk vurdering av effekten av forventet økning i transporten med mindre fartøyer. Stortingsproposisjonen, slik samfunns målet er formulert, gir ikke noe klart svar på dette og derfor ingen troverdig begrunnelse for prosjektet.

## Lofast 2

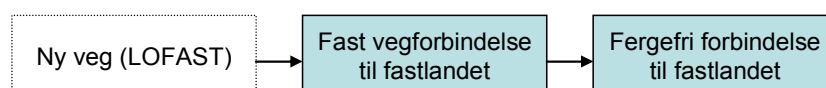
Det tredje prosjektet er ”Lofast 2” som gjelder fastlandsforbindelsen til Lofoten fra Raftsundet til Gullsfjord, og som er planlagt ferdig ved utgangen av 2007. Stortingsproposisjonen omtaler resultatmålet konkret, det vil si ”om lag 30 kilometer ny veg inklusive 4 tunneler og 2 større og 9 mindre bruer”. Imidlertid finner en knapt noe effektmål eller strategisk mål for prosjektet i dokumentet. Det nærmeste en kommer en begrunnelse er følgende utsagn:





*”LOFAST vil gi en region med ca. 25 000 innbyggere fast vegforbindelse til fastlandet, og dermed gi fergefri forbindelse til fastlandet for Lofoten.”*

Denne setningen inneholder tre meningsbærende elementer: (1) veien som skal bygges, (2) fast veiforbindelse til fastlandet og (3) fergefri forbindelse til fastlandet. Lenkeordene ”gi” og ”dermed gi” fastslår årsak-virkning-sammenheng mellom disse som følger:



Den nye veien er selvsagt ikke et strategisk mål men bare resultatmålet. De to andre målene er ikke annet enn to forskjellige uttrykk for det samme og hele utsagnet og kausaliteten blir følgelig meningsløs. Begge målene realiseres i det øyeblikk veien er bygget. Det betyr at prosjektet ikke har noe overordnet mål. Det at 25000 innbyggere får nytte av veien er meget sannsynlig, men stortingsproposisjonen gir ingen beskrivelse av hvilken effekt en mener dette vil gi, som kan begrunne hvorfor prosjektet bør gjennomføres. Dette er heller ikke uvanlig i offentlige beslutningsdokumenter.

Som nevnt innledningsvis benytter en i disse eksemplene vurdering av kausalitet og sannsynlighet som hjelpemiddel. Begge baseres på skjønn. Årsak-virkningskjeden er en logisk konstruksjon, det vil si basert på erfaring og sunn fornuft. Sannsynlighetsvurderingen som følger er nødvendig for å teste rimeligheten eller styrken i leddene i kjeden.

Utgangspunktet i alle tre eksemplene er prosjektenes resultatmål, som det er lett å identifisere - det er prosjektets leveranse ved overlevering. Sannsynlighetstesten på godheten av resultatmålet er ganske enkelt at sannsynligheten for realisering skal være lik 1, gitt de ressursene som står til rådighet for gjennomføringen. I eksemplene er resultatmålet henholdsvis flyanskaffelsen, skipstunnelen og veiforbindelsen med broer og tunneler.

Når det gjelder realiseringen av øvrige mål er sannsynligheten i prinsippet mindre enn 1. Dette fordi realiseringen av effektmål og samfunns mål avhenger ikke bare av prosjektet, men av andre forhold som ligger utenfor prosjektet og som det er knyttet usikkerhet til. Et realistisk mål er et mål der prosjektet gir et vesentlig bidrag til realiseringen, det vil si at sannsynligheten for realisering er relativt høy. Dersom sannsynligheten for realisering er lav, innebærer det at målet er for ambisiøst i forhold til prosjektet, og at en bør velge et annet mål på et lavere ambisjonsnivå.

Vurderingen i tilfellet kampflyprosjektet var at bidraget fra prosjektet var marginalt i forhold til for eksempel et overordnet mål som nasjonal selvstendighet. Det ble derfor ut fra en sannsynlighetsvurdering foreslått at målene burde erstattes med et effektmål og et strategisk mål som var mer realistiske, i dette tilfellet henholdsvis *troverdige forsvarsevne* og *nasjonal sikkerhet*.

I eksemplet Stad skipstunnel er det rimelig å anta at denne med meget høy sannsynlighet vil øke sikkerheten for brukerne fordi de unngår overfarten over Stadhavet. Dette betyr at ”økt sikkerhet” er for lite ambisiøst som effektmål. Det er trolig noe mer usikkerhet om sjøtransporten vil øke vesentlig, det avhenger blant annet av behovet og avgiftsnivået. Dersom en antar at sannsynligheten for realisering av målet er tilstrekkelig høy, gitt at tunnelen er bygget, vil dette være et hensiktsmessig effektmål. Ennå mer usikkerhet er knyttet til i hvilken grad prosjektet vil bidra til økt sysselsetting og bosetting. Dersom sannsynligheten for realisering av disse målene er lav, gitt at tunnelen er bygget, er det grunn til å lete etter andre mål som kan gi en bedre begrunnelse for prosjektet. Dersom dette ikke er tilfellet bør prosjektet forkastes.

I veiprojektet er det også en sannsynlighetsvurdering som ligger til grunn for den reduksjonen som ble foretatt. Sannsynligheten for ny vei er 1, gitt at prosjektet realiseres. Det samme gjelder med hensyn til at vegforbindelsen skal være fergefri. Det betyr to ting: Det finnes ikke noen årsakssammenheng mellom målene, og derfor heller ikke noe som gir en overordnet begrunnelse for investeringen.

## 10.8 Konklusjon

Kvalitativ informasjon er ikke noe en velger å bruke i en gitt situasjon, det er noe vi alle bruker i vår daglige kommunikasjon. Troverdigheten av innholdet i kommunikasjonen avhenger av hva informasjonen bygger på. Mye av informasjonen vi bruker er basert på underliggende systematisk analyse av fakta, ofte kvantitativ informasjon. Andre deler av informasjonen bygger på antakelse og skjønn. I en innledende vurdering av slike forhold som behov og antatt effekt av mulige prosjektalternativer vil en i betydelig grad være henvist til å bruke kvalitativ informasjon som bygger på antakelser og skjønn. Dette er ikke bare en ulempe. Det at en i stor grad opererer med begreper og ikke med tall kan gjøre det lettere å visualisere og drøfte komplekse forhold. Presisjonsnivået i slike tilfeller blir følgelig lavt. Dette gjør at en må være spesielt oppmerksom på å sikre definisjonsmessig validitet for de begrepene som brukes for å sikre at den informasjonen som skal benyttes som utgangspunkt for konsept- og prosjektutviklingen videre er mest mulig entydig og konsistent. Mye kvalitativ informasjon fremstår i form av komplekse, sammensatte utsagn. Analyse av slik informasjon forutsetter at den relevante teksten brytes ned i sine enkelte meningsbærende elementer. Mye av analysen består i *kategorisering* (for eksempel med SWOT-analyse), i å identifisere eller bygge *struktur* (for eksempel med log-frame-analyse), eller å vurdere *egenskap* (for eksempel i risikoanalyse). I slike tilfeller er vurdering av kausalitet og sannsynlighet viktige hjelpemidler, og bruken av dette er illustrert i dette kapitlet, ovenfor.

## 10.9 Referanser

Bazerman, Max, 1994, *Judgment in Managerial Decision Making*, John Wiley & Sons, Inc., New York

Bennett, Deborah, 2004, *Logic made easy*, Penguin Books, London

Goodwin, Paul and Wright, George, 1996, *Decision Analysis for Management Judgment*, John Wiley & Sons, Inc., London

- Gullvåg, Ingemund, 1990, *Rasjonalitet, forståelse og forklaring. Innføring i argumentasjonsteori, logikk og vitenskapsfilosofi*, Tapir Akademisk Forlag, Trondheim
- Hansen, Sven Ove, 2003, *Konsten att vara vetenskaplig*, Kompendium, Filosofienheten, Kungliga Tekniska Högskolan, Stockholm
- Hellevik, Ottar, 1991: *Forskningsmetode i sosiologi og statsvitenskap*, Universitetsforlaget, Oslo.
- Hendricks Vincent F., 2006, *Thoughts 2 Talk. A Crash Course in Reflection and Expression*, [www.formalphilosophy.com](http://www.formalphilosophy.com), Automatic Press
- Klakegg, Ole Jonny, 2004, *Målformulering i store statlige investeringsprosjekt*, Concept rapport Nr. 6, Concept-programmet, NTNU, Trondheim
- Næss, Petter m.fl., 2004, *Bedre utforming av store offentlige investeringsprosjekter. Vurdering av behov, mål og effekt i tidligfasen*, Concept rapport Nr. 9, Concept-programmet, NTNU, Trondheim
- Olsson, Henny og Sörensen, Stefan, 2003: *Forskningsprosessen. Kvalitative og kvantitative perspektiver*. Gyldendal Akademisk, Oslo ISBN 82-05-30185-9
- Samset, Knut, 1999, *The Logical Framework Approach (LFA), Handbook for objectives-oriented planning*, 4th edition, Norad, Oslo
- Samset, Knut, 2001, *Prosjektvurdering i tidligfase, Fokus på konseptet*, Tapir Akademisk Forlag, Trondheim
- Samset, Knut, 2006: Design of high-uncertainty projects in international aid, paper presentert på konferansen "PROMAC 2006", Sydney
- Teigen, Karl Halvor, 2006: *Skjønn og skivebom. Hvordan vi bedømmer usikkerhet*, Foredrag, Norsk Senter for Prosjektledelse, Oslo



## Del 4

# Anvendelse av informasjon – Fra informasjon til beslutning



## 11 EN VERKTØYKASSE FOR ANALYSE I PROSJEKTERS TIDLIGFASE

BJØRN ANDERSEN\*

### 11.1 Innledning

Så langt har boken sett på utfordringer i prosjekters tidligfase og hvordan man kan fremskaffe informasjon av god kvalitet i denne fasen. Denne informasjonen blir først nyttig om den anvendes som datagrunnlag for ulike analyser, som i sin tur utgjør grunnlaget for å fatte beslutninger. Hensikten med dette kapittelet er å presentere ulike analyseteknikker som kan anvendes til dette formålet.

Det finnes et stort antall ulike analyseteknikker tilgjengelig. Disse teknikkene har typisk vært utviklet for ulike formål, deretter tatt i bruk i andre sammenhenger, gjerne i noe tilpasset form. Så vidt undertegnede bekjent finnes det dog ingen verktøy utviklet spesifikt for prosjekters tidligfase. Dermed må det gjøres et utvalg av egnede analyseverktøy fra andre fagområder.

Typiske behov for verktøy til analyse og beslutning i et prosjekts tidligfase omfatter (minst) å kunne:

- Forutse effekter (positive og negative) av ulike beslutningsalternativer.
- Evaluere/sammenlikne ulike beslutningsalternativer.
- Vurdere klimaet i prosjektet og dets omgivelser og krefter for og imot gjennomføring av ulike beslutningsalternativer.
- Analysere alternativets effekt på de ønskede mål/effekter prosjektet skal oppnå.

Relevante analyseteknikker for disse ulike formålene eksisterer både innenfor fagområder som strategisk planlegging, kvalitetsledelse, samfunnsøkonomiske analyser, endringsledelse og også generell prosjektledelse. Mange teknikker anvendes etter hvert innenfor et stort antall områder og det kan være vanskelig å lenger vite hvor de egentlig har sitt opphav. For vår del er det heldigvis ikke spesielt viktig. Hensikten har vært å finne et antall analyseteknikker/verktøy som dekker de ovennevnte behov. Det endelige utvalget som presenteres her er selvsagt et resultat av subjektiv siling og utvelgelse, og man kan bare håpe at leseren vil være enig i at de fyller behovet.

---

\* Bjørn Andersen er forskningssjef ved SINTEF

Verktøyene vil for øvrig presenteres i en rekkefølge som tilsvarer hvilket av disse behovene de adresserer. For hvert verktøy omtales:

- Teknikkens logikk og anvendelser i tidligfase vurderinger i prosjekter.
- Praktisk gjennomføring av analyser ved bruk av teknikken.
- Et kort og sterkt forenklet eksempel på bruk av teknikken i et prosjekt.
- ”Advarende pekefinger” om begrensninger i bruk av teknikken.

## 11.2 Teknikker for å forutse effekter av ulike beslutningsalternativer

En svært utbredt problemstilling i prosjekters tidligfase er å måtte velge mellom flere ulike beslutningsalternativer, for eksempel ulike løsningskonsepter, forskjellige partnere eller geografiske lokasjoner. Et sentralt grunnlag for å fatte en slik beslutning er innsikt i hvilke fordeler og ulemper de respektive alternativer forventes å kunne medføre. Tre teknikker som fra ulike synsvinkler belyser dette spørsmålet er:

- Interessentanalyse kombinert med Kano-modellen for behovsanalyse, for å identifisere interessenter for et prosjekt og de ulike konseptalternativer samt interessentenes uttalte og ikke-uttalte behov og forventninger.
- SWOT-analyse, svært enkel systematikk for å få frem styrker/svakheter/muligheter/trusler ved hvert beslutningsalternativ.
- Usikkerhetsanalyse, med en enkel logikk for identifisering av usikkerhetselementer og klassifisering av disse i forhold til sannsynlighet for å inntreffe og konsekvens.

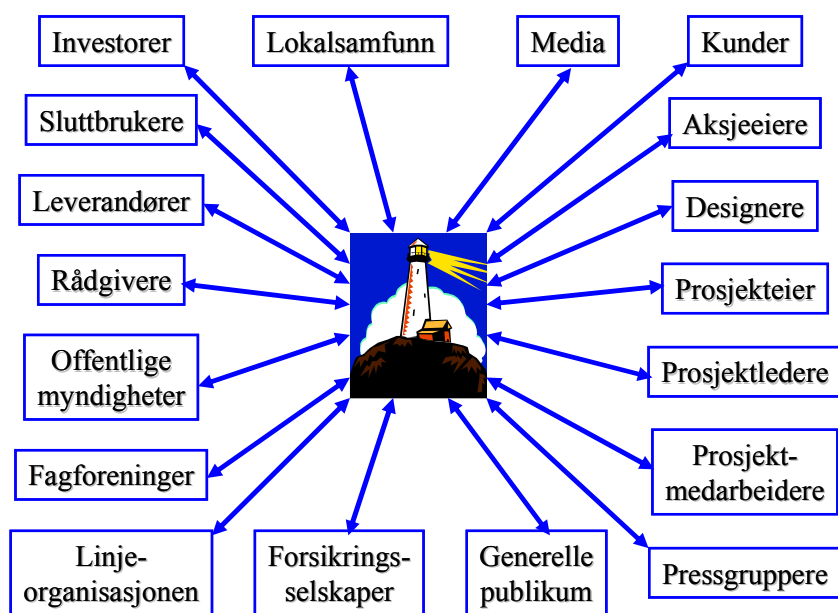
### Interessent- og behovsanalyse

Ethvert prosjekt er “omringet” av et antall organisasjoner eller individer som direkte eller indirekte deltar i eller påvirker utforming og utførelse av prosjektet og som selv påvirkes av dets utfall. Disse benevnes med et fellesbegrep interessenter (engelsk stakeholders), som av PMI defineres som (PMI, 2000):

*Individuals and organizations who are actively involved in the project, or whose interests may be positively or negatively affected as a result of project execution or successful project completion.*

Noen typiske interessenter for et prosjekt er vist i Figur 11-1 , hvor noen av disse ikke vil ha en aktiv rolle i forhold til prosjektet før i gjennomføringsfasen eller enda senere.





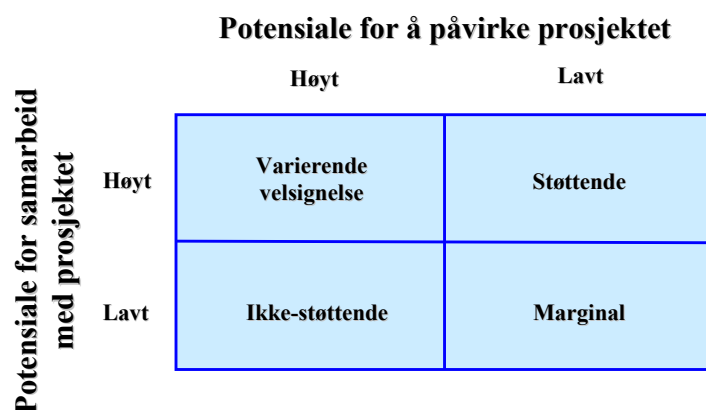
Figur 11-1 Prosjektinteressenter

En rekke empiriske funn viser tydelig at interessenter spiller en avgjørende rolle i utforming og gjennomføring av prosjekter. Analyser av et stort antall prosjekter viser også at interessenter skaper omtrent like mye vanskeligheter som de støtter prosjektet. Spørreundersøkelser har vist at de fleste prosjekter med problemer i ettertid innser at altfor lite arbeid ble investert i forståelse av interessenter og ikke minst utvikling av tiltak for å håndtere dem. Dermed representerer interessenter en svært viktig ”analyseenhet” i et prosjekts tidlige fase, og ved vurdering av ulike beslutningsalternativer vil en analyse av hvordan prosjektets nåværende og fremtidige interessenter vil forholde seg til de ulike alternativene være et nyttig hjelpemiddel. Dette vil da typisk innebære å kartlegge prosjektets interessenter og deres behov og deretter forsøke å forutse interessentenes reaksjon på hvert av de ulike alternativene som vurderes.

I praksis vil dette normalt innebære følgende trinn i analysen (Andersen og Fagerhaug, 2002):

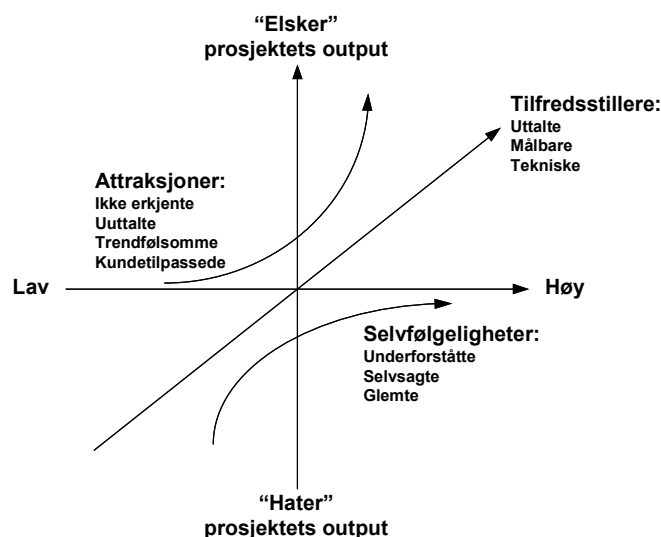
1. Identifisering av prosjektets eksisterende og antatte fremtidige interessenter.
2. Klassifisering av interessentene i forhold til viktighet og forventet holdning til prosjektet.
3. Identifisering av de ulike interessenters behov og forventninger i forhold til prosjektet og dets resultater.
4. Analyse av interessentenes preferanser i forhold til de beslutningsalternativer som foreligger.

Identifisering av interessenter gjøres normalt simpelthen gjennom brainstorming, helst med deltakere fra ulike deler av prosjektets organisasjon eller initiativtakere. En slik øvelse vil ofte resultere i en lang liste av interessenter, og det er åpenbart at ikke alle disse er like viktige for prosjektet. For å begrense omfanget av analysen er det derfor ønskelig å utpeke de få interessenter som vil være viktigst å vurdere. Til dette finnes det flere sett av kriterier, der en av de mest anvendte metodene bygger på matrisen i Figur 11-2.



Figur 11-2 Firefeltsmatrise for klassifisering av interessenter

Interessentene som plasserer seg i de to venstre feltene vil normalt representere de aktørene det er viktigst å sikre seg at aksepterer det valg som fattes blant beslutningsalternativene. I den videre analysen er det derfor ofte tilstrekkelig å arbeide videre med bare disse. Neste trinn er å skape innsikt i hvilke behov og forventninger disse viktige interessentene har i forhold til prosjektet og dets utvikling. Denne vurderingen kan med fordel baseres på den såkalte Kano-modellen, som tydeliggjør at interessenters behov bør vurderes innenfor tre kategorier; selvfølgeligheter, tilfredsstillere og attraksjoner (Akao, 1990). Disse gruppene illustreres i Figur 11-3, der Kanos poeng er at man ikke må overse behov i de lett glemte kategoriene selvfølgeligheter og attraksjoner. Fullstendig innsikt i en interessents holdning til prosjektet fordrer at man tenker gjennom behovselementer innenfor hver kategori.



Figur 11-3 Tre kategorier av behov

Etter å ha kartlagt forventninger og behov ved bruk av Kano-modellen, avsluttes analysen med å anvende denne kunnskapen til å analysere hvordan den enkelte interessent vil vurdere hvert av de foreliggende beslutningsalternativene. Dette kan enten gjøres "verbalt" i form av å diskutere hvordan man forventer at de ulike interessentene, gitt deres identifiserte

forventninger og behov, vil reagere på hvert av alternativene. Vil man være enda mer systematisk kan dette settes opp i en tabell med interessenter langs ene akse og beslutningsalternativer langs den andre. I hver rute nedtegnes så hvordan man antar den enkelte interessent vil respondere på det ulike alternativ. Uansett fremgangsmåte vil utfallet være en innsikt i hvordan man kan forvente at prosjektets interessenter vil like eller mislike alternativene som vurderes. Dette utgjør dermed et grunnlag for å fatte en beslutning, ut fra en målsetning om å maksimere tilfredsheten hos de viktigste og mest innflytelsesrike interessentene.

En kommune vurderer utbygging av et nytt, større boligfelt. Tre ulike geografiske områder er identifisert som egnede alternativer:

- Nord, som ligger øde til og vil kreve bygging av en lang atkomstvei, men ellers ikke vil berøre verken andre boligområder eller næringsvirksomhet.
- Øst, som vil medføre en utvidelse av et eksisterende boligområde, men både vil kreve utvidelse av eksisterende veinett og flytting av en mindre industrivirksomhet.
- Sør, hvor veinett og annen infrastruktur er velegnet, men som vil medføre konvertering av en park/rekreasjonsområde til boligformål.

Kommunen erkjenner at de må forstå hvordan ulike aktører i kommunen vil forholde seg til de ulike alternativene og gjennomfører en interessentanalyse. I første omgang brainstormes de mest åpenbare interessentene:

- De fremtidige beboerne i boligfeltet som planlegges etablert.
- Innbyggerne, i kommunen generelt og i nærhet til de områder som vil berøres av de ulike områdene.
- Næringslivet, i kommunen generelt og i nærhet til de områder som vil berøres av de ulike områdene.
- Entreprenører og andre som vil bidra til selve utbyggingen.
- Grunneiere som sitter på tomter som vil berøres av en utbygging.
- Media, som kan komme til å omtale planene og gjennomføringen i lokalpressen.
- Interesseorganisasjoner som kan komme til å fatte interesse for prosjektet, for eksempel miljøvernorganisasjoner.

Man er klar over at det vil finnes en rekke andre, men nøyer seg med disse. Videre forsøker man å avgjøre hvilken kategori av interessent de ulike representerer og hvilke forventninger de vil ha til prosjektet. Dette er oppsummert i tabellen under.

<b>Tabell 1. Interessenter og behov for nytt boligfelt</b>				
<b>Interessenter</b>	<b>Kategori</b>	<b>Selvfølgeligheter</b>	<b>Tilfredsstillere</b>	<b>Attraksjoner</b>
<b>Fremtidige beboere</b>	<b>Støttende</b>	<b>Atkomst til bolig</b>	<b>Tomtestørrelse, uforstyrret beliggenhet</b>	<b>Utsikt</b>
<b>Innbyggerne</b>	<b>Ikke-støttende</b>	<b>Nytt felt medfører ingen ulemper</b>	<b>Prosjektet gjennomføres uten negative konsekvenser for kommunens økonomi</b>	<b>Nytt boligfelt fører til bedret servicetilbud i kommunen</b>
<b>Næringslivet</b>	<b>Varierende velsignelse</b>	<b>Nytt felt medfører ingen ulemper</b>	<b>Økt omsetning som følge av økt tilflytning</b>	<b>Deltakelse i oppdrag ved utbyggingen</b>
<b>Entreprenører</b>	<b>Varierende velsignelse</b>	<b>Alle reguleringer og tillatelser er på plass</b>	<b>Økonomisk gevinst som følge av oppdrag</b>	<b>Læring, markedsføring grunnet oppdrag</b>
<b>Grunneiere</b>	<b>Varierende velsignelse</b>	<b>Ikke bli ofre for ekspropriasjon</b>	<b>Pris oppnådd for avgitte tomter</b>	<b>Bli involvert i planlegging av utbyggingen</b>
<b>Media</b>	<b>Varierende velsignelse</b>	<b>Få tilgang på relevant informasjon</b>	<b>Adgang til møter og diskusjoner</b>	<b>Økt avissalg grunnet interesse for prosjektet</b>
<b>Interesseorganisasjoner</b>	<b>Marginal</b>	<b>Få tilgang på informasjon man har krav på</b>	<b>Tas med i høringer</b>	<b>Få reell påvirkning på planene</b>

Ut fra dette finner man at spesielt innbyggerne i kommunen, næringslivet, grunneiere og aktører som skal delta i utbyggingen er viktige å forholde seg til på en ryddig måte. For alternativ Nord finner man at de eneste negative reaksjoner dette kan føre til er at prosjektkostnadene blir høyere grunnet behovet for ny vei, mens det ellers vil vurderes som gunstig av de aller fleste. Alternativ Øst vil sannsynligvis bli mottatt negativt av både beboerne i det eksisterende boligfeltet samt næringsvirksomheten som må flyttes. Sør vil ganske sikkert skape sterke, negative reaksjoner fra både innbyggere generelt, media og interesseorganisasjoner.

Gitt at utbyggingskostnadene for alternativ Nord blir signifikant mye høyere enn for Øst, konkluderer man med at Nord er å foretrekke.

Dette enkle eksempelet illustrerer ganske godt noen av begrensningene ved en slik interessentanalyse. For det første vil vurderingene rundt både holdning til samarbeid og påvirkningsevne til interessentene og ikke minst deres forventninger og behov være basert på subjektive antakelser hos dem som gjennomfører analysen. Disse kan være høyst unøyaktige og dermed gi et unøyaktig eller direkte feil resultat fra analysen. Dette kan imidlertid motvirkes ved å gjennomføre intervjuer med noen av eller alle interessentene for å sikre at antakelsene blir så nøyaktige som mulige. Videre vil analysen til slutt bestå av et stort antall ulike vurderinger, sannsynligvis med mange faktorer som trekker i ulike retninger for de forskjellige alternativene. Derfor kan det være vanskelig å trekke en entydig konklusjon fra analysen. Resultatet er dermed at en slik analyse ofte bidrar til å skape en bedre innsikt i beslutningssituasjonen, men at den ofte bør suppleres med andre typer vurderinger for å kunne fatte en beslutning.

## SWOT-analyse

SWOT-analyse har fått navnet sitt som et akronym bygget på de fire engelske analyseperspektivene *styrker* (Strengths), *svakheter* (Weaknesses), *muligheter* (Opportunities) og *trusler* (Threats). Analysen har sitt opphav i strategisk planlegging, som et verktøy for å identifisere faktorer en virksomhet må forholde seg til, innenfor disse fire perspektivene, men har blitt tatt i bruk for en rekke ulike andre formål, inkludert prosjektledelse (Lewis, 1999). Også i analyser i en tidligfase i et prosjekt kan SWOT-teknikken være til nytte, da som en enkel logikk for å strukturere styrker/svakheter/muligheter/trusler ved hvert beslutningsalternativ. Selve analysen vil da ikke frembringe noen direkte konklusjoner om hvilket alternativ som bør velges, men det vil bidra til å øke innsikten i hvert alternativ og sannsynlig utvikling for hvert av disse.

Gjennomføring av en SWOT-analyse er for øvrig svært enkelt. Det innebærer stort sett bare en brainstorming rundt de fire perspektivene. Alle råd sier at dette bør skje i en gruppe, gjerne med ulike disipliner/fagområder involvert. Oppsummert i trinn blir da SWOT-analyse i tidligfase som følger:

1. Etabler en analysegruppe som representerer ulik kompetanse og bakgrunn for å gjennomføre analysen.
2. For hvert beslutningsalternativ som foreligger, brainstorm styrker, svakheter, muligheter og trusler.
3. Oppsummer de fremkomne faktorer i en tabell med fire felter, ett for hvert av perspektivene.

Kommunen fra forrige eksempel kunne gjerne brukt SWOT-analyse til å vurdere de tre alternativene, noe som beskrives under.

Som beskrevet foreligger det tre alternativer for utbygging i kommunen:

- Nord, som ligger øde til og vil kreve bygging av en lang atkomstvei, men ellers ikke vil berøre verken andre boligområder eller næringsvirksomhet.
- Øst, som vil medføre en utvidelse av et eksisterende boligområde, men både vil kreve utvidelse av eksisterende veinett og flytting av en mindre industrivirksomhet.
- Sør, hvor veinett og annen infrastruktur er velegnet, men som vil medføre konvertering av en park/rekreasjonsområde til boligformål.

For å gjøre en SWOT-analyse av disse etableres en arbeidsgruppe bestående av leder for plan- og bygningsetaten, rådmannen, to eksperter på henholdsvis vei/samferdsel og miljøspørsmål samt to politikere fra kommunestyret.

Gruppen møtes og tilbringer en halv dag med å identifisere faktorer innenfor hvert av analyseperspektivene og for hvert av de tre alternativene. Disse er oppsummert i de tre tabellene under:

Tabell 2 SWOT-faktorer for alternativ Nord

<p><b>Styrker</b></p> <ul style="list-style-type: none"> <li>• Berører ikke eksisterende boligbebyggelse eller næringsvirksomhet</li> <li>• Gir stor frihet for utforming av det nye boligfeltet</li> </ul>	<p><b>Svakheter</b></p> <ul style="list-style-type: none"> <li>• Vil kreve bygging av nye atkomstvei, som kan bli kortbart</li> <li>• Det nye boligfeltet kan bli liggende øde til med lang vei til annen infrastruktur/tilbud</li> </ul>
<p><b>Muligheter</b></p> <ul style="list-style-type: none"> <li>• Det nye boligfeltet kan representere starten på større utbygging i dette ellers ubenyttede området</li> <li>• Mangel på eksisterende tilbud, som butikk, kan skape grunnlag for ny næringsvirksomhet</li> </ul>	<p><b>Trusler</b></p> <ul style="list-style-type: none"> <li>• Mulige protester fra ulike hold rundt å omregulere et ubenyttet område til boligformål</li> <li>• Vanskelig å tiltrekke seg boligkjøpere grunnet øde beliggenhet</li> </ul>

Tabell 3 SWOT-faktorer for alternativ Øst

<p><b>Styrker</b></p> <ul style="list-style-type: none"> <li>• Effektiv utnyttelse av areal og infrastruktur ved å utvide eksisterende boligfelt</li> <li>• Reguleringsplaner for slik utvidelse har vært utredet ved etableringen av det eksisterende feltet</li> </ul>	<p><b>Svakheter</b></p> <ul style="list-style-type: none"> <li>• Vil kreve utvidelse av atkomstvei, som kan bli kortbart</li> <li>• Vil kreve flytting av eksisterende industrivirksomhet</li> </ul>
<p><b>Muligheter</b></p> <ul style="list-style-type: none"> <li>• Økt antall beboere i området vil kunne bidra til full utnyttelse av overkapasitet i områdets skole</li> </ul>	<p><b>Trusler</b></p> <ul style="list-style-type: none"> <li>• Flytting av eksisterende industrivirksomhet kan bli en "delikat" sak</li> <li>• Dagens beboere i området kan komme til å protestere mot utvidelsen</li> </ul>

Tabell 4 SWOT-faktorer for alternativ Sør

<p><b>Styrker</b></p> <ul style="list-style-type: none"> <li>• Effektiv utnyttelse av eksisterende infrastruktur</li> <li>• Nytt område gir frihet til å utforme boligfeltet på best mulig måte</li> </ul>	<p><b>Svakheter</b></p> <ul style="list-style-type: none"> <li>• Vil ødelegge et eksisterende rekreasjonsområde</li> <li>• Det nye boligfeltet vil bli liggende noe avsides til i forhold til kommunens mer sentrale områder</li> </ul>
<p><b>Muligheter</b></p> <ul style="list-style-type: none"> <li>• Lokaliseringen gir rom for å integrere boligfeltet i naturen</li> <li>• Kan åpne for større aktivitet i denne delen av kommunen</li> </ul>	<p><b>Trusler</b></p> <ul style="list-style-type: none"> <li>• Sterke protester fra interesseorganisasjoner mot sanering av parkområde</li> </ul>

Ved å veie alle disse faktorene mot hverandre konkluderer gruppen med at den vil anbefale alternativ Nord.

På samme måte som for interessent- og behovsanalyse ser man at en SWOT-analyse bidrar til å få frem i lyset ulike vurderinger ved de alternativene som foreligger. Samtidig gir ikke analysen noen direkte anbefaling, noe som kan føre til en form for "beslutningsparalyse" etter

å ha brukt tid på å identifisere faktorer i ulike retninger. Igjen er det da viktig å forstå at de teknikkene som presenteres i dette kapittelet har ulike anvendelser og at man gjerne bør bruke flere verktøy for å til slutt komme frem til en beslutning.

### **Usikkerhetsanalyse**

Usikkerhetsanalyse, eller snarere usikkerhetsstyring, er i seg selv et omfattende tema som dekker både konkrete analysetilnærminger, psykologiske forhold rundt risiko og muligheter, prinsipper for løpende overvåkning av usikkerhet, osv. Av dette er den mest grunnleggende teknikken for usikkerhetsanalyse av størst interesse når det gjelder anvendelse i en tidligfase. Denne består av en enkel metode for å både identifisere usikkerhetselementer, vurdere hvor sannsynlig det er at hvert element vil inntreffe, angi konsekvensen om det skulle inntreffe samt utvikle tiltak for å unngå at negative elementer (risiko) skal inntreffe og øke sannsynligheten for at positiv usikkerhet (muligheter) kan utnyttes.

I så måte kan metoden minne om SWOT-analyse ved at den fokuserer på både trusler og muligheter. Samtidig går usikkerhetsanalyse mye lengre enn SWOT-analysen ved at den forsøker å kvantifisere både sannsynligheter og konsekvenser. Den vil dermed kunne supplere en SWOT-analyse med mer detaljert informasjon. I et prosjekts tidligfase vil usikkerhetsanalyse typisk kunne anvendes til å skape en bedre forståelse for hvilken usikkerhet ulike beslutningsalternativer kan medføre. Dette vil, på samme måte som de to foregående verktøyene, dermed bidra til at prosjektet får bedre innsikt i alternativene før man fatter en beslutning.

Gjennomføringsmessig følger trinnene naturligvis denne logikken, slik at man for hvert beslutningsalternativ som vurderes må (Husby, Klakegg, Kilde, Berntsen, Samset og Torp, 1999):

1. Identifisere usikkerhetselementer, både av positiv og negativ karakter.
2. For hvert usikkerhetselement, vurdere sannsynligheten for at elementet vil inntreffe.
3. For hvert usikkerhetselement, angi konsekvensen for prosjektet om det skulle inntreffe.
4. Oppsummere de identifiserte usikkerhetselementene i en usikkerhetsmatrise.
5. For de viktigste usikkerhetselementene, vurder om tiltak kan iverksettes som kan endre sannsynligheten for at de skal inntreffe eller konsekvensene av dem.

Med usikkerhetselementer menes alle typer hendelser eller utvikling man antar kan inntreffe og som vil påvirke prosjektet. Eksempler kan være diskrete hendelser eller forhold som regjeringsskifte, bortfall av en konkurrerende tilbyder eller funn av leirgrunn på utbyggingssted. Videre kan det være rammebetingelser som endres langs en kontinuerlig skala, for eksempel lavere rente, høyere priser i entreprenørmarkedet eller høyere etterspørsel etter den vare eller tjeneste produktets sluttleveranser skal frembringe.

For å angi sannsynligheten for at et usikkerhetselement skal inntreffe og konsekvensen for prosjektet om det skulle inntreffe, anvendes normalt en skala fra 1-5. Her betyr verdien 1 henholdsvis svært lav sannsynlighet og svært liten konsekvens, 5 at det ganske sikkert vil

inntreffe og at hele prosjektet kan stå i fare, mens verdiene mellom angir en kontinuerlig skala mellom disse ytterpunktene. Når man har tilordnet disse to vurderingene til hvert usikkerhetselement, plasseres elementene i en usikkerhetsmatrise. Avhengig av plasseringen i matrisen får hvert element en rangering som vist i Figur 11-4. Det er åpenbart de elementer som plasserer seg innenfor det røde feltet i øvre, høyre del av matrisen som bør tas med videre i vurderingen.

5: Svært høy sannsynlighet, nesten sikkert					
4: Høy sannsynlighet					
3: Middels sannsynlighet					
2: Lav sannsynlighet					
1: Svært lav sannsynlighet, nesten umulig					
<i>Sannsynlighet / Konsekvens</i>	1: Nesten ingen effekt	2: Liten effekt	3: Middels effekt	4: Stor effekt	5: Hele prosjektet i fare

*Middels viktige usikkerhetselementer*

*Kritiske usikkerhetselementer*

*Neglisjerbare usikkerhetselementer*

Figur 11-4 Prinsipiell usikkerhetsmatrise

Så langt i analysen vil det da foreligge en usikkerhetsmatrise for hvert av beslutningsalternativene. For å opparbeide en enda bedre innsikt i hvert av alternativene er siste trinn i analysen å vurdere om de kritiske usikkerhetselementene kan modifiseres ved å iverksette ulike tiltak. Om egnede tiltak enten kan redusere sannsynligheten eller konsekvensen av negative elementer eller øke dem for positive, vil dette endre bildet av beslutningsalternativene. Til slutt vil man sitte med et mer detaljert inntrykk av den usikkerhet hvert alternativ innebærer, og burde dermed være i en bedre posisjon til å vurdere alternativene opp mot hverandre.

Et bredbåndsselskap vurderer utbygging av en ny høyhastighetslinje for å øke tilbudet til en hel bydel i en større by. Alternativet er simpelthen nullalternativet, det vil si å ikke bygge en slik linje. For å skape innsikt i usikkerheten et slikt prosjekt vil være beheftet med har selskapet besluttet å gjennomføre en usikkerhetsanalyse.

Analysen gjennomføres med et lite antall personer fra den etablerte prosjektgruppen, assistert av en ekstern konsulent. Denne resulterer blant annet i de følgende usikkerhetselementene:

- En annen konkurrent bygger ut et alternativt tilbud.
- Båndbreddebehovet øker utover det linjen kan overføre.



- Det generelle prisnivået på bredbånd i Norge går betydelig ned.
- En større andel av husstandene ønsker å knytte seg til bredbånd.
- Entreprenørkostnader øker som følge av høy aktivitet i BA-bransjen.
- Det kan oppstå problemer med traséadgang for deler av linjen.

I tabellen under er sannsynlighet og konsekvens angitt slik gruppen vurderer dem:

Tabell 5 Usikkerhetslementer

#	Usikkerhetslement	Sannsynlighet	Konsekvens
1	En annen konkurrent bygger ut et alternativt tilbud	3	4
2	Båndbreddebehovet øker utover det linjen kan overføre	2	4
3	Det generelle prisnivået på bredbånd i Norge går betydelig ned	3	3
4	En større andel av husstandene ønsker å knytte seg til bredbånd	2	2
5	Entreprenørkostnader øker som følge av høy aktivitet i BA-bransjen	2	3
6	Det kan oppstå problemer med traséadgang for deler av linjen	1	5

Når dette oppsummeres i en usikkerhetsmatrise blir resultatet som vist i Figur 11-5.

5: Svært høy sannsynlighet, nesten sikkert					
4: Høy sannsynlighet					
3: Middels sannsynlighet			3	1	
2: Lav sannsynlighet		4	5	2	
1: Svært lav sannsynlighet, nesten umulig					5
Sannsynlighet / Konsekvens	1: Nesten ingen effekt	2: Liten effekt	3: Middels effekt	4: Stor effekt	5: Hele prosjektet i fare

Figur 11-5 Usikkerhetsmatrise

Fra denne matrisen finner prosjektgruppen at et slik prosjekt ikke vil stå overfor spesielt kritiske usikkerhetslementer. Dermed synes det fornuftig å gå videre med arbeidet.

Som for de foregående teknikkene er også usikkerhetsanalyse beheftet med den ulempe at den baserer seg på subjektive antakelser som i stor grad påvirker utfallet av analysen. Dette må man ha i bakhodet ved anvendelser av metoden og ikke ta resultatet for en absolutt sannhet. Ellers er en fordel med teknikken at de identifiserte usikkerhetsselementer kan anvendes for usikkerhetsstyring på et senere tidspunkt dersom prosjektet og ett av beslutningsalternativene besluttes gjennomført.

### **Teknikker egnet for å evaluere/sammenlikne ulike beslutningsalternativer**

Som nevnt for teknikkene for å vurdere enkeltalternativer gir hver slik analyse kun innsikt i hvert av alternativene, men fører ikke til noen rangering eller sammenlikning av alternativene. Dette behovet vil raskt melde seg i en tidligfase-vurdering, og i denne seksjonen presenteres verktøy for å kunne foreta slike sammenlikninger. Tre slike teknikker er:

- Normalisering, kjent fra analyser i benchmarking, der hensikten er å sikre at man sammenlikner størrelser som reelt er sammenliknbare.
- Radardiagram, et diagram for å sammenlikne helhetlige profiler for ulike alternativer, i ett og samme bilde, ved bruk av profilkurver langs ulike vurderingsdimensjoner.
- Sensitivitetsanalyse, tradisjonell kvantitativ vurdering av ulike alternativers følsomhet for endringer i rammebetingelser/parametre i forhold til forventet effekt.

### ***Normalisering***

Når man sammenlikner ulike beslutningsalternativer vil normalt en betydelig andel av de faktorene man sammenlikner være av en kvantitativ natur. Slike faktorer kan for eksempel være ulike kostnader, etterspørselsdata, trafikkestimater, valutaforhold, osv. En utfordring man ofte vil stå overfor ved slik sammenlikning er at de kvantitative størrelsene man ser på ikke er sammenliknbare. Ser man på totale utbyggingskostnader for ulike alternative samferdselsløsninger, kan direkte sammenlikning være unøyaktig fordi hver løsning vil håndtere et ulikt antall reisende, og man heller burde analysert kostnader per reisende. Slike unøyaktigheter ser man for mange ulike typer sammenliknende analyser, og normalisering innebærer å sikre at de størrelser som sammenliknes er relevante (spøkefullt sies det at om man ikke klarer å få til sammenlikning av epler med epler, hvertfall epler med eple-ekvivalenter). I så måte er ikke normalisering en konkret teknikk, men mer et prinsipp om å regne om størrelser for å justere for forhold som gjør sammenlikning unøyaktig (APQC, 1993).

Noen faktorer som typisk bidrar til at beslutningsalternativer ikke direkte kan sammenliknes er:

- Ulikt løsningsomfang, det vil si hvilke oppgaver som håndteres av det enkelte beslutningsalternativ eller graden av vertikal og horisontal integrasjon.
- Ulik størrelse, det vil si antall ansatte, omsetning, enheter, osv. som vil være involvert.
- Ulike markedsforutsetninger, med hensyn til geografisk plassering, størrelse, adferd, forventninger, osv.
- Ulikt kostnadsnivå, på grunn av for eksempel eiendomspriser, statsstøtte og tilskudd, lønnsnivå, osv.

Anvendelsesmessig er det vanskelig å sette opp et antall trinn når det gjelder normalisering. Rent praksis vil normalisering simpelthen innebære at det for hvert beslutningsalternativ som vurderes foretas en omregning til å sikre at alternativene er reelt sammenliknbare. Hvilke omregninger som gjøres vil variere fra det enkelte prosjekt til prosjekt. Dette illustreres kanskje best med et enkelt eksempel.

Det vurderes å etablere et nytt stadium i en større norsk ny. Flere alternative lokalisering og utforminger diskuteres, blant annet om anlegget skal være utendørs eller under tak, dekke bare fotball eller også friidrett, skøyter, osv., ha rom for butikker og kontorer, legge til rette for leiligheter, osv. For å sammenlikne alternativene har en foreløpig prosjektgruppe sammenliknet utbyggingskostnadene for hvert av de fem konkrete alternative utformingene som foreligger:

- Alternativ A: 288 millioner kroner
- Alternativ B: 156 millioner kroner
- Alternativ C: 132 millioner kroner
- Alternativ D: 102 millioner kroner
- Alternativ E: 88 millioner kroner

Naturlig nok vil utbyggingskostnadene for de ulike konseptene variere sterkt avhengig av hvor omfattende det gjøres, slik disse tallene viser. Skulle man fatte en beslutning utelukkende basert på disse tallene og et ønske om lavest mulig kostnader, ville naturligvis alternativ E bli foretrukket. Gruppen er heldigvis fullstendig klar over at nytten for byens innbyggere, idrettslag, næringsliv og andre interessenter vil variere sterkt med omfanget av anlegget. Det besluttes derfor å normalisere alternativene ved å beregne kostnader per forventet bruker som vil dra nytte av det nye anlegget:

- Alternativ A: 6.600 kroner per bruker
- Alternativ B: 5.460 kroner per bruker
- Alternativ C: 8.880 kroner per bruker
- Alternativ D: 8.720 kroner per bruker
- Alternativ E: 6.180 kroner per bruker

Disse tallene viser at alternativ B, som er et av de mest omfattende alternativene for stadiumutforming, vil minimere kostnadene per bruker og dermed ut fra dette alternativet utgjør det mest attraktive alternativet.

Faren ved en slik normalisering er selvsagt at man fester stor lit til de omregnede tallene, nettopp fordi man har foretatt en normalisering, mens omregningen egentlig ikke er relevant. Kanskje er faktoren kroner per bruker i eksempelet over slett ikke noe godt kriterium for å vurdere de foreliggende alternativene, men tallene får troverdighet gjennom at de er normalisert. Derfor bør slike omregninger vurderes nøye før normalisering tas i bruk.

## Radardiagram

Som vist i eksempelet rundt stadiumutbygging er ofte mye av de data som foreligger for hvert beslutningsalternativ av kvantitativ karakter. Sammenlikning av disse i lister/tabeller/matriser er selvsagt en relevant og nyttig tilnærming. Samtidig er det et velkjent faktum at grafiske fremstillinger av data ofte gjør det lettere å tolke data eller se mønstre. Radardiagram er et enkelt verktøy for å sammenlikne helhetlige dataprofiler, altså mange faktorer/kriterier samtidig, for ulike beslutningsalternativer i ett og samme bilde, ved bruk av profilkurver. I et prosjekts tidlige fase vil dette være nyttig nettopp i den typen situasjoner det foregående eksempelet på en svært forenklet måte representerer.

For å konstruere og anvende et radardiagram følges de følgende stegene (under forutsetning av at de data som skal fremstilles i diagrammet allerede foreligger) (Andersen, 1998):

1. Bestem hvor mange variable/faktorer som skal inkluderes i diagrammet og tilordne én per eike (akse) i diagrammet.
2. Avhengig av den enkelte variabel, del aksene inn i logiske segmenter ved å la hver akse benytte den måleenhet som passer for variabelen. Husk også på at aksene skal retningsgis slik at jo lenger fra sentrum av diagrammet et datapunkt er, jo bedre/gunstigere er punktet. Dette betyr at "negative" faktorer som kostnader, tidsbruk, osv. får høyeste verdier nærmest sentrum og "positive" faktorer som nytte, anvendelse, osv. får høyeste verdier lengst fra sentrum.
3. Plott inn datapunktene i diagrammet langs de riktige aksene, gjerne ved å bruke ulike symboler eller farger for å skille data for de ulike beslutningsalternativene fra hverandre.
4. Trekke linjer mellom datapunktene for hvert beslutningsalternativ, slik at det dannes profiler for hvert alternativ.
5. Studer diagrammet for å foreta en helhetlig vurdering av forskjellene mellom alternativene langs de ulike kriteriene.

I stedet for å beskrive disse trinnene teoretisk er det mer illustrerende å gjennomgå dem ved bruk av et enkelt eksempel, som bygger videre på forrige case rundt bygging av et stadium.

Prosjektgruppen som vurderer alternativer for utbygging av nytt stadium innser at kriteriet kostnader per bruker er for snevert som beslutningsgrunnlag. Gruppen utvikler derfor flere andre faktorer:

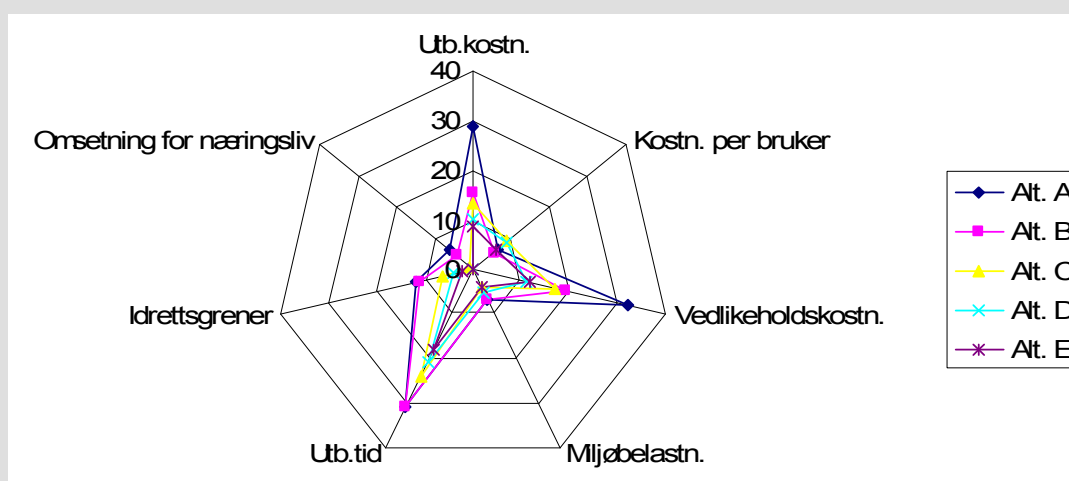
- Årlige vedlikeholdskostnader.
- En vurdering av miljømessig belastning for hvert alternativ.
- Utbyggingstid.
- Antall idrettsgrener som kan anvende anlegget.
- Estimert omsetning for næringsliv med lokalisering i anlegget.

De kvantitative data for alle kriteriene er vist i Tabell 6..

Tabell 6 Underlagsdata for radardiagram

	Alt. A	Alt. B	Alt. C	Alt. D	Alt. E
Utbyggingskostnad ( 10 millioner kroner)	28,8	15,6	13,2	10,2	8,8
Kostnad per bruker (1000 kroner per bruker)	6,6	5,46	8,88	8,72	6,18
Årlige vedlikeholds-kostnader (100 tusen kroner)	32	19	17	11	12
Miljømessig belastning (komposittfaktor langs skala 1-10)	7	7	4	5	4
Utbyggingstid (måneder)	31	31	24	21	18
Antall idrettsgrener som kan anvende anlegget (antall)	12	11	6	4	2
Estimert årlig omsetning for næringsliv med lokalisering i anlegget (10 millioner kroner)	6,3	4,2	0,8	0	0

Man ser snart at dette, ikke minst om flere kriterier skulle legges til, er uoversiktlig. For å lette tolkningen av dataene beslutter gruppen å sette opp et radardiagram. Hvert kriterium tilordnes til hver sin skala, retningen på skalaene snus i riktig retning og diagrammet settes opp som vist i Figur 11-6.



Figur 11-6 Radardiagram for stadiumutbygging

Grunnet størrelsen på diagrammet kan det være vanskelig å se alle profilene like godt, men når gruppen viser dette på en stor skjerm se de at alternativene A og B synes å være totalt sett de som representerer flest styrker.

## Sensitivitetsanalyse

Når man sammenlikner ulike alternativer, som for eksempel i radardiagrammet over, antar man ofte at ulike størrelser er kjente og faste, og dermed kan sammenliknes direkte. Dessverre er dette ofte en altfor lettvent forenkling. Selv ved små endringer i forutsetninger kan forhold som kostnader, varigheter, osv. lett endres betydelig. Sensitivitetsanalyse, gjerne også kalt følsomhetsanalyse, er en analyseteknikk som anvendes for å vurdere hvor følsomme ulike estimater er for slike endringer i forutsetninger de bygger på. I så måte er metoden nokså nært knyttet til usikkerhetsanalyse, som jo nettopp søker å fange opp forhold i prosjektet en frykter

kan endres og dermed påvirke det, for eksempel økning i kostnadsnivå, endret etterspørsel, osv.

Som en teknikk for anvendelse i tidligfasen av prosjekter kan sensitivitetsanalyse derfor gjerne kombineres med/bygge på en usikkerhetsanalyse. Hensikten blir da å supplere sammenlikningen av antatte, faste størrelser for ulike beslutningsalternativer med en analyse av følsomheten til disse størrelsene. Ved store forskjeller i følsomhet kan man få frem et bilde der det tilsynelatende mest attraktive alternativet er beheftet med så stor usikkerhet i forhold til estimatenes nøyaktighet og påvirkning fra underliggende faktorer, at et alternativ vurdert som mindre attraktivt likevel bør velges ut fra lavere følsomhet. En sensitivitetsanalyse kan for øvrig gjennomføres på ulike måter, først og fremst enten ved å variere én og én parameter av gangen eller ved å fremstille endringer av mange parametre i samme analyse (Rolstadås, 2001). Førstnevnte er selvsagt den enkleste varianten, mens sistnevnte gir et rikere bilde. Vi vil derfor her ta utgangspunkt i en multivariabel analyse.

Trinnene for å gjennomføre en slik sensitivitetsanalyse er:

1. Bestem hvilken vurderingsfaktor ved prosjektet det ønskes vurdert følsomheten til. Dersom flere faktorer ønskes analysert må dette gjøres ved å gjenta analysen.
2. Bestem hvilke underliggende parametre som skal inkluderes i analysen.
3. Beregn endringer i vurderingsfaktoren som følge av ulike endringer i de underliggende parametre.
4. Fremstill dette i et følsomhetsdiagram.
5. Etter å ha gjennomført samme analyse for både flere vurderingsfaktorer og flere beslutningsalternativer, vurderes følsomheten til de ulike alternativene opp mot hverandre.

Som for radardiagrammet i forrige seksjon er det enklere å forklare hvordan analysen gjennomføres ved å bruke et eksempel, og vi bygger derfor videre på stadiumeksempelet for å vise en forenklet sensitivitetsanalyse.

Prosjektgruppen vil se nærmere på utformingsalternativene A og B og spesielt hvor følsomme utbyggingstiden er for ulike endringer i rammebetingelser. I utgangspunktet er denne vurdert til å være 31 måneder for begge alternativene, og dette er svært nær den lengste byggeperioden en kan akseptere i forhold til at det skal arrangeres et større mesterskap i anlegget kort tid etter. Dersom ett av alternativene viser stor følsomhet i retning av lengre byggetid, vil dette telle i klar negativ retning.

Gruppen går dermed videre med å identifisere et antall underliggende parametre en antar kan påvirke byggetiden:

- Kapasitetsutnyttelse totalt sett i byggebransjen i regionen.
- Leveringstid for sentrale byggeelementer med lang produksjonstid.
- Antall måneder med arbeid som kun kan foregå sommerstid.

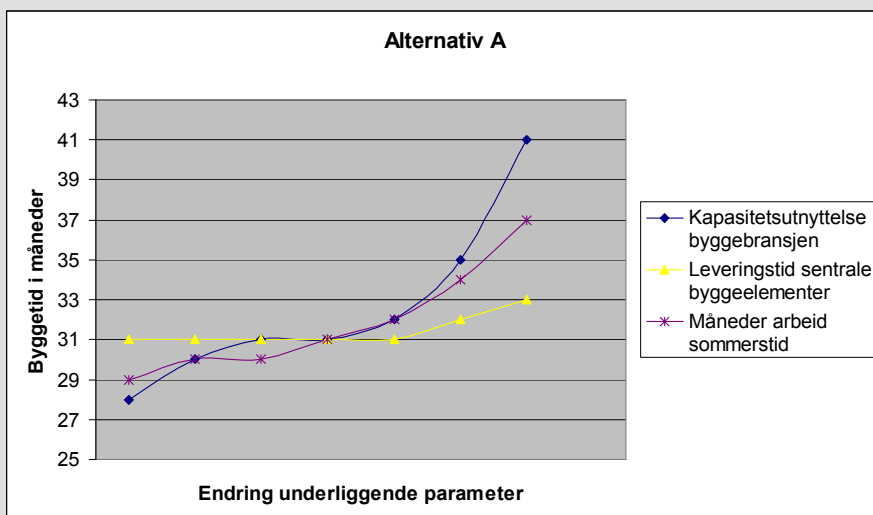
For hver av disse foretas en analyse av effekten på byggetid. Dette gjøres ved å variere parameteren i begge retninger fra antatt forventningsverdi i prosentvis sprang og beregne hvordan dette øker eller reduserer tiden. For eksempel finner man da for

kapasitetsutnyttelsen i bransjen at endringer bort fra de antatte 65% vil føre til følgende effekter:

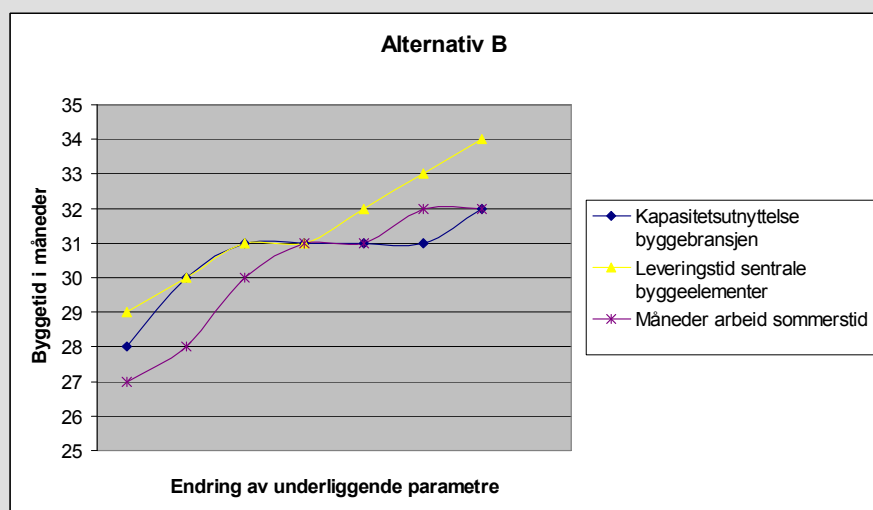
**Tabell 7 Beregninger for følsomhetsanalyse**

Kapasitetsutnyttelse byggebransjen	Byggetid
Basisestimat 65%	31
Endring +10%	32
Endring +20%	35
Endring +30%	41
Endring -10%	31
Endring -20%	30
Endring -30%	28

Tilsvarende beregninger gjøres for de to andre parametrene og det samme gjøres for alternativ B. Totalt sett gir dette to sett med data som fremstilles i følsomhetsdiagrammene i Figur 11-7 og Figur 11-8.



**Figur 11-7 Følsomhetsdiagram for alternativ A**



**Figur 11-8 Følsomhetsdiagram for alternativ B**

Ved å studere de to diagrammene, og legge merke til at skalaen på y-aksen er forskjellig, innser gruppen at alternativ A er langt mer følsomt for økning i byggetid, og foretrekker dermed alternativ B.

Som for alle slike kvantitative teknikker gjelder det selvsagt også for sensitivitetsanalyse at kvaliteten på analysen ikke blir bedre enn de tallene den baseres på. I dette tilfellet foretok prosjektgruppen mer eller mindre kvalifiserte vurderingen av hvordan ulike endringer ville påvirke byggetiden. Er disse nøyaktige er analysen nokså pålitelig, mens konklusjonen selvsagt blir feil om dette er tilfeldige antakelser. Kvalitetssikring av de underliggende data er dermed en viktig forutsetning for denne typen analyser.

### 11.3 Teknikker egnet for å vurdere krefter for og imot gjennomføring av ulike alternativer

Når man sammenlikner ulike beslutningsalternativer er selvsagt de ”rasjonelle” forholdene ved hvert alternativ av overordnet betydning. Samtidig finnes det veldig mange eksempler på at et nøytralt sett beste alternativ likevel kommer uheldig ut, ved at aksepten for alternativet blant ansatte, brukere eller andre interessenter ikke er til stede. Om en virksomhet kan velge to ulike teknologier for en investering, der det rimeligste og kanskje på andre måter beste medfører et redusert antall arbeidsplasser og det andre ikke gjør det, er det enkelt å forestille seg at de ansatte og deres fagforeninger vil velge det siste. Og tilsvarende at implementering av det første kan møte problemer, dersom man likevel velger dette. I så måte er det rådende klimaet for gjennomføring av de ulike alternativer, med krefter for og imot, viktig informasjon for å rangere alternativene. Vi vil her behandle en slik teknikk, kalt kraftfeltanalyse.

#### Kraftfeltanalyse

Kraftfeltanalyse er basert på antakelsen om at enhver situasjon er et resultat av krefter for og i mot nåtilstanden og som er i balanse med hverandre styrkemessig. En økning eller reduksjon i styrken på noen av kreftene vil skape en forandring, noe man i endringsledelse bruker til å skape forandringer i ønsket retning. I et prosjekts tidligfase vil teknikken primært kunne brukes til å simpelthen skape en oversikt over disse kreftene, for derigjennom å gi et bilde av hvilke utfordringer ulike beslutningsalternativer eventuelt vil måtte forventes å støte på (men selvsagt kan denne innsikten også være nyttig ved senere gjennomføring av prosjektet). Fremgangsmåten ved kraftfeltanalyse er (Andersen og Fagerhaug, 2001):

1. For hvert beslutningsalternativ, brainstorm alle mulige krefter, i prosjektets omgivelser, innad i prosjektets organisasjon og hos interessenter, som kan tenkes å virke for eller imot alternativet og dets implementering gjennom prosjektet.
2. Vurder styrken på hver av kreftene og tegn dem inn i et kraftfeltdiagram. Lengden på hver pil i diagrammet angir styrken på kraften den representerer.
3. For hver kraft, men spesielt for de sterkere kreftene, vurderes om det kan settes inn tiltak som kan øke de positive kreftene som virker for og redusere de negative som virker imot.

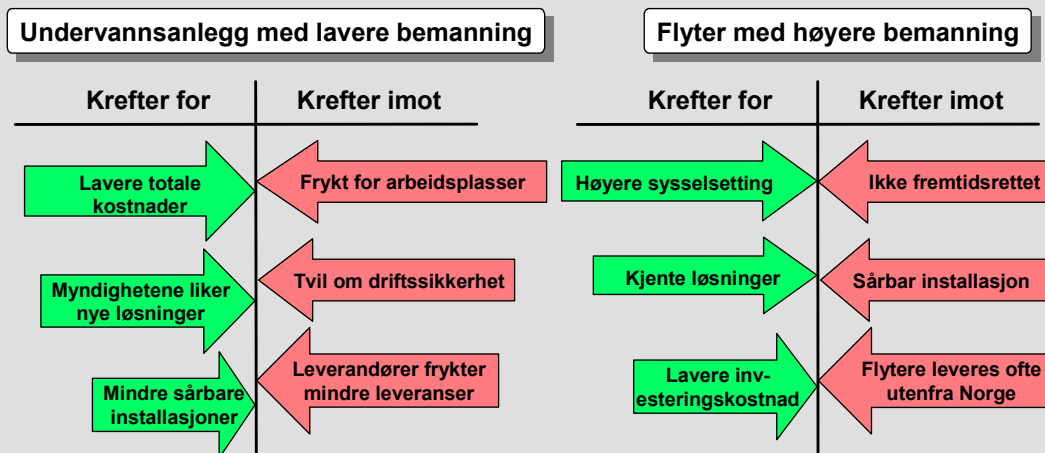


Den innsikt dette skaper vil så være ytterligere en supplering til analyser gjennomført ut fra andre synsvinkler.

Et oljeselskap skal bygge ut et nytt oljefelt i Nordsjøen og har etter hvert snevret de alternative utbyggingskonseptene ned til to:

- Bruk av undervannsanlegg, som koples til en nærliggende plattform for prosessering, og der undervannsanlegget styres fra en landbasert kontrollsentral. Denne modellen representerer noen teknologiske utfordringer og vil ha en høyere utbyggingskostnad, men vil over feltets livssyklus innebære en betydelig lavere total kostnad.
- Bruk av en flytende plattform for boring og prosessering, med boligkvarter for mannskap. Investeringen vil være noe lavere enn for alternativet med undervannsanlegg, men vil gi betydelig høyere driftskostnader over feltets livslengde.

Ut fra kostnads- og følsomhetsanalyser er selskapet innstilt på å velge undervannsanlegget, men ser stadige tegn på at ansatte, andre aktører i bransjen og politikere er skeptiske til den typen fjerndrift med sterkt redusert bemanning denne løsningen representerer. Man ønsker derfor å gjennomføre en kraftfeltanalyse for å vurdere styrken på spesielt de eventuelle motkrefter man vil stå overfor om dette velges. Figur 11-9 viser kraftfeltdiagrammer for begge alternativene.



Figur 11-9 Kraftfeltdiagram for begge løsningene

Ut fra dette bildet ser gruppen at undervannsanlegget nok vil møte større motstand enn en flytende plattform. Dette tolker man dit hen at man enten bør vurdere om det likevel bør velges det siste, eller at man ved valg av undervannsanlegg bør fokuseres betydelig på å motvirke de sterke motkreftene gjennom informasjon, dialog og medvirkning.

Kraftfeltanalyse er ikke, selv om den kanskje gir inntrykk av å basere seg på de vitenskapelige prinsipper i statikk, ikke en svært analytisk teknikk. De krefter som identifiseres har bakgrunn i gruppens antakelser og kunnskap, og styrken på dem og måten den symboliseres er svært grove vurderinger. Resultatene fra analysen bør dermed, som for mange andre teknikker, være ett av mange innspill til beslutning.

## 11.4 Teknikker egnet for å analysere alternativens effekt på ønskede mål/effekter

De verktøy som har vært presentert så langt er typisk karakterisert ved at de belyser ulike beslutningsalternativer fra forskjellige synsvinkler, for eksempel interesser eller usikkerhet. Etter å ha skapt en bred innsikt i de ulike alternativene og deres styrker og svakheter, vil det etter hvert normalt være et behov for å oppsummere analysene i en mer helhetlig vurdering. Fordeler og ulemper ved alternativene må veies opp mot hverandre slik at man forhåpentligvis kan identifisere det alternativet som totalt sett peker seg ut som det gunstigste. Det er da selvsagt viktig å holde vurderingene av hvert alternativ opp mot de mål man har med prosjektet og de effekter man ønsker det skal gi, på kort og lang sikt.

Til dette finnes det flere ulike analysemetoder å velge mellom, med varierende grad av kompleksitet, spesielt i forhold til matematiske tilnærminger. Her har vi valgt å ta med én metode, som ofte refereres til som flermålsanalyse, men som også har flere betegnelser.

### Flermålsanalyse/kriterietesting

Flermålsanalyse er ikke en entydig teknikk, men snarere en fellesbetegnelse på ulike tilnærminger der hovedhensikten er å vurdere ulike beslutningsalternativers påvirkning på et sett av flere målsetninger. En slik vurdering kan gjøres på flere ulike måter, noe det finnes utallige eksempler på i ulike utredninger og rapporter fra prosjekter. Mange bruker ulike grafiske fremstillinger med piler av ulikt utseende/tykkelse for å få frem en helhet i et slikt bilde, mens andre varianter baserer seg på bruk av vektninger av kriterier og kvantitative analyser. Vi vil her presentere en form for slik vektet, kvantitativ analyse som ofte omtales som kriterietesting.

Under vurderinger i et prosjekts tidlige fase vil dette verktøyet altså typisk komme til anvendelse på et punkt der de ulike beslutningsalternativer har vært analysert ved hjelp av andre, egnede teknikker. Nå ønsker man å se helheten i hvordan disse slår ut på prosjektets målsetninger.

Proseduren for å gjennomføre en flermålsanalyse ved bruk av kriterietesting er som følger:

1. Analysen gjøres gjennom bruk av en matrise. Prosjektets målsetninger plasseres i hver sin kolonne i matrisen og tilordnes, om dette er ønskelig, ulik tyngde videre i analysen gjennom å sette vekttall som uttrykker deres relative viktighet i forhold til hverandre. Ofte brukes verdier fra 1 til 10, der lavest verdi er mindre viktig, høy verdi kritisk, men en annen skal naturligvis benyttes. Beslutningsalternativene plasseres så i hver sin rad i matrisen.
2. For hvert beslutningsalternativ gjøres en vurdering av graden av bidrag det har til oppfyllelse av hvert enkelt av målsetningene. Igjen kan det brukes ulike skalaer for disse vurderingene, alt fra kontinuerlige skalaer fra 1-3 eller 1-10 til diskontinuerlige der man for eksempel bare anvender verdiene 1, 3 og 9. Uansett betyr en lav verdi svakt bidrag til målsetningen og høy verdi et sterkt bidrag.
3. Faktoren for påvirkningsgrad multipliseres med målsetningens vekttall, og produktet plasseres i de respektive matrisecellene i krysningspunktet mellom alternativ og målsetning.

4. For hvert beslutningsalternativ summeres produktene horisontalt i matrisen og totalsummen plasseres i matrisens høyre kolonne. Denne verdien angir nå det enkelte beslutningsalternativs bidrag til det samlede settet av målsetninger for prosjektet, korrigert for viktigheten av disse gjennom vektallene. Jo høyere score, desto større grunn til å anta at et alternativ vil bidra til å oppfylle målene.

Som sådan er selve analyseteknikken enkel og består kun av å sette vurderingsfaktorer og summere et antall multiplikasjoner. I sum er resultatet av analysen mer slagkraftig enn som så, nettopp fordi det gjøres et relativt stort antall enkeltvurderinger som minsker risikoen for feilvurderinger, og fordi bruken av vektall sikrer at kritiske målsetninger får høyere gjennomslagskraft i analysen. Som vanlig illustrerer vi bruken med et eksempel.

Det skal bygges en ny jernbanestrekning i Oslo-området, og man har valget mellom hele fire reelle alternative trasévalg:

- Vestre alternativ.
- Midtre alternativ.
- Østre alternativ.
- Tunnelalternativ.

Gruppen som utreder dette består av representanter for utbygger, trafikkoperatører og kommunen, og den har allerede gjennomført flere ulike analyser av alternativene. Disse har gitt en del innsikt i situasjonen, men man ønsker nå å konkludere valget basert på en flermålsanalyse.

Målsetningene/effektene med prosjektet har lenge vært definert og fremkommer i matrisen der analysen er gjennomført, Figur 11-10. Disse tilordnes vektall fra 1 til 10. Alternativenes effekt på disse er vurdert ved bruk av faktorene 1, 3 og 9, og det hele kulminerer i summene til høyre.

Mål	Minimere reisetid	Redusere sporkrysninger	Øke passasjertallet	Eliminere krysning av vei	Akseptable utbyggingskostnader	Unngå ødeleggelse av grøntområde	Sum
Vekting	8	4	7	6	6	5	
Vestre	8	36	21	54	18	5	142
Midtre	24	12	21	6	54	15	132
Østre	24	4	21	18	18	45	130
Tunnel	72	36	21	18	6	15	168

Figur 11-10 Flermålsanalyse for trasévalg

Gruppen er litt skuffet over at differansene mellom alternativene ikke ble større, da dette ville gitt en klarere indikasjon for valget. Like fullt er man enige om tunnelalternativet totalt sett kommer såpass mye bedre ut enn de andre at man, selv med de høyeste kostnadene, velger å gå videre med dette.

Det har vært gjentatt noen ganger, kanskje til det kjedsommelige, men også denne teknikken resulterer i et tallresultat. Erfaringsmessig vet vi at slike tall ofte oppfattes som absolutte størrelser, til tross for at de er et resultat av høyst subjektive vurderinger. De bør derfor tolkes med forsiktighet. Man skal også være klar over at valget av skala for både vektall og

påvirkningsfaktorer kan influere sterkt på resultatet av analysen. Ofte forsøker man å gjennomføre samme analyse med ulike skalaer, for å se hvordan dette påvirker analysen. Om dette ikke medfører for mye ekstraarbeid, kan det være en anbefalt fremgangsmåte.

## 11.5 Avslutning

De presenterte teknikkene burde utgjøre en verktøykiste av et passende omfang for prosjektgrupper som er i ferd med å foreta analyser i en tidligfase. Til sammen dekker de bredden i den typer analyser man normalt har behov for å gjennomføre, og de er stort sett alle relativt enkle teknikker som lett kan tas i bruk selv uten foregående erfaring med dem. Samtidig må vi tilføye at det finnes mange andre verktøy som kunne vært inkludert i dette utvalget. Føler man at ett eller flere analyseperspektiver ikke dekkes av disse, anbefaler vi at man ser seg om etter andre metoder i tillegg.

## 11.6 Referanser

Akao, Yoji (editor) (1990) *Quality Function Deployment: Integrating Customer Requirements Into Product Design*, Productivity Press, Cambridge, Massachusetts, USA.

Andersen, Bjørn (1998) *The Business Process Improvement Toolbox*, ASQ Quality Press, Milwaukee, Wisconsin, USA.

Andersen, Bjørn og Fagerhaug, Tom (2002) *Root Cause Analysis: Simplified Tools and Techniques*, ASQ Quality Press, Milwaukee, Wisconsin, USA.

Andersen, Bjørn og Fagerhaug, Tom (2002) *Performance Measurement Explained: Designing and Implementing Your State-of-the-Art System*, ASQ Quality Press, Milwaukee, Wisconsin, USA.

American Productivity & Quality Center (1993) **The Benchmarking Management Guide**, Productivity Press, Portland, Oregon, USA.

Husby, Otto, Klakegg, Ole Jonny, Kilde, Halvard S., Torp, Torp, Berntsen, Stein og Samset, Knut (1999) *Usikkerhet som gevinst*. Usikkerhetsstyring i prosjekt, NSP.

Lewis, James P. (1999) *The Project Manager's Desk Reference*, McGraw-Hill, New York, USA.

Project Management Institute (2000) *A Guide to the Project Management Body of Knowledge*, PMI, Newtown Square, Pennsylvania, USA.

Rolstadås, Asbjørn (2001) *Praktisk prosjektstyring (3. utgave)*, Tapir Forlag, Trondheim, Norge.

## 12 FUZZY LOGIKK I TIDLIGFASEVURDERINGER

KJELL AUSTENG\*

### 12.1 Innledning

Fuzzy logic er en fellesbetegnelse på flere beslutnings- eller styringsmodeller basert på teorien om diffuse mengde (fuzzy sets). Felles for alle disse modellene er at de er laget for å håndtere situasjoner i virkelighetens verden, og at de bevisst søker å utnytte den menneskelige vurderingsevne.

De fleste av metodene som er basert på subjektive vurderinger, benytter ulike former for betraktninger rundt sannsynlighetsmodeller (probabilistiske modeller). Fuzzy logic-metodene er ikke basert på sannsynlighet, men benytter såkalte **mulighetsmodeller (possibilistiske modeller)** for å beskrive virkeligheten. Dette blir nærmere beskrevet siden.

Felles for modeller basert på fuzzy logic er at de er tilpasset menneskelig måte å tenke på. De gir rom for omtrentlighet, usikkerhet i anslagene basert på subjektive vurderinger, og de kan brukes i etablering av beslutningsstøtte.

Fuzzy logic er å betrakte som et ekspertsystem, men til forskjell fra tradisjonell tenkning hvor et utsagn er vurdert som enten riktig eller galt, så er fuzzy logic i stand til å ta vare på utsagn som "kanskje riktig" eller "delvis galt".

”Fuzzy thinking”, også innen prosjektvirksomhet, innebærer en erkjennelse om at "everything is a matter of degree", dvs. at det aldri er noe som er enten helt svart eller helt hvitt, men at man hele tiden må forholde seg til «shades of grey».

Den mest utbredte anvendelsen av fuzzy logic har imidlertid frem til nå vært i forbindelse med forskjellige styrings- og reguleringsystemer av særlig elektrisk utstyr hvor det er spørsmål om å sammenholde informasjon fra flere kilder (sensorer) med divergerende krav til respons, og ut fra dette trekke riktige slutninger.

Hvis fuzzy logic er velegnet til styring av avanserte vaskemaskiner og ventilasjonsanlegg, kan kanskje den samme teorien anvendes for styring av prosjekter.

Styring av prosjekter består ofte av å registrere tingenes tilstand i form av «gråtonede» utsagn (godt, bra, lenge, snart, tilfredsstillende), og å treffe beslutninger om pådrag av ressurser som i sin tur kan ha ujevn kapasitet.

Til slutt i dette kapitlet vil vi presentere bruk av fuzzy logic brukt i nytteanalyse, og hvordan dette kan danne et fullverdig konsept for alternativvalg.

Den som har fått æren for å være opphavsmannen til utviklingen av teoriene om Fuzzy logic er professor Lotfi Zadeh ved University of California at Berkley. Han skrev i 1965 en artikkel

---

\* Kjell Austeng er 1. amanuensis ved Institutt for bygg- og anleggsteknikk, NTNU

kalt "Fuzzy Sets", og har siden hele tiden bidratt til fornyelse og videreutvikling av både det teoretiske grunnlaget og tilhørende praktiske applikasjoner.

I starten fikk tankene til Zadeh en noe blandet mottakelse, inkludert den skepsis som nye ideer vanligvis vekker.

I de siste 20 årene har japanerne ledet an i den praktiske utviklingen. De store japanske industriselskapene kappes om å utvikle mer eller mindre intelligente maskiner, og fuzzy logic har blitt en integrert del av kybernetisk forskning. Et av de mest kompliserte systemer basert på fuzzy logic er styrings- og sikkerhetssystemet til undergrunnsbanenettet i den japanske byen Sendai.

Innenfor den store sekken "prosjektledelse" vil, slik vi ser det, fuzzy logic ha sin mest naturlige anvendelse som hjelpemiddel til å bygge såkalte ekspertsystemer.

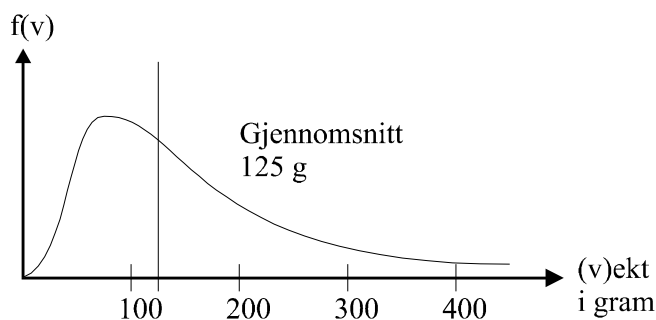
## 12.2 Sentrale begreper og generelle beskrivelser

Før vi kan forklare og eksemplifisere hva Fuzzy logic egentlig innebærer er det nødvendig å avklare noen sentrale begreper. I vedlegg A: Terminologi helt sist i kapitlet er det i tillegg en oversikt over betydningen av en del ord og uttrykk.

### Sannsynlighet, mulighet, grad av medlemskap i en mengde.

Det er naturlig å starte med å forklare sammenhengene og forskjellene mellom sannsynlighetsfordelinger og mulighetsfunksjoner.

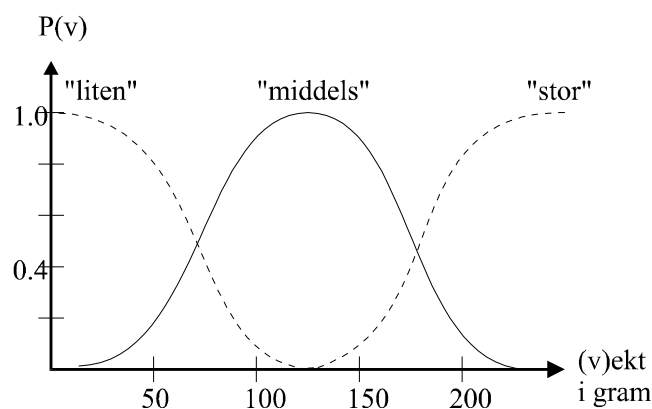
En mann skal sortere poteter. Fordelingen av potetene er slik at sannsynlig vekt for en tilfeldig valgt potet fra haugen har en tetthetsfunksjon som vist i figur 12-1.



Figur 12-1 Sannsynlighetsfordeling.

Sorteringen foregår på den måten at han vurderer vekten på potetene og først grupperer i klassene «under gjennomsnittet» og «over gjennomsnittet». Fra klassen «over gjennomsnittet» blir så sortert ut en gruppe kalt «stor» og fra «under gjennomsnittet» blir sortert ut en gruppe kalt «liten». Det som da blir igjen slås sammen i gruppen «middels».

Hvis vi nå prøver å beskrive *muligheten* for at en tilfeldig valgt potet vil havne i gruppen «liten», «middels» eller «stor» blir den kanskje noenlunde som i Figur 12-2.



Figur 12-2 Eksempel på mulighetsfunksjoner.

Denne «mulighetsfunksjonen» gir et bilde av i hvilken grad en potet med en viss vekt kan sies å være medlem i de forskjellige gruppene.

Denne fremstillingsmåten blir heretter kalt medlemskapsfunksjoner, hvor fullt medlemskap har verdien 1.0 og intet medlemskap har verdien 0.

Vi ser at konstruksjonen av slike medlemskapsfunksjoner er basert på vurderinger, noen ganger på rent subjektivt grunnlag. Vurderingene går på i hvilken grad et element kan sies å være medlem i en bestemt mengde. Vi kjenner her klart igjen tankegangen som ligger til grunn for bayesisk statistikk, hvor det etableres sannsynlighetsfordelinger ved hjelp av subjektive vurderinger og ”grad av tiltro .....”.

En variabelverdi kan i forskjellig grad være medlem i flere grupper, men fullt medlem kan den bare være i én gruppe.

### Diffus mengde.

Fuzzy logic er som tidligere nevnt sprunget ut fra ideen om diffuse mengder. En diffus mengde er en mengde som ikke har skarp definert grense. Den kan, i motsetning til faste eller klassiske mengder, inneholde elementer som delvis er medlem av mengden.

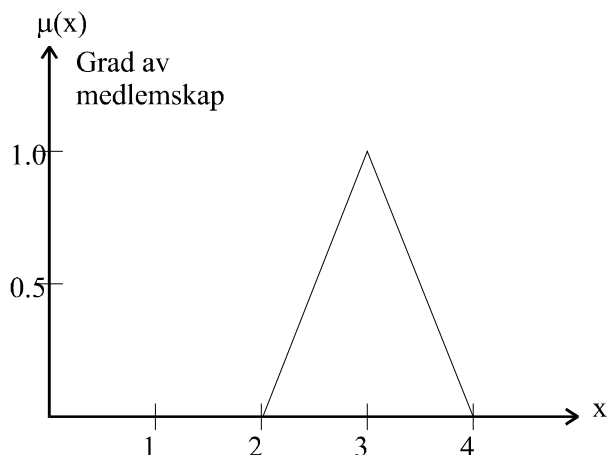
Graden av medlemskap kan variere kontinuerlig i områder 0 til 1.

I klassisk mengdelære er et element enten fullt medlem av mengden, eller ikke medlem i det hele tatt.

D.v.s. graden av medlemskap i mengden er enten 0 eller 1. Grad av medlemskap i diffuse mengder er gjerne vurderinger på subjektivt grunnlag, ofte situasjonsbetinget, og avhengig av øynene som ser.

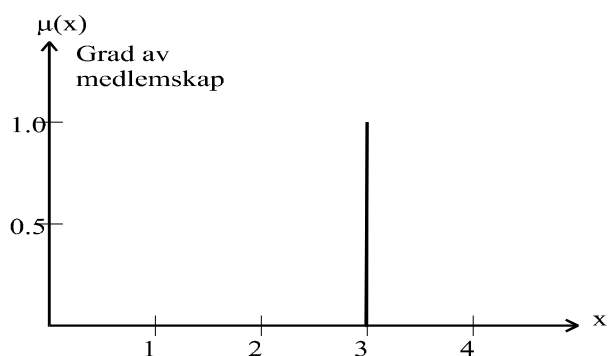
De diffuse mengdene vi skal se på her er diffuse tallstørrelser (cirkastørrelser), på engelsk snakker man om "fuzzy numbers", og mengder knyttet til såkalte lingvistiske variabler.

Eksempel på en diffus tallstørrelse er mengder av ca. 3 ( $\{\text{ca. } 3\}$ ), som i en gitt sammenheng kan tenkes å være definert slik at punktene på tallinjen fra 2 til 3 har grad av medlemskap som øker lineært fra 0 til 1, og at graden av medlemskap deretter minker lineært ned til 0 i punkt 4.



Figur 12-3 Eksempel på "fuzzy number".

Vi har nå dannet oss et bilde som viser sammenhengen mellom tallstørrelser og grad av medlemskap i den diffuse mengden  $A = \{\text{ca. } 3\}$ . Dette kalles en medlemskapsfunksjon og skrives  $\mu_A(\mathbf{x})$ . Funksjonen for den klassiske mengde  $\{3\}$  vil for sammenlikningens skyld se ut som i figur 12-4.



Figur 12-4 Medlemskapsfunksjon for en fast tallstørrelse.

### Lingvistisk variable.

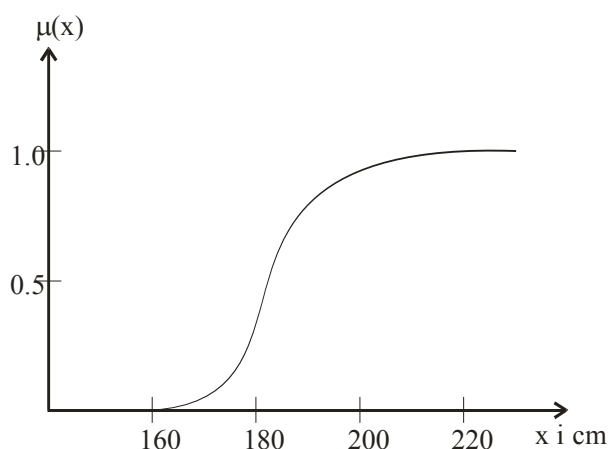
Fuzzy logic kan beskrives som er en måte å tenke på. Tanker uttrykkes i ord, og det er behov for en type mengder som på best mulig måte gjenspeiler variasjoner som ligger i verbale utsagn.



I årevis har byggerne av ekspertsystem prøvd å oversette naturlige språkbeskrivelser til program der regler fullstendig beskriver systemets oppførsel. Fuzzy logic starter med, og bygger på, et sett av brukerdefinerte språkregler. Variablene i slike språkregler kalles for *Lingvistisk variable*.

Et mye brukt eksempel på en mengde med lingvistisk variabel er **B = {høye menn}**. Hvem som blir ansett for å være høy er helt klart avhengig av hvor vi befinner oss på kloden, altså øyne som ser, men også avhengig av hva vi forøvrig sammenlikner med. En mann på 185 cm vil bli betraktet som ganske høy blant den jevne befolkning, mens 185 cm neppe virker særlig imponerende i basketballkretser.

Medlemskapsfunksjonen  $\mu_B(x)$  kan tenkes å se ut som i figuren nedenfor.



Figur 12-5 Medlemskapsfunksjon i mengden av høye menn.

Før vi går videre skal vi kan summere opp det som er omhandlet så langt:

- Diffuse mengder beskriver diffuse begreper (små poteter, høye menn, ca. 3).
- Bruk av diffuse mengder er en erkjennelse av muligheter for delvis medlemskap i mengder.
- En medlemskapsfunksjon i tilknytning til en diffus mengde gir inputverdien "grad av medlemskap" til de regelbaserte logiske operasjonene.
- Graden av medlemskap angis i størrelse mellom 0 og 1.
- Lingvistisk variabel: For eksempel størrelsen på poteter.
- Lingvistiske variabler er delt i størrelser. Disse størrelsene er gitt semantiske merkelapper (liten, middels, stor), og størrelsene kalles semantiske variabler. Til hver semantisk variabel knyttes en diffus mengde, og en medlemskapsfunksjon

### Modellering med fuzzy logic.

Fuzzy logic brukes til å modellere (kontinuerlige) sammenhenger mellom input og resultat.

Hvis vi har én input-kilde, vil resultatet i sin enkleste form fremstå som en kurve i et todimensjonalt aksesystem. I mer kompliserte modeller kan output-verdiene generere nye funksjonssammenhenger og gi opphav til nye resultatmodeller. Hvis vi har to input-kilder, vil vi tilsvarende få en modell i tre dimensjoner, og resultatet vises i sin enkleste form som en romkurve eller som en mer eller mindre krum skive i rommet. Også her kan output-verdiene tilordnes nye funksjoner og således generere rommodeller med innhold.

Med mer enn to input-kilder vil vi få multidimensjonale rom. Disse er, som kjent, vanskelig å illustrere, så selv om de aller fleste anvendelsene er basert på mer enn to input-kilder, vil vi i det etterfølgende begrense oss til det tredimensjonale rom.

Vi vil i det etterfølgende vise hvordan vi ved hjelp av fuzzy logic, uten å benytte noen form for matematisk kurvetilpasning, kan beskrive matematiske sammenhenger.

Fordelen med å bruke fuzzy logic fremfor matematisk tilpasning for å beskrive funksjoner øker kraftig når antall input-kilder øker.

### **Hvorfor Fuzzy logic?**

- Konseptet "Fuzzy logic" er rimelig lett å forstå.
- Det som gjør fuzzy tiltrekkende er at man benytter en tilnærming til problemstillingene som føles naturlig.
- Fuzzy logic er en svært fleksibel tilnæringsmåte.
- Fuzzy logic er utviklet for å behandle upresise data.
- Erkjennelsen om at "everything is a matter of degree" er bygd inn i prosessen, i stedet for det som er mer vanlig, det å ta hensyn til dette som en øvelse til slutt.
- Fuzzy logic brukes til å modellere ikkelineære sammenhenger.
- Det kan lages fuzzy systemer som beskriver hvilken som helst sammenheng mellom inputdata og resultat.
- Fuzzy logic er utviklet for å utnytte kunnskap og erfaring hos eksperter.
- Ideen med medlemskapsfunksjoner og logiske operasjoner er i realiteten et ekspertsystem som krever viten om temaet for analysen og de sammenhenger som gjelder.
- Fuzzy logic kan brukes sammen med eller som et supplement til andre analyseteknikker.
- Fuzzy logic er basert på bruk av vanlig, dagligdags språkbruk.
- Dette er kanskje det viktigste punktet, og det som gjør fuzzy logic til et svært kraftig verktøy. Vi skal siden se på bruk av såkalte lingvistiske variabler og behandling av disse ved bruk av adjektiver, adverb og semantiske nyanser.

## 12.3 Prosessene.

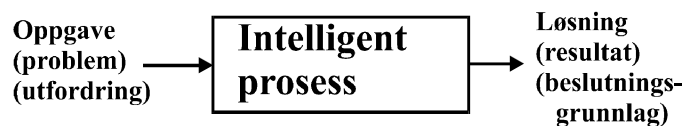
I dette delkapitlet skal vi beskrive hvordan et problem kan løses ved hjelp av fuzzy logic.

Vi skal beskrive hva vi ønsker, hvordan systemet er tenkt å virke, definisjon av input- og outputvariabler medlemskapsfunksjoner, systemoperatorene, utvikling av kontrollregler, simulering av modeller og justering av parameterne, samt integrering og implementering til en brukerapplikasjon.

Vi vil videre gå gjennom arbeidsgangen i en tenkt problemløsning, og samtidig beskrive delprosessene i det vi kaller for fuzzymaskinen.

### Beskriv systemet.

Generelt er det slik at når man står overfor en oppgave eller utfordring ønsker man ved en eller annen intelligent prosess å komme fram til et mulig resultat som kan brukes som beslutningsgrunnlag eller styringssignal.



Eksempel:

Oppgave: F.eks. valg av alternativ for veiutløsning ved hjelp av konsekvensanalyse.

Prosess:

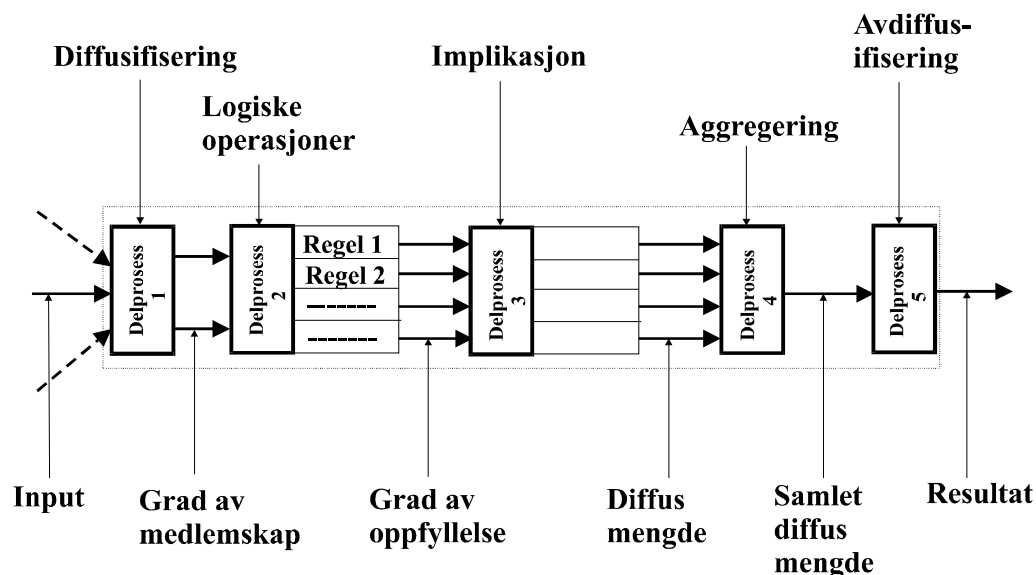
- a. Kostnadsoverslag
- b. Beregning av nytte (det som lar seg verdisette) og utregning av nytte/kostnad.
- c. Vurdering av usikkerhet i beregningene i a) og b).
- d. Vurdering av de verdiparameterne om ikke lar seg måle i kroner.
- e. Sammenstilling av resultatene for alle alternativene.

Løsning: Beslutning om vektlegging av de forhold som kan måles i kroner kontra de "myke" verdiene og rangering av alternativene.

Før vi tar for oss et eksempel skal vi beskrive gangen i en fuzzy logic-prosess.

## Fuzzymaskinen

I det etterfølgende skal vi beskrive fuzzymaskinen og ta for oss de delprosessene som er knyttet til denne.



Figur 12-6 Fuzzymaskinen.

Etter at systemet er beskrevet, og input til fuzzymaskinen dermed er gitt, blir neste skrittene i problemløsningen som beskrevet i det følgende.

### Spesifikasjon av intervallene for alle input til og output fra fuzzymaskinen.

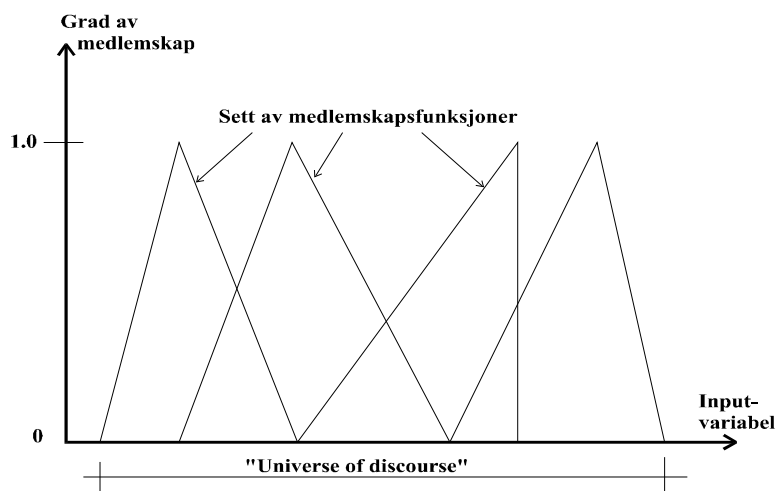
Det er svært sjelden at grensene for input og output er  $\langle -\infty, +\infty \rangle$ .

I alle praktiske sammenhenger må vi finne formålstjenlige grenser å arbeide innenfor. I all engelskspråklig litteratur om fuzzy logic snakker man om "universe of discourse". Vi vil snakke om "definert intervall".

Input består av eksakte størrelser innenfor definerte intervaller for hver parameter som påvirker resultatet. Ut fra delprosess 1 kommer en tallstørrelse mellom 0 og 1 som viser i hvilken grad input er medlem i diffuse mengder. Resultatet (output) er én eksakt størrelse for hvert sett av input.

### Etablere medlemskapsfunksjoner.

Først må det etableres et sett av medlemskapsfunksjoner for hver parameter som vurderes for input, og i tillegg må det etableres et sett av funksjoner som dekker mengden av konsekvenser.



Figur 12-7 Medlemskapsfunksjoner.

Hvor mange funksjoner, hvilken form de har, og hvilket område de dekker er et spørsmål om vurdering basert på kunnskap innenfor det område input hentes fra. Et krav er at funksjonene i settet til sammen skal dekke hele det definerte intervallet.

Heske and Heske [2] har satt opp følgende forslag til regler for plassering og overlapping av medlemskapsfunksjonene:

- I. Der det er mulig og ønskelig med mer presise angivelser av de lingvistiske variablene brukes flere og tettere funksjoner.
- II. Ethvert punkt i det definerte intervallet bør være medlem i minst én diffus mengde.
- III. De fleste punkter i det definerte intervallet bør være medlem i to diffuse mengder.
- IV. Funksjonene bør være slik at ikke to eller flere funksjoner definerer fullt medlemskap på samme sted. (Toppene i medlemskapsfunksjonene bør ikke falle sammen).
- V. Der hvor funksjonene overlapper hverandre bør summen av medlemskap ligge i nærheten av 1,0.
- VI. Bare få punkter i det definerte intervallet bør være medlem i så mye som tre diffuse mengder, og ingen punkter bør ha mer enn tre medlemskap.

For egen del vil vi føye til:

- VII. Settet av medlemskapsfunksjonen må være slik at det er mulig å fortelle det man ønsker ved hjelp av if-then-betingelsene. Særlig er det viktig (og vanskelig) å sette funksjonen slik at endepunktene i resultatpresentasjonen havner riktig.

### Formen på medlemskapsfunksjonene.

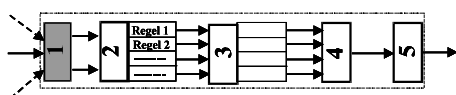
Det er et utall av mulige former på medlemskapsfunksjoner, og det kan være fristende å prøve å beskrive den virkelige situasjonen med høyest mulig presisjon. Vi tror imidlertid at hvis det rettes for mye oppmerksomhet mot en nøyaktig modellering av graden av medlemskap, vil

dette overskygge de vanskeligste elementene i analysen, som f.eks. input og logiske sammenhenger. Eksotiske medlemskapsfunksjoner er slett ikke nødvendig for å gi en god nok beskrivelse av virkeligheten. Rettlinjede funksjoner som trekanten og trapeser, og i noen tilfeller S-kurver og Gausskurver, hvor arbeidet legges i å definere utstrekningen, er godt nok.

Erfaring fra eksperimentering med et dataprogram er at forskjellen på lineære medlemskapsmodeller og forsøk på ikkelineære tilpasninger er liten, og langt overskygges av usikkerheten i input. Det som imidlertid gir synbare utslag er endringer i intervallengden for de forskjellige semantiske variabler<sup>39</sup>.

Jeg tror også det er viktig å ha for øyet at unødige presisjon koster. Dette fordi høy presisjon krever stor tilgang på nøyaktig informasjon, som noen ganger til og med ikke er tilgjengelig, og fordi høy presisjon gjerne gjør bruken av metoden mer komplisert og derved mindre brukervennlig.

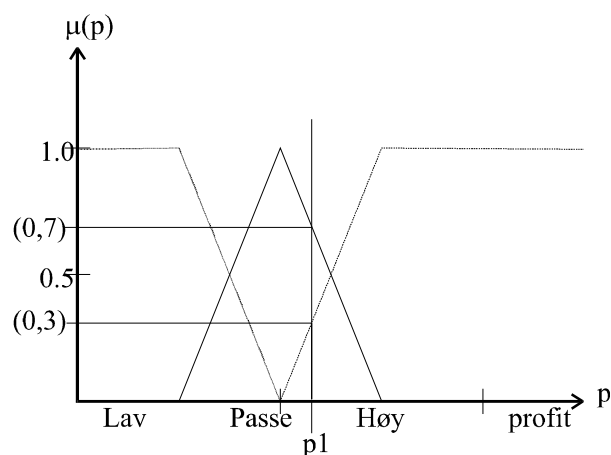
### Delprosess 1: Diffusifisering



For å vise eksempler på de forskjellige delprosessene, har vi valgt å se på hvorvidt et prosjekt er attraktivt for en investor som en avveining mot mulig profitt og mulig risiko.

Altså: Attraktivitet =  $f(\text{profitt}, \text{risiko})$ .

Først skal vi vise medlemskapsfunksjonene for parameterne hvor  $p_1$  og  $r_1$  er input til henholdsvis profitt og risiko:

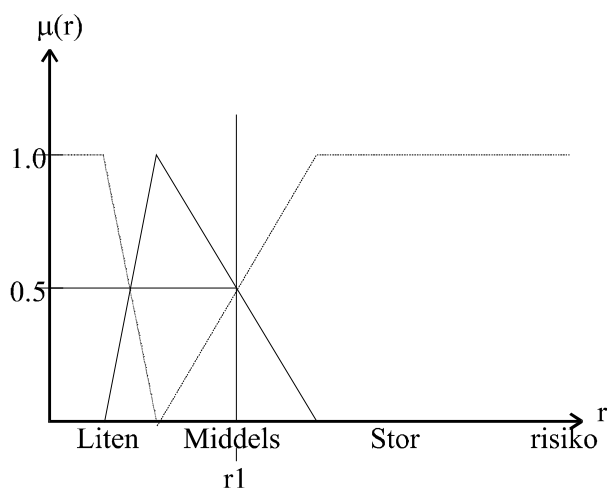


<sup>39</sup> (Semantisk variabel: En lingvistisk variabel består av flere beskrivelsesord gjerne i form av adjektiver, (liten, middels, stor, enorm). Et slikt beskrivelsesord med tilhørende medlemskapsfunksjon kalles en semantisk variabel.)

Figur 12-8 Medlemskapsfunksjoner for profitt og grader av medlemskap.

Fig. 12-8 viser at for en bestemt profitt kan vi definere medlemskap i flere mengder. Eksempelvis er  $p_1$  70 % "passe" og 30 % "høy".

Fra delprosess 1 til delprosess 2 blir overført to størrelser,  $\mu_{\text{Passe}}(p_1) = 0,7$  og  $\mu_{\text{Høy}}(p_1) = 0,3$ .



Figur 12-9 Medlemskapsfunksjoner for risiko og grader av medlemskap.

Vi leser av at  $r_1$  er 50 % "middels" og 50 % "stor".

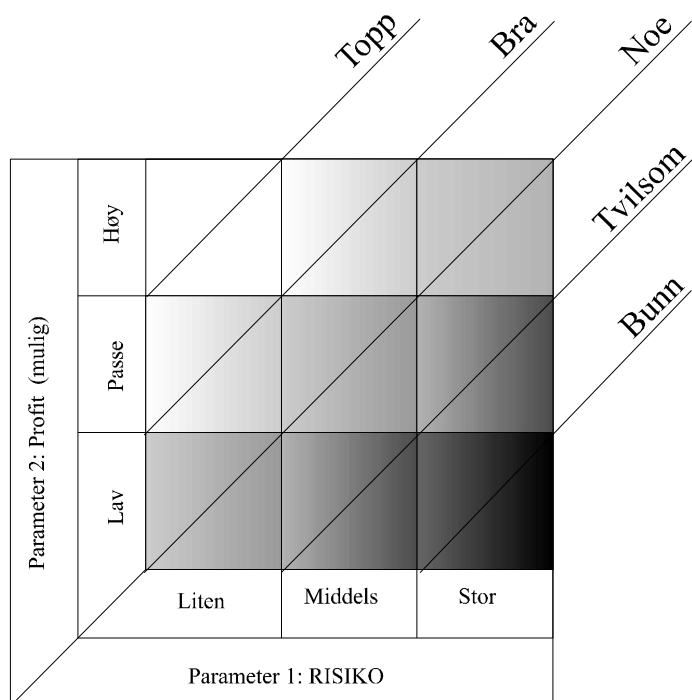
Fra delprosess 1 til delprosess 2 overføres  $\mu_{\text{Middels}}(r_1) = 0,5$  og  $\mu_{\text{Stor}}(r_1) = 0,5$ .

### Konstruere reglene for de logiske operasjonene.

I delprosess 2 blir resultatene fra diffusifiserings-prosessen kjørt gjennom en sil av betingelser i form av IF-THEN-regler, og for hver regel blir det beregnet i hvor stor grad den er oppfylt.

For å sette opp reglene slik at de dekker virkeligheten kreves kunnskaper om påvirkningsfaktorene (input), sammenhengen mellom dem og konsekvensene av variasjon; altså fagkunnskap innenfor problemområdet. I tillegg kreves kunnskap innen matematisk mengdelære og om hvordan logiske operatører (AND, OR, NOT) virker.

Et bra hjelpemiddel hvis input kommer fra to kilder, er en såkalt konsekvensmatrise. Se fig. 12-10. Denne viser sammenhengen mellom påvirkningsparameterne og en konsekvens.



Figur 12-10 Konsekvensmatrise.

Eksemplet i figuren viser sammenhengen mellom størrelsen på den profitt som kan hentes ut av et prosjekt og den risiko man tar, og hvor attraktivt prosjektet er for investeringer. Attraktiviteten er i dette tilfellet gradert i "topp", "bra", "noe", "tvilksom" og "bunn". Vi vil siden komme tilbake til medlemskapsfunksjoner for disse semantiske variablene.

Vi vil se at det er nødvendig med en vekselvirkning mellom utviklingen av konsekvensfunksjonene, matrisen og operasjonsreglene.

### Delprosess 2: Logiske operasjoner.

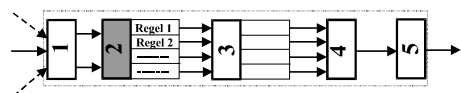
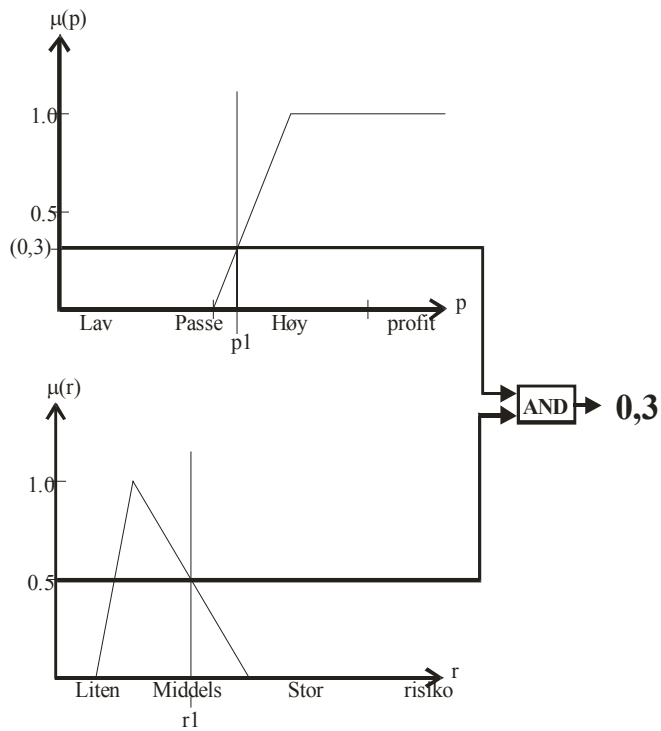


Fig. 12-11 og fig. 12-12 viser eksempler på hvordan AND- og OR-operatorene virker.

I begge figurene er det forutsatt at profitt har verdien  $p_1$ , som gir 30 % medlemskap i mengden "Høy". Risiko har verdien  $r_1$  som gir 50 % medlemskap i mengden "Middels".

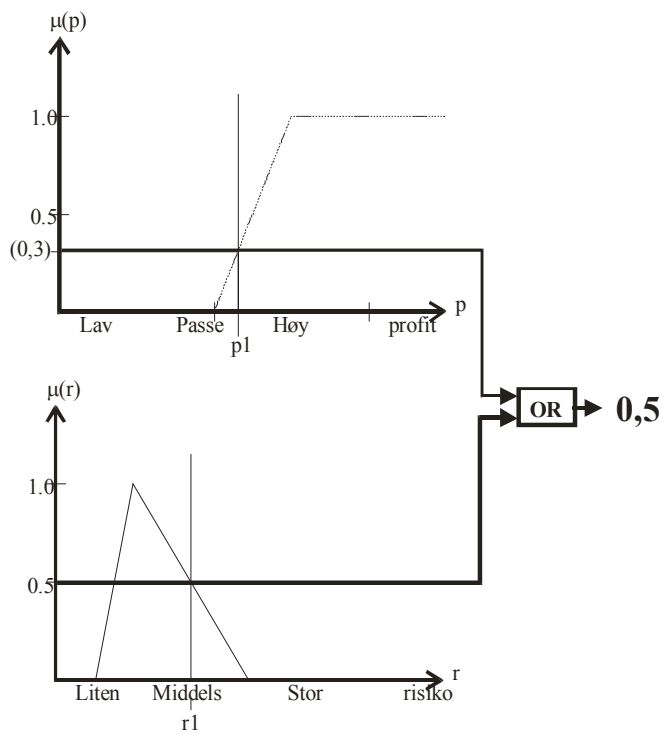
Regelen "If (profitt is **Høy**) AND (risiko is **Middels**) then ....." gir et snitt, (profitt  $\cap$  risiko). Dette snittet overfører en verdi fra delprosess 2 til delprosess 3.





Figur 12-11 Output fra AND-operasjon.

Regelen "If (profitt is **Høy**) OR (risiko is **Middels**) then ....." gir en union, (profitt  $\cup$  risiko).

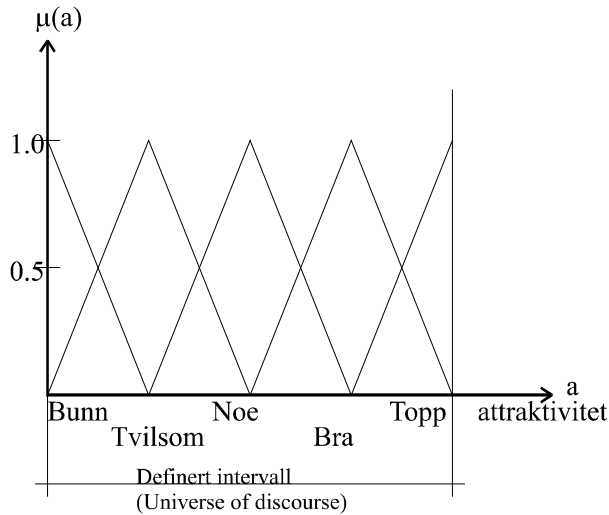


Figur 12-12 Output fra OR-operasjon.

Det overføres ett tall fra delprosess 2 til delprosess 3 for hver regel.

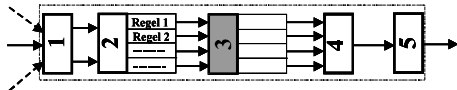
**Definere implikasjonsreglene.**

I fig. 12-10, som viser konsekvensmatrisen, så vi hvordan konsekvensene er funksjoner av parameterne profitt og risiko. Fig. 12-13 viser et forslag til medlemskapsfunksjoner for de forskjellige semantiske variablene.

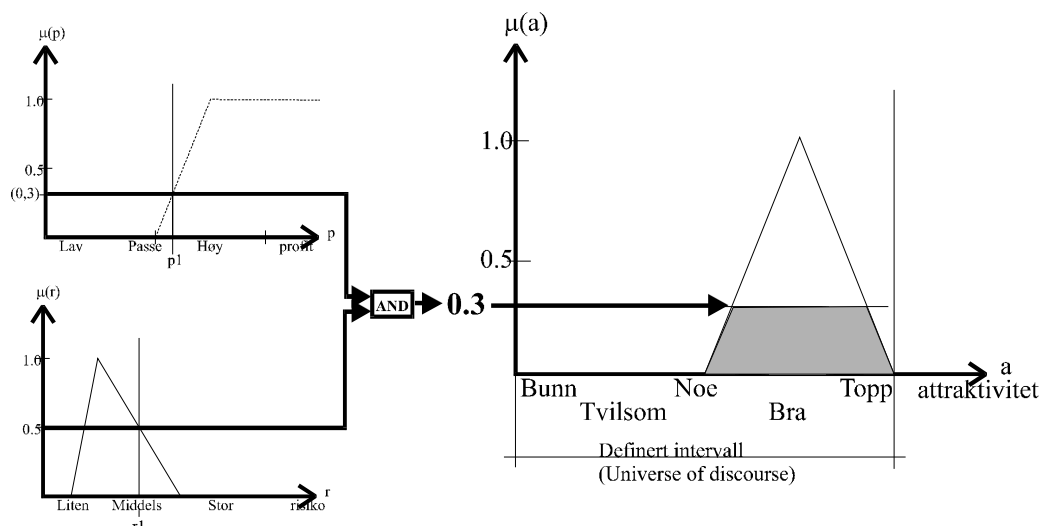


Figur 12-13 Medlemskapsfunksjoner for attraktivitet.

**Delprosess 3: Implikasjon**



Regelen "If (profitt is Høy) AND (risiko is Middels) then (attraktivitet is Bra)" vil, hvis vi uttrykker AND med en min-operator, implisere konsekvensen "Bra" i en grad av 0,3

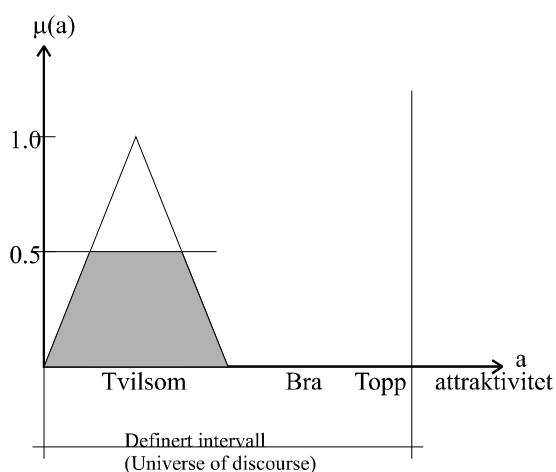


Figur 12-14 Sammenheng mellom input og konsekvens. Implikasjon ved avkorting.

Det skraverte feltet i fig. 12-14 er et mål på hvor mye av den diffuse mengden "Bra" som er med i den totale konsekvensmengden. Fra delprosess 3 til delprosess 4 overføres diffuse mengder som angir grader av konsekvens. Det overføres én slik mengde for hver If-then-regel.

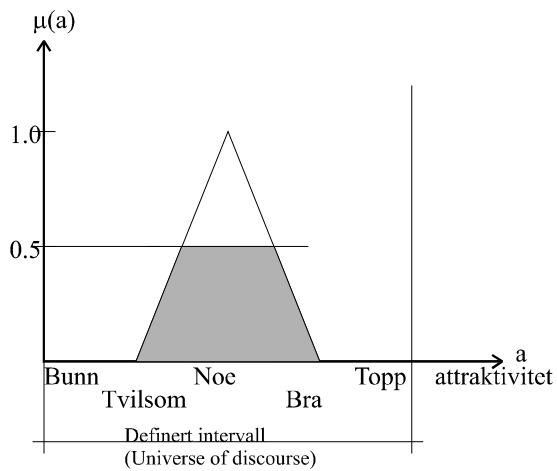
Før vi går over til å vise hvordan konsekvensene aggregeres skal vi utvikle et par implikasjoner til.

Regelen "If (profitt is Passe) AND(risiko is Stor) then (attraktivitet is Tvilsom)", se matrisen fig. 12-10, vil tilsvarende til det vi har vist ovenfor implisere konsekvensen "Tvilsom" i en grad av  $\min(0.7,0.5) = 0,5$ , og den diffuse mengden som angir hvor attraktivt prosjektet er (konsekvensmengde) som vist i fig. 12-15.



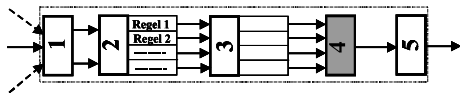
Figur 12-15 Implikasjon av "profitt er Passe og risiko er Stor".

Regelen "If (profitt is Passe) AND (risiko is Middels) then (attraktivitet is Noe)" vil implisere en konsekvensmengde som vist i fig. 12-16.

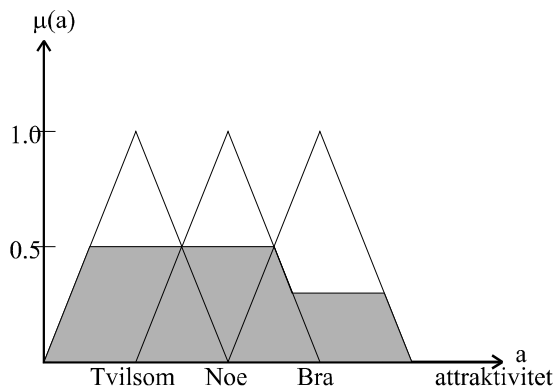


Figur 12-16 Implikasjon av "profitt er Passe og risiko er Middels".

#### Delprosess 4: Aggregering

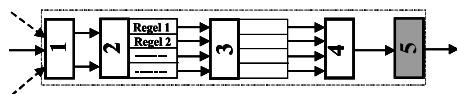


Sammenstilling av konsekvensmengdene kalles aggregering. Dette er vist i fig. 12-17, og representerer, som vi ser, unionen av de aggregerte mengdene.

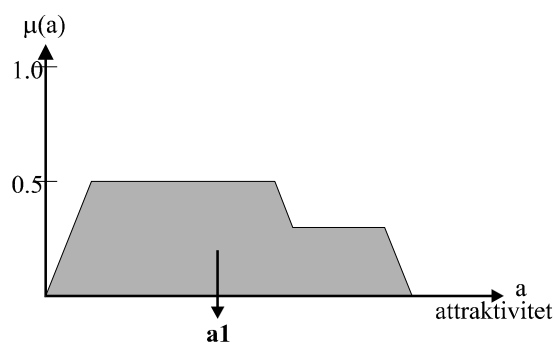


Figur 12-17 Aggregerte konsekvensmengder.

### Delprosess 5: Avdiffusifisering.



I delprosess 5 trekkes det mest representative element ut av denne aggregerte mengde i form av en tallstørrelse innenfor det definerte intervallet for konsekvensen. Dette kalles å avdiffusifisere resultatet. Det mest vanlige er å benytte tyngdepunktet i mengden.



Figur 12-18 Tyngdepunktet i aggregert mengde.

Størrelsen  $a_1$  er resultatet fra fuzzymaskinen med input  $p_1$  og  $r_1$ , og vil i sin enkleste form definere et punkt i rommet med koordinatene  $(p_1, r_1, a_1)$ .

Funksjonen  $a = f(p,r)$  vil da beskrive en flate i rommet.

## 12.4 Eksempel på bruk av fuzzy logic

### Presentasjon av casen

Caseprosjektet er grundig beskrevet i vedlegg 1: Statens vegvesen. Håndbok-140: "Konsekvensanalyser. Del III. Eksempel". [10]

Oppgaven består i valg blant tre alternative løsninger for veiutløsning til øya Vassøy utenfor Stavanger. I tillegg har man alltid "nullalternativet", d.v.s. å ha det som før.

**Alternativ 0** er eksisterende ferje fra Stavanger til Vassøy.

**Alternativ 1** er å gå fra Bjørnøy via Langøy, som dermed også får veiutløsning, og videre via Hestholmen og bru til Vassøy

**Alternativ 2** går noe lenger sør fra Bjørnøy via Roaldsøy, som alt har veiforbindelse, og med bru til Vassøy.

**Alternativ 3** er undersjøisk tunnel fra Bjørnøy til Vassøy.

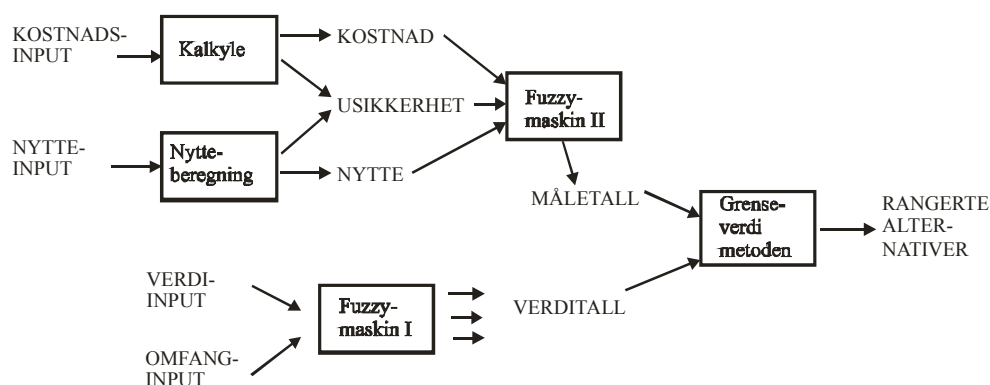
Konsekvensanalysen består i å regne ut kostnad og nytte i forhold til nullalternativet,

og veie netto nytte (nytte - kostnad) opp mot usikkerheten i overslagene, og de konsekvensene som ikke lar seg prissette. Det siste består i hovedsak av kvalitative nytteparametre og miljøbelastninger.

Eksemplet er tilpasset<sup>40</sup> de justeringene som er gjort i Håndbok – 140 i 2005. [11]

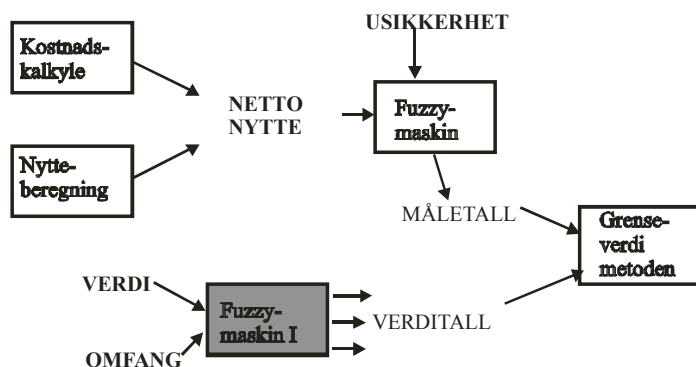
### Prosessene i eksemplet

Figur 12-19 viser systembeskrivelsen for oppgaveløsning.



Figur 12-19 Systemskisse for konsekvensanalyse.

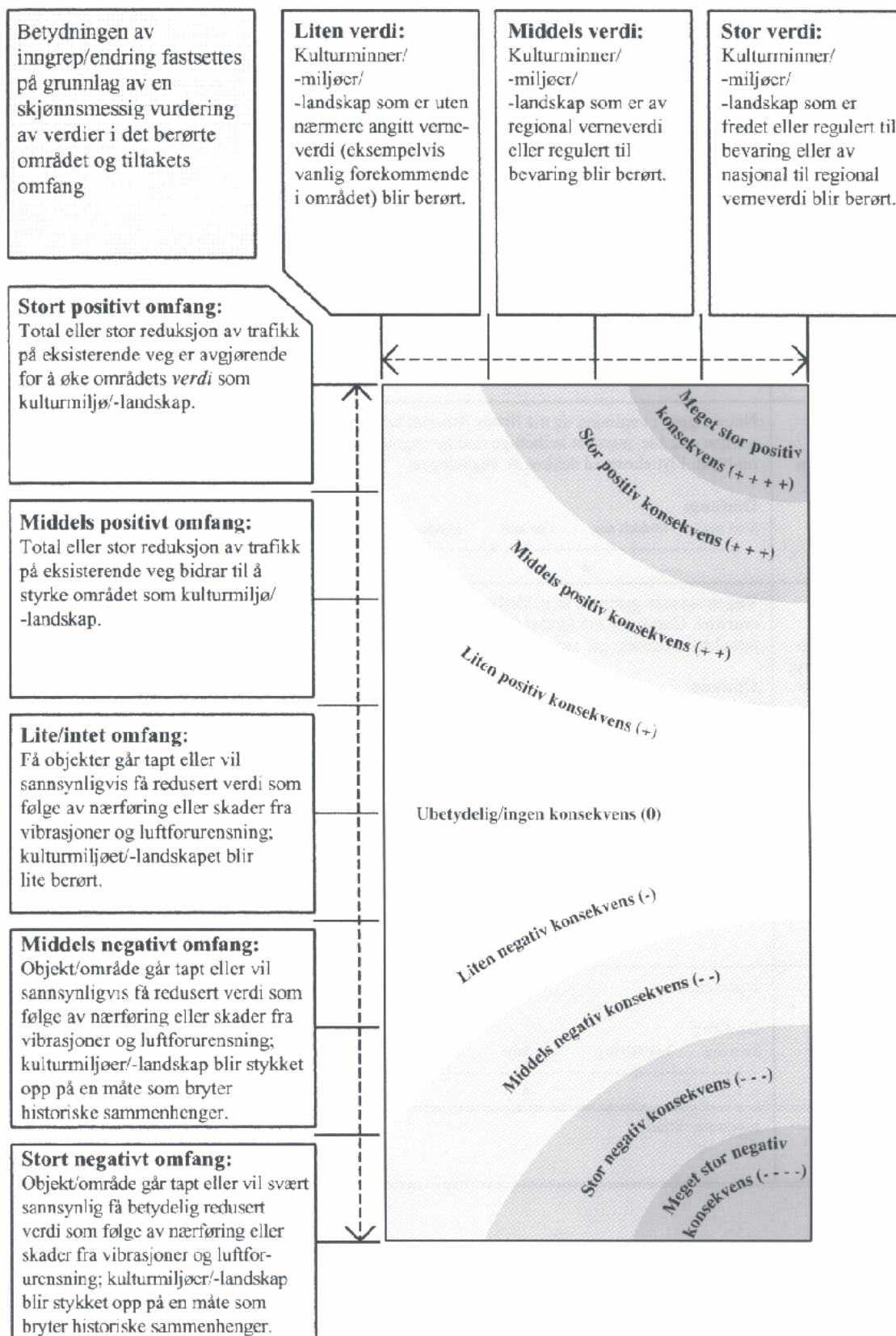
### Fuzzymaskin I: Forhold som ikke er målbare i penger.



Som vi ser, tas der her utgangspunkt i vurderinger om verdi og omfang av de såkalt ikke-prissatte konsekvenser knyttet til vegalternativene.

Vurderingene i bilag 1, og klassifisering av konsekvensene tar utgangspunkt i konsekvensmatrisen vist i figur 12-20.

<sup>40</sup> Denne tilpasningen er gjort på skjønn av artikkelforfatteren, og representerer ingen ”offisiell” nyvurdering av prosjektalternativene.

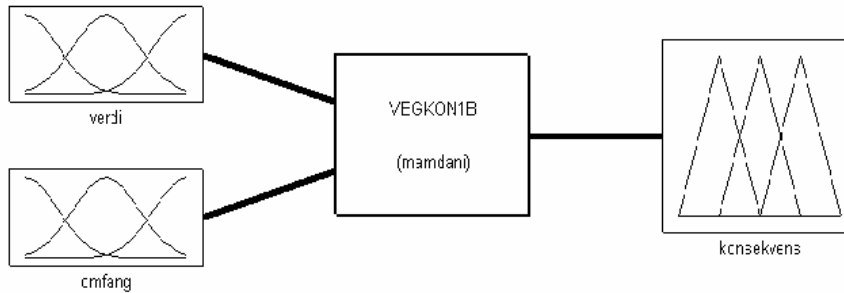


Figur 12-20 Konsekvensmatrise som beskriver hvordan ikke-prissatte konsekvenser skal klassifiseres i henhold til vegvesenets nye retningslinjer for konsekvensanalyse.

Vi skal siden vise hvordan denne matrisen kan bygges opp i fuzzy logic ved bruk av et dataprogram.

I det følgende skal vi beskrive fuzzymaskinen slik den bygges opp i dataprogrammet "**Fuzzy Logic Toolbox with Matlab™**".

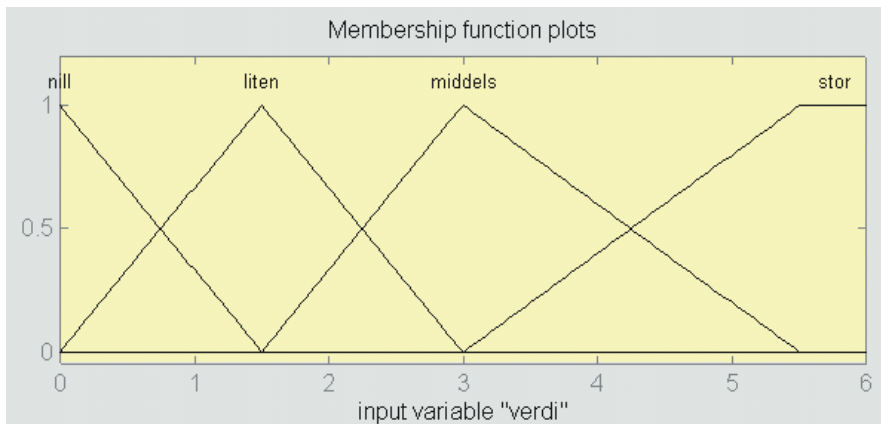
Figur 12-21 viser fuzzymaskinen med tilhørende input/output-definisjoner.



Figur 12-21 Input-variabler og tilhørende konsekvens i fuzzymaskin I.

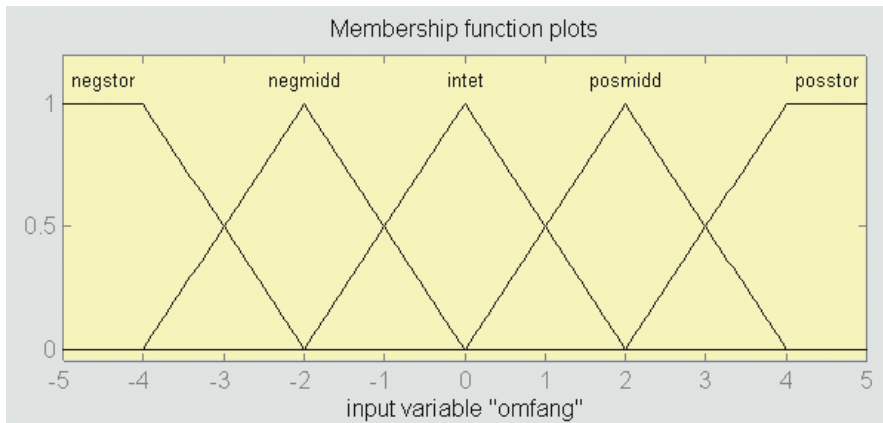
I figur 12-22 ser vi hvordan input-variabel "verdi" er definert med mengde, semantiske variabler og medlemskapsfunksjoner.

Figur 12-23 og figur 12.24 vises det samme for henholdsvis input-variabel "omfang" og output-variabel "konsekvens".

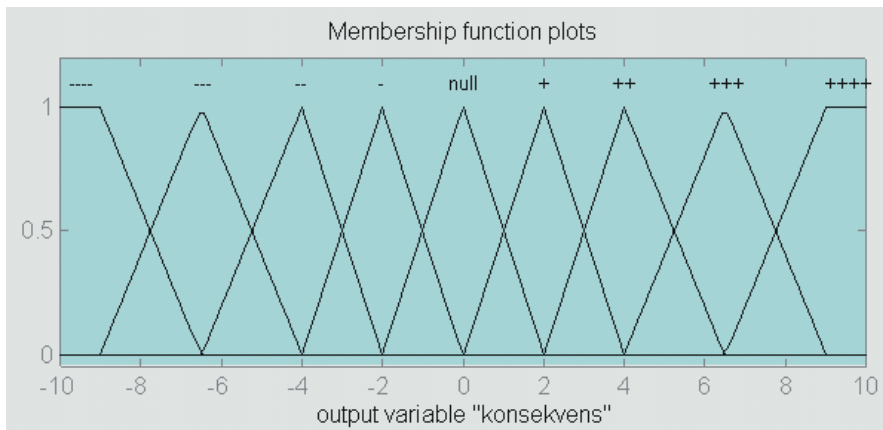


Figur 12-22 Input-variabel "verdi".





Figur 12-23 Input-variabel "omfang".



Figur 12-24 Output-variabel "konsekvens".

"Verdi" har et definert intervall fra 0 til 6. Variabelen "omfang" har et definert intervall fra -5 til +5.

Variabelen "konsekvens" gir ut en tallstørrelse som skal være sammenlignbart med resultatet fra den andre fuzzymaskinen. Intervallet er i begge tilfeller fra -10 til +10.

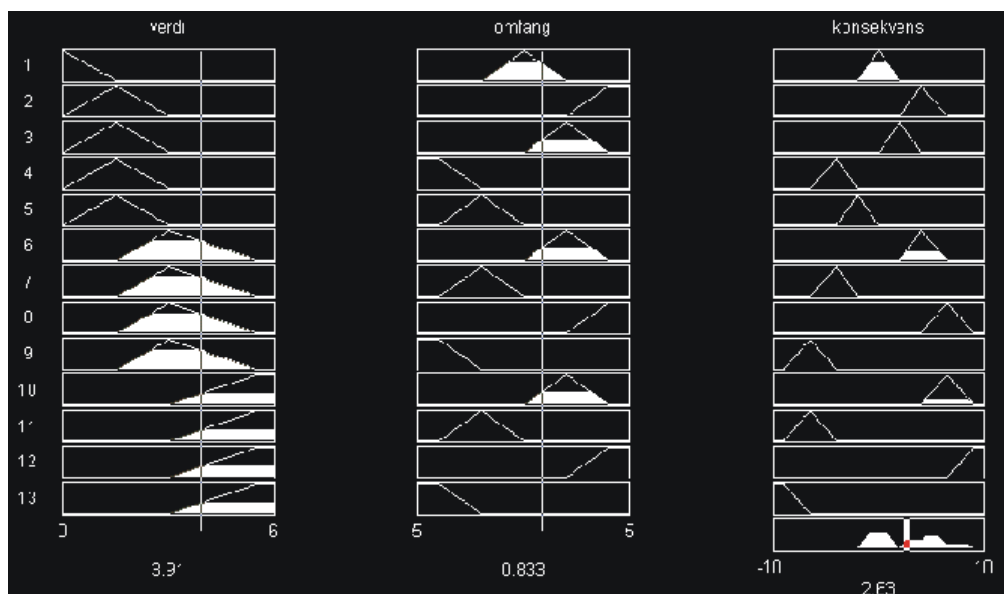
Reglene som gir grunnlag for de logiske operasjonene er vist i figur 12-25.

1. If (verdi is null) or (omfang is intet) then (konsekvens is null) (1)
2. If (verdi is liten) and (omfang is posstor) then (konsekvens is ++) (1)
3. If (verdi is liten) and (omfang is posmidd) then (konsekvens is +) (1)
4. If (verdi is liten) and (omfang is negstor) then (konsekvens is --) (1)
5. If (verdi is liten) and (omfang is negmidd) then (konsekvens is -) (1)
6. If (verdi is middels) and (omfang is posmidd) then (konsekvens is ++) (1)
7. If (verdi is middels) and (omfang is negmidd) then (konsekvens is --) (1)
8. If (verdi is middels) and (omfang is posstor) then (konsekvens is +++ ) (1)
9. If (verdi is middels) and (omfang is negstor) then (konsekvens is ---) (1)
10. If (verdi is stor) and (omfang is posmidd) then (konsekvens is +++ ) (1)
11. If (verdi is stor) and (omfang is negmidd) then (konsekvens is ---) (1)
12. If (verdi is stor) and (omfang is posstor) then (konsekvens is ++++) (1)
13. If (verdi is stor) and (omfang is negstor) then (konsekvens is ----) (1)

Figur 12-25 If-then-reglene i fuzzymaskin I.

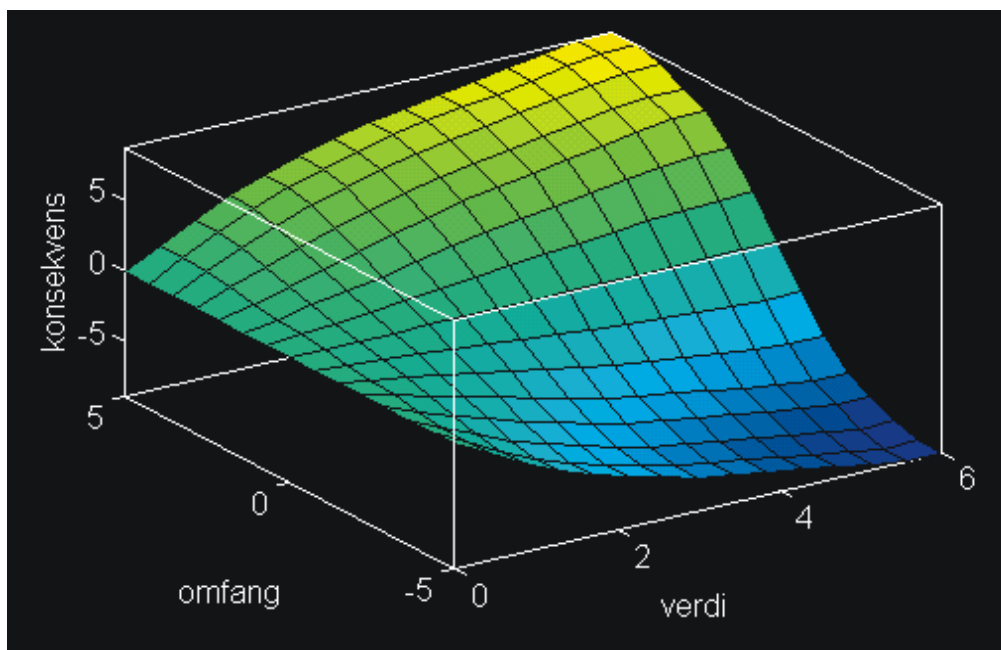
Hvordan toolboxen viser sammenhengen mellom input, regler og impliserte konsekvenser ser vi i figur 12-26. Figuren viser også sammenslåing av de diffuse mengdene som representerer alle konsekvensene, og hvordan de sammenslåtte mengdene blir omgjort til et resultat, (avdiffusifisert).

For et gitt sett av input for hvert alternativ får vi et resultat, som vi i dette tilfellet kaller alternativets verditall.



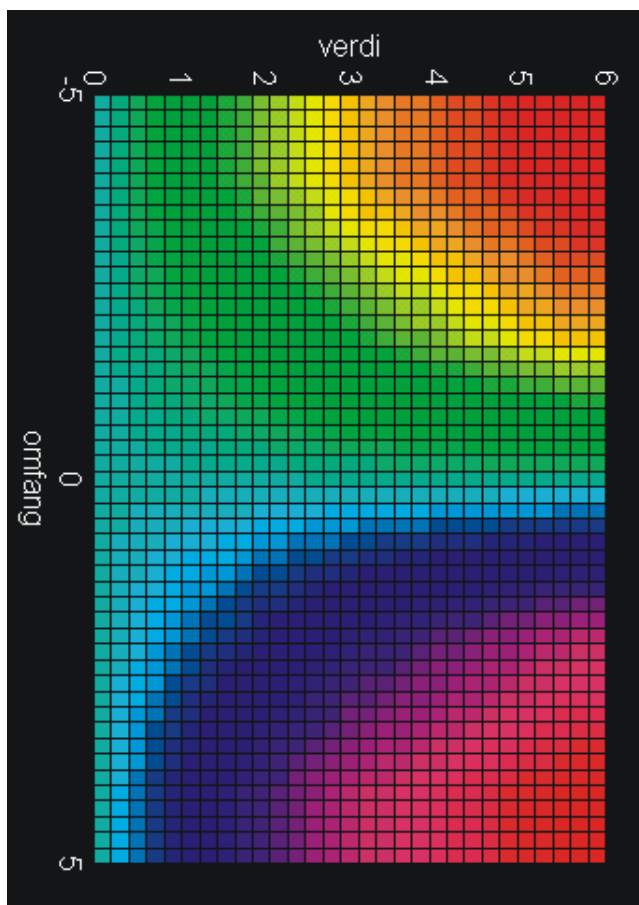
Figur 12-26 Fuzzymaskin I. Samlet oversikt.

Sammenhengen mellom alle mulige input-kombinasjoner og tilhørende resultater er vist som en flate i rommet i figur 12-27, og som en konsekvensmatrise i figur 12-28.



Figur 12-27 Resultatfunksjonen fra fuzzymaskin I.

Figur 12-28 gir en todimensjonal fremstilling av romflaten i figur 12-27. De grafiske fremstillingene i dataprogrammet er tilpasset bruk av farger. Dette gjør at ”høydekurvene” ikke kommer til sin rett i svart/hvitt-fremstilling, men vi ser tydelig sammenhengen med konsekvensmatrisen i figur 12-20.



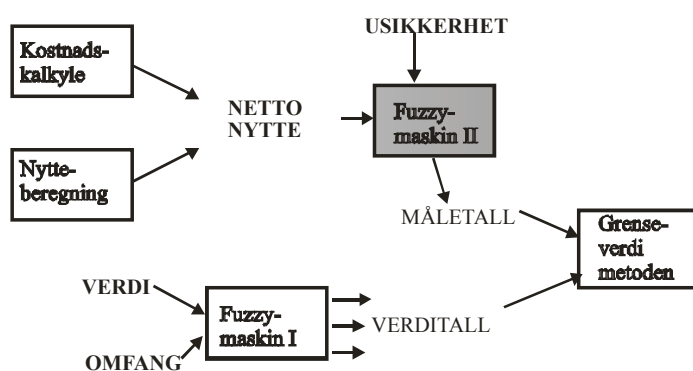
Figur 12-28 Resultatmatrisen fra fuzzymaskin I.

I tabell 12-1 er gjengitt resultatet for hver konsekvens som er vurdert. Input til fuzzymaskinen er hentet fra [10].

Tabell 12-1 Vekting av resultatene fra fuzzymaskin I:

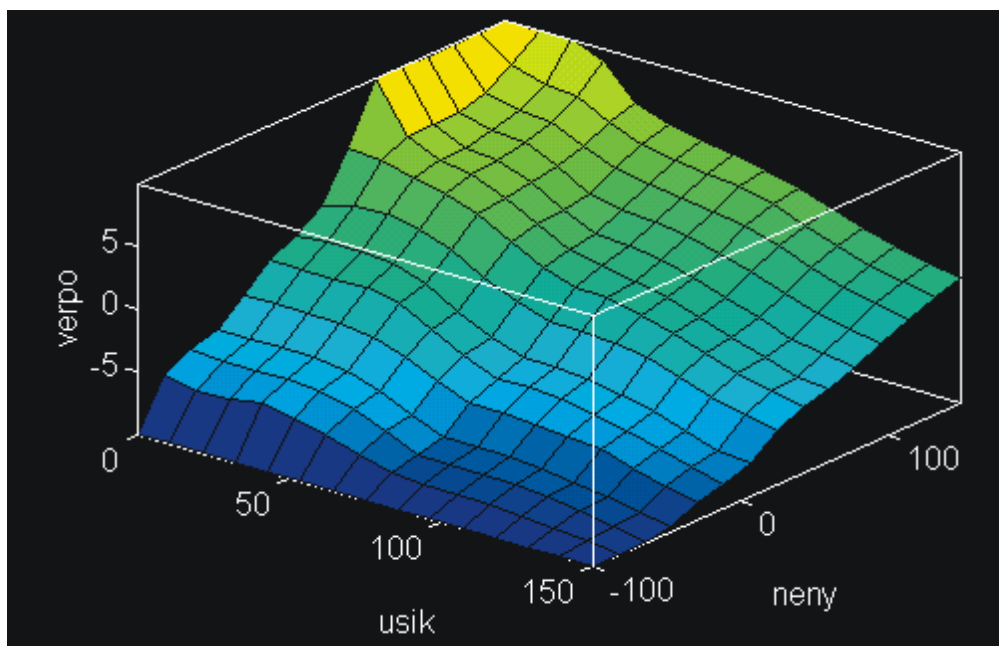
<i>Konsekvens for:</i>	<i>Alternativ 1</i>	<i>Alternativ 2</i>	<i>Alternativ 3</i>
<i>Landskapsbilde</i>	-6,5	-5,0	-3,0
<i>Nærmiljø og friluftsliv</i>	-3,0	-4,5	-3,5
<i>Kulturmiljø</i>	-2,5	-5,0	+1,0
<i>Naturmiljø</i>	-8,0	-5,0	-3,0
<i>Naturressurser</i>	-1,0	0	0
<i>Sum</i>	-21,0	-19,5	-8,5
<i>Verditall (samlet verdi som overføres til grenseverdimetoden)</i>	-4,2	-3,9	-1,7

### Fuzzymaskin II: Omgjøring av kombinasjonen netto nytte og usikkerhet til måletall.



Her tar vi utgangspunkt i resultatet fra nytteberegningen og kostnadskalkylene og finner netto nytte. Dette utgjør, sammen med en usikkerhetsvurdering, input i fuzzymaskinen.

Både netto nytte ("nyen") og usikkerhet målt i standardavvik ("usik") er angitt i mill. NOK.



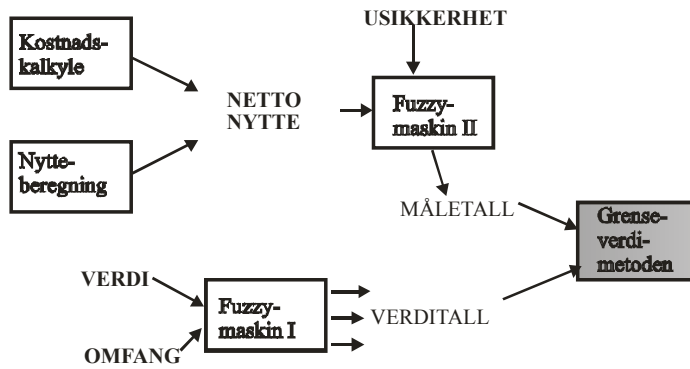
Figur 12-29 Resultatfunksjon fra fuzzymaskin II.

Tabell 12-2. Resultat fra fuzzymaskin II.

Alternativ 1			Alternativ 2			Alternativ 3		
Netto nytte	Usikkerhet	Måletall	Netto nytte	Usikkerhet	Måletall	Netto nytte	Usikkerhet	Måletall
114	75	<b>3.0</b>	87.5	75	<b>1.5</b>	-12.6	75	<b>-5.0</b>

Måletallene, hhv 3.0, 1.5 og -5.0, overføres til grenseverdimetoden.

### Rangering av alternativene.



Rangering av alternativene skjer etter grenseverdimetoden hvor resultatet fra begge fuzzymaskinene blir underlagt en vektning.

$v_I$  representerer den vekt som legges på de ikke-prissatte forhold, og  $v_{II}$  er den vekt som legges på de økonomiske forholdene.

$$v_I + v_{II} = 1.0.$$

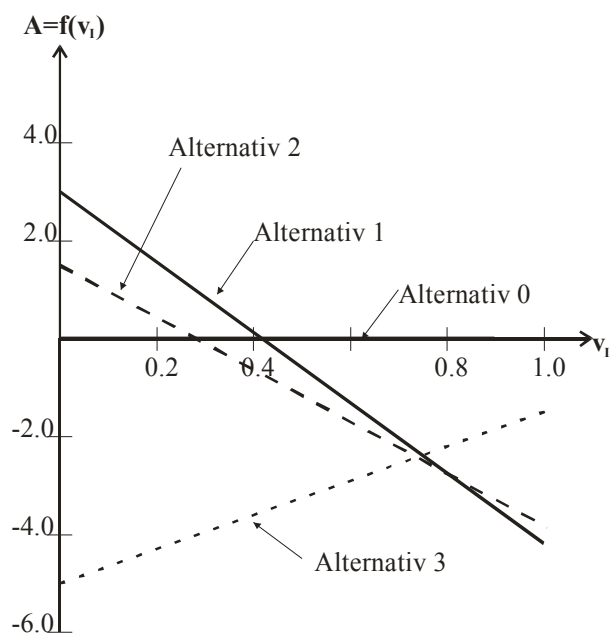
Dette gir følgende funksjoner:

$$\text{Alternativ 0: } \underline{A_0} = 0$$

$$\text{Alternativ 1: } \underline{A_1} = -4,2v_I + 3,0v_{II} = \underline{3,0 - 7,2 v_I}$$

$$\text{Alternativ 2: } \underline{A_2} = -3,9v_I + 1,5v_{II} = \underline{1,5 - 5,4 v_I}$$

$$\text{Alternativ 3: } \underline{A_3} = -1,7v_I - 5,0v_{II} = \underline{-5,0 + 3,3 v_I}$$



Figur 12-30 Alternativenes totalpoeng som funksjon av hvor stor vekt det legges på de ikke-prissatte faktorene.

Vi ser at alternativ 1 er det beste så lenge  $v_1$  er mindre enn ca. 0,4, mens nullalternativet overtar som det beste alternativet når  $v_1$  er større enn 0,4. Hvis vi ser bort fra nullalternativet vil alternativet med tunnel under fjorden være det beste når de ikke-økonomiske forhold gis en vekt på mer enn ca. 0,75. Alternativ 2 vil aldri være et aktuelt valg med de vurderinger som er gjort ovenfor.

I dette eksemplet er fuzzy logic brukt som hjelp for å bygge en datorisert beslutningsmodell basert på et ekspertsystem. Input som ekspertene må skaffe er nytte, kostnad og usikkerhet for de faktorene som kan måles i penger, og vurderinger om verdier i det berørte området og omfang av de konsekvensene som ikke kan måles i penger. Output fra beslutningsmodellen vil være en del av beslutningsgrunnlaget for, i dette tilfellet, valg av ett blant fire alternativer.

Ved hjelp av fuzzy logic er det forbausende enkelt å bygge denne typen datamodeller. I tillegg til enkelheten ligger det en betydelig egenverdi i det forhold at modellen krever en omforent avgrensning av diffuse, situasjonsbetingede størrelser som "stor", "liten", "mye", "ofte" osv. En annen fordel ved bruk av fuzzy logic er at det er relativt enkelt å lage modeller som behandler input fra mange kilder.

Svakheten er at det kreves en datatoolbox for å lage og kjøre modellene.



## 12.5 Referanser

1. Ross, Timothy J: Fuzzy Logic with Engineering Applications McGraw-Hill Inc. 1995
2. Heske, Ted and Heske, Jill Fuzzy Logic for Real World Annabooks 1996  
Neparent: Design
3. McNeill, Martin F. and Fuzzy Logic. A Practical AP Proffesional 1996  
Thro, Ellen: Approach
4. Kasko, Bart: Fuzzy thinking Harper Collins 1994
5. Cox, Earl: The Fuzzy Systems Handbook AP Professional 1994
6. Manual til Fuzzy Logic The MathWorks 1995  
Toolbox i tilknytning til Matlab
7. Nørgaard Larsen, Asker og Tradition og begær. En Centrum Forlag 1993  
Dahlgaard, Su Mi Park: kulturhistorie om Østens og Vestens vei til år 2000
8. Holm, Frank Henning Byggeøkonomi Universitetsforlaget 1993
9. Holberg, Ludvig: Komedier: Jeppe paa Bjerget, Gyldendal Norsk Forlag  
Den politiske Kandestøber og 1976  
Erasmus Montanus
10. Vegdirektoratet Konsekvensanalyser Statens Håndbokserien nr.140  
Vegvesen 1996
11. Vegdirektoratet Konsekvensanalyser Statens Håndbokserien nr.140  
Vegvesen 2005

## 12.6 Vedlegg A: Terminologi

Jeg finner det nødvendig å avklare en del ord og uttrykk som er brukt i dette kapitlet. Noen av disse begrepene er spesielle for temaet fuzzy logic:

*Aggregering:* Sette sammen alle diffuse mengder som blir implisert fra If-then-regler til en samlet diffus mengde.

*Avdiffusifisering:* Å finne en tallstørrelse som representerer en diffus mengde.

*Bayesisk sannsynlighet:* Grad av tiltro basert på rasjonelle, subjektive vurderinger.

*Definert intervall:* (*Universe of discourse*). Det området av en fuzzy variabel som er av interesse i den aktuelle analysen.

*Diffuse tall:* (*Fuzzy number*). Tallstørrelser som representerer et intervall. F.eks. ca 3 kan være intervallet  $\langle 2, 4 \rangle$ .

*Diffus mengde:* (*Fuzzy set*). En mengde som ikke krever at elementene enten bare er i mengden eller bare utenfor mengden, men som tillater at elementer kan være bare delvis innenfor; altså tillater grader av medlemskap.

En diffus mengde er en utvidelse av det klassiske mengdebegrepet. Hvis  $X$  er det definerte intervall (universe of discourse) og elementene er betegnet med  $x$ , og  $A$  en diffus mengde knyttet til en semantisk variabel, er den diffuse mengden  $A$  av  $X$  definert som det sett av ordnede par,  $(x, \mu_A(x))$ , slik at  $A = \{x, \mu_A(x) \mid x \in X\}$ .  $\mu_A(x)$  kalles medlemskapsfunksjonen for  $x$  i  $A$ .

*Diffusifisering:* Å gjøre input i form av tallstørrelser om til grad av medlemskap i diffus mengde.

*Ekspertsystem:* Datasystem som prosesserer menneskelig kunnskap.

*Fuzzymaskin:* Det sett av prosesser som benyttes for å gjennomføre en analyse ved hjelp av fuzzy logic.

*Fuzzy variabel:* Variabel som deles i mer eller mindre overlappende intervaller som hver har tilknyttet en diffus mengde med tilhørende medlemskapsfunksjon. Typiske Fuzzy variabler er diffuse tall og lingvistiske variabler.

*Implikasjon:* Følge av logisk operasjon. I fuzzymaskinen er implikasjon en egen delprosess som gir fra seg en konsekvens i form av en diffus mengde.

*Konsekvens:* Resultat av logisk operasjon basert på If-then-regler. Presenteres ved diffuse mengder knyttet til semanter.

*Lingvistisk variabel:* Fuzzy variabel som er beskrevet med ord, gjerne adjektiv eller adverb (stor, fort)

*Semantisk variabel:* Navnet på et intervall innenfor en lingvistisk variabel tilknyttet en diffus mengde.

## 13 VERDI OG KONSEKVENSER AV NY INFORMASJON

KJELL J. SUNNEVÅG\*

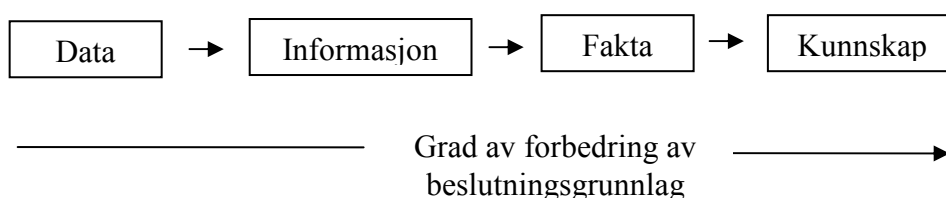
### 13.1 Innledning

I alle faser av et prosjekt vil det være usikkerhet forbundet med faktorene som inngår i vurderingen av prosjektet, det være seg kostnader, inntekter og andre konsekvenser av gjennomføringen av prosjektet, samt konsekvenser av alternativet, dvs. å ikke gjennomføre prosjektet. Desto lengre tid det er fra beslutningstidspunktet til prosjektet settes i gang, og mye av usikkerheten blir avslørt, desto større vil utfallsrommet være. På den annen side vil det også hele tiden komme ny informasjon som bidrar til å oppdatere utfallsrommet. Ny informasjon kan øke så vel som minske utfallsrommet.

Fokus i dette kapittelet vil være på hvordan man håndterer ny informasjon som kommer etter hvert som de ulike beslutningspunktene nærmer seg, og informasjon som først vil avsløres etter at beslutning om investering er tatt. Sentralt i denne drøftingen vil selvsagt bayesiansk metodikk stå. Men før vi kommer inn på en bayesiansk tilnærming til ny informasjon – og *verdien* av ny informasjon – vil vi imidlertid kort definere selve informasjonsbegrepet, og den mer tradisjonelle tilnærmingen til beskrivelse av usikkerhet. Kapittelet avsluttes med en gjennomgang av hvordan informasjonsavsløring som skjer etter at prosjektet er igangsatt kan påvirke både innfasing og valg av prosjektkonsept. Perspektivet her er en realopsjonstilnærming til prosjektvurdering.

### 13.2 Hva er informasjon?

Informasjon er et begrep med mange betydninger, avhengig av kontekst. Data er bare en grov form for informasjon, og representerer ikke kunnskap i seg selv. Sekvensen fra data til kunnskap er illustrert i figuren under.



Figur 13-1. Fra data til kunnskap

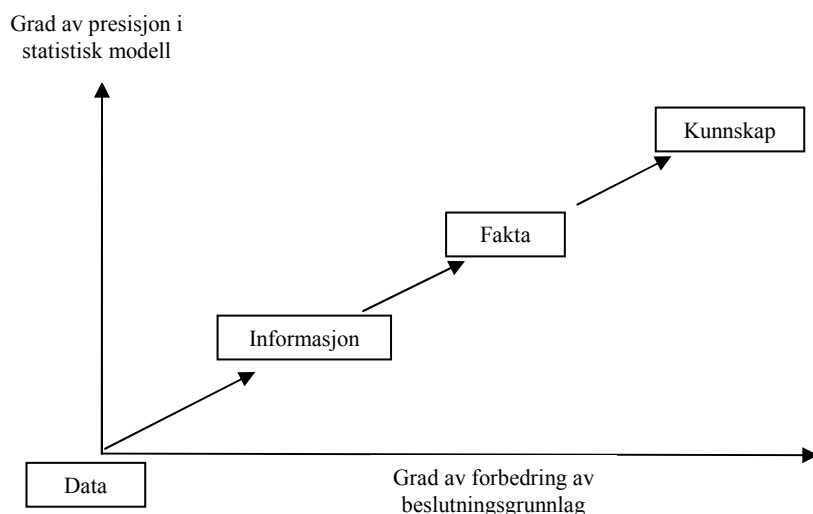
Data blir informasjon når det blir relevant for beslutningsproblemet. Informasjon blir fakta når data kan understøtte dette. Fakta er det som dataene avslører. En vanlig forståelse av begrepet kunnskap er at det representerer en form for bevissthet omkring, eller forståelse av fakta,

\* Kjell J. Sunnevåg er forskningssjef ved SNF

etablerte sannheter eller informasjon ervervet gjennom erfaring eller læring (*a posteriori*), eller selvinnsikt (*a priori*). I denne sammenheng kan vi altså si at fakta blir kunnskap når det har blitt benyttet på en vellykket (eller ikke vellykket) måte i en avsluttet beslutningsprosess.

Svært ofte er det betydelig usikkerhet knyttet til i hvilken grad de data som er tilgjengelige, er relevant for beslutningsproblemet. Data er representert i form av *signal* som inneholder *støy*, og som må tolkes ut fra kontekst. Derfor vil det ofte være nødvendig å legge en statistisk modell til grunn for en sannsynlighetsmodellering og tolkning av dataene.

Siden en mer presis statistisk modell også innebærer en forbedring av grunnlaget for å fatte beslutninger, kan vi tilpasse figuren med prosessen fra data til kunnskap, der vi legger inn graden av presisjon i den underliggende statistiske modellen:



Figur 13-2 Grad av forbedring i beslutningsundrelag som funksjon av grad av presisjon i statistisk modell.

En *statistisk modell* kan tenkes på som et matematisk redskap som det er hensiktsmessig å benytte ved beskrivelsen av store mengder data. Grunnlaget vil typisk være et begrenset sett med data. Den statistiske modellen kan enten formuleres uten data for en påfølgende etterprøving med et innsamlet materiale, eller modellen kan ta utgangspunkt i en tilpasning til et sett med måldata. *Statistisk inferens* går i korthet ut på om vi kan trekke konklusjoner som går ut over datagrunnlaget – og om statistisk signifikans (utsagnskraft) kan knyttes til disse konklusjonene etter at vi har tatt hensyn til tilfeldig variasjon som en feilkilde.

Ved å ta i bruk en statistisk modell og sannsynlighetsmodellering, vil kunnskapen være basert på et systematisk grunnlag med utsagnskraft. Den kunnskapen som anvendes i beslutningsgrunnlaget uttrykkes da gjerne sammen med en grad av statistisk konfidens.

Dersom vi tar i betraktning den usikkerheten som er knyttet til beslutningsgrunnlaget, vil sannsynligheten for *ex ante* ”gode beslutninger” øke med tilgangen på ”god informasjon”. Man kan ikke ta gode beslutninger før man har tilgang på tilstrekkelig ”god informasjon” og således tilstrekkelig kunnskap. Sannsynligheten for at det foreligger tilstrekkelig ”god informasjon” øker med i hvilken grad vi er i stand til å tilrettelegge og tolke informasjonsgrunnlaget.

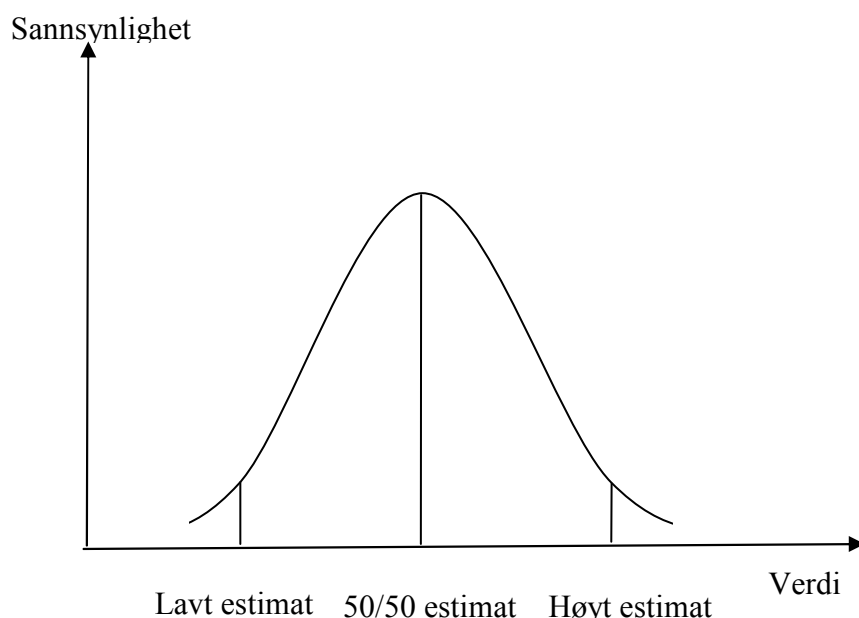
I denne sammenheng er det et viktig skille mellom kvalitativ og kvantitativ informasjon. *Kvalitativ informasjon* er informasjon som er deskriptiv i sin karakter, og som relaterer seg til, eller involverer kvalitet eller type. Noen benytter begrepet ”kvalitative data” for å representere kategoriske data, mens andre gjerne knytter det til data av mer verbal eller berettende karakter. Slike data kan være samlet inn gjennom intervjuer, spørreskjema med mulighet for svar av kvalitativ karakter eller andre mindre strukturerte situasjoner. Kvalitative data er gjerne representert i form av tekst, lyd eller bilde.

Dette står i motsetning til *kvantitativ informasjon*, som uttrykkes i form av en mengde eller relaterer seg til eller kan gjøres til gjenstand for måling. Kvantitative data omfatter måling av intervall eller grad, f.eks. IQ eller temperatur.

Det vil måtte legges ulike tilnæringer til grunn for analyse av henholdsvis kvalitative og kvantitative data. Kvalitative analytiske metoder kan f.eks. være teksttolking, tekstkoding og dekonstruksjon. Kvantitative analytiske metoder vil typisk være ulike statistiske teknikker.

### 13.3 Tradisjonell tilnærming til å beskrive usikkerhet

Når usikkerhet skal beskrives tas det vanligvis utgangspunkt i det som er usikkerhetsspennet i de enkelte faktorene. Ved ulike statistiske metoder beregnes så estimater for de enkelte faktorene, og den tilhørende usikkerheten for totalprosjektet. I NOU 1999:11 kan vi lese at det vanligste er å angi kostnadsestimater som et 50/50 estimat, noe som betyr at det er like stor sannsynlighet for at de endelige kostnadene blir høyere enn det angitte estimatet, som at de endelige kostnadene blir mindre enn estimatet. For å angi usikkerhetsmomentet i estimatet, angis et høyt og et lavt estimat i tillegg til 50/50 estimatet. Dette er illustrert i figuren under.

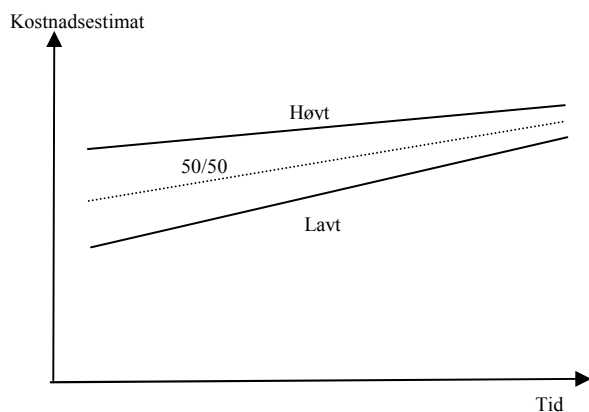


Figur 13-3 Illustrasjon av usikkerhetsspenn

Usikkerhetsspennet vil snevres inn etter hvert som mer informasjon om innholdet i prosjekt blir kjent og konkretisert gjennom planlegging og detaljering. I NOU 1999:11 nevnes det at

når oljeselskapet sender en plan for utbygging og drift til myndighetene, en såkalt PUD, vil det totale usikkerhetsspennet være nede i +/- 20%, mens det for enkelte kostnadskomponenter vil være helt oppe i +/- 40%.

Vi kan også merke oss at kravene til detaljering på estimatene avhenger av prosjektfasen. Etter hver som prosjektet ”modner” vil det bli utarbeidet mer detaljerte estimater, med et høyere konfidensnivå. Dette kan illustreres ved hjelp av figuren under:



Figur 13-4. Utvikling i usikkerhetsspenn over tid.

Kravene til detaljering er beskrevet i klassifikasjonssystemer som er etablert for enklere å kunne kommunisere hvilken grunnlagsinformasjon, estimeringsmetode, og usikkerhetsvurdering som kostnadsestimatene baserer seg på.

I NOU'en pekes det på at i estimeringssystemet inngår også en *database med grunnlagsdata*, herunder:

- Erfaringsdata fra gjennomførte prosjekter og forventet teknisk og markedsmessig utvikling.
- Estimater fra leverandører.
- Fortløpende erfaringer fra prosjekter i gjennomføring, både operatørens egne og partneropererte.
- Vurderinger med hensyn til usikkerhetsfaktorer (teknisk, forutsetninger, marked).

Usikkerheten i kostnadsestimatene er imidlertid ikke bare avhengig av modenhet og detaljeringsnivå i prosjektet, men også av kvaliteten i grunnlagsdataene. Kvaliteten kan bedres ved at det etableres databaser, der erfaringsdata oppdateres løpende. Slik kan en på det tidspunktet det leveres en plan for prosjektet som myndighetene skal ta stilling, etablere et best mulig datagrunnlag for estimering av kostnader og en tilhørende usikkerhetsvurdering.

På bakgrunn av analyse av historiske kvantitative data i en slik database, eller ekstrahering av mer ekspertbaserte kvalitative erfaringsdata, og selvfølgelig en tilpasning til den aktuelle kontekst, kan man etablere forventningsverdier og usikkerhetsspenn for de ulike aktuelle variablene som inngår i beslutningsgrunnlaget. Imidlertid vil det over tid komme ny informasjon som kan påvirke våre opprinnelige estimater. Ny informasjon kan komme

gjennom *aktiv* informasjonsinnhenting, for eksempel basert på nye analyser. Ny informasjon kan også komme gjennom *passiv* informasjonsavsløring, for eksempel utvikling i valutakurser eller råvarepriser, Spørsmålet er så hvordan slik informasjon skal håndteres på en faglig korrekt måte.

## 13.4 Verdi av ny informasjon

Det er viktig å huske på at dersom ny informasjon *ikke* vil påvirke beslutningen, vil den heller ikke ha noen verdi. I så fall bør man heller ikke være villig til å betale for innhenting av mer informasjon. For eksempel, dersom et bestemt prosjekt fremstår som det mest attraktive ut fra de to beslutningskriteriene som miljøutslipp og kostnader, vil det være lite hensiktsmessig å benytte tid og energi på å finne ut hvordan prosjektet skårer på andre variable, dersom disse ikke inngår blant beslutningskriteriene. Det som er av verdi, er å få et bedre informasjonsgrunnlag i forhold til det som er sentrale beslutningsvariable. Her vil begrep som a priori fordeling, a posteriori fordeling og Bayes' formel stå sentralt. Dette vil jeg komme tilbake til under. Først skal vi se nærmere på hvordan vi kan finne verdien av *ny* informasjon i et beslutningsproblem. Det vil være nødvendig å sette dette inn i en slik kontekst, nettopp fordi ny informasjon bare har en verdi i den grad den vil kunne påvirke beslutninger. Deretter vil vi komme inn på metoder for kvalitetssikring av informasjon.

Vi kan skille mellom to typer ny informasjon, perfekt og ikke perfekt ny informasjon. Perfekt ny informasjon vil være 100% sikker, mens det vil hefte usikkerhet ved ikke perfekt ny informasjon.

### Verdien av perfekt informasjon

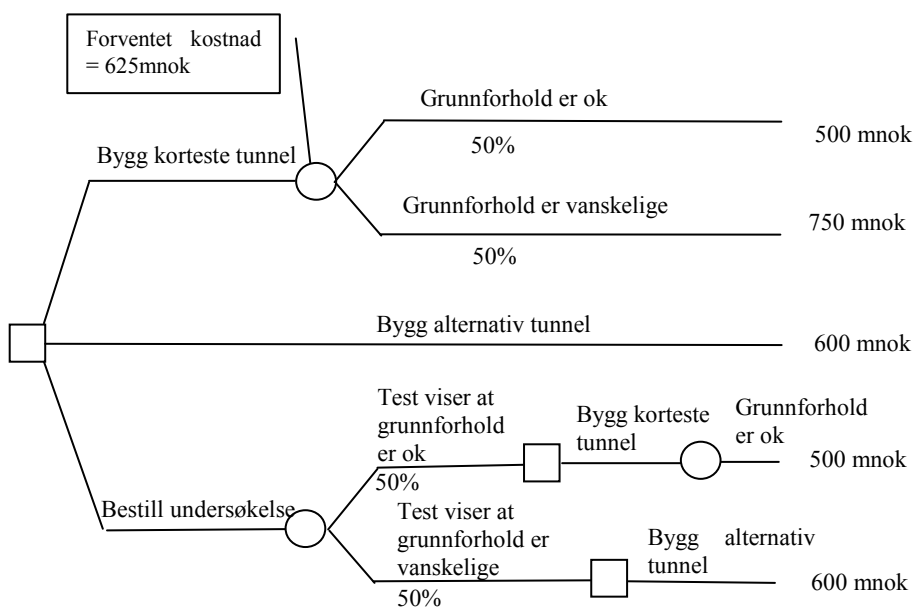
La oss ta følgende eksempel som utgangspunkt: En entreprenør skal bygge en lang og kostnadskreven tunnel. Det er en viss usikkerhet omkring de geologiske forholdene, og entreprenøren har to alternative valg med hensyn til utsprenging, det ene vesentlig lengre og mer kostbart enn det andre. Entreprenøren kan imidlertid bestille en geologisk undersøkelse som med 100% sikkerhet kan gi svar på om de geologiske forholdene skaper problemer for bygging av det korte alternativet eller ikke. En slik undersøkelse er imidlertid kostbar og tidkrevende. Hvis entreprenøren velger det korteste alternativet, og de geologiske forholdene viser seg å være vanskelige, vil dette medføre en betydelig ekstrautgift.<sup>41</sup>

Anta at det er 50% sannsynlig at de geologiske forholdene er uproblematiske. Den korte tunnelen vil koste 500 mnok, mens det lange alternativet vil koste 600 mnok. Hvis det viser seg at de geologiske forholdene er vanskelige, vil det påløpe ekstrakostnader på 250 mnok for den korte tunnelen. Spørsmålet er så hvor mye entreprenøren maksimalt er villig til å betale for en slik geologisk undersøkelse. Vi kan sette opp følgende beslutningstre.

---

<sup>41</sup> Eksempelet er tilpasset fra Goodwin og Wright (2004).





Figur 13-5. Beslutningstre for tunnelprosjekt

Dersom informasjon fra en geologisk undersøkelse ikke er mulig å fremskaffe, vil entreprenøren velge å bygge den alternative tunneltraseen siden de forventede kostnadene ved dette er lavest, dvs.  $0,5 \cdot 500 + 0,5 \cdot 650 = 625$  mnok sammenlignet med 600 mnok. Utfordringen nå er å finne verdien av perfekt informasjon, noe som da vil utgjøre et sammenligningsgrunnlag for hva vi er villig til å betale for en undersøkelse.

I utgangspunktet antar vi at det er 50% sannsynlig at grunnforholdene er vanskelige. Det er altså en tilsvarende sannsynlighet for at en undersøkelse vil vise dette. Sammenligningsgrunnlaget blir dermed den forventede verdien ved de to mulige utfallene dersom undersøkelsen gjennomføres. Dette er  $0,5 \cdot 500 + 0,5 \cdot 600 = 550$ .

Vi kan da sette opp følgende sammenstilling:

Forventet kostnad uten perfekt informasjon = 625 mnok

- Forventet kostnad med perfekt informasjon = 550 mnok

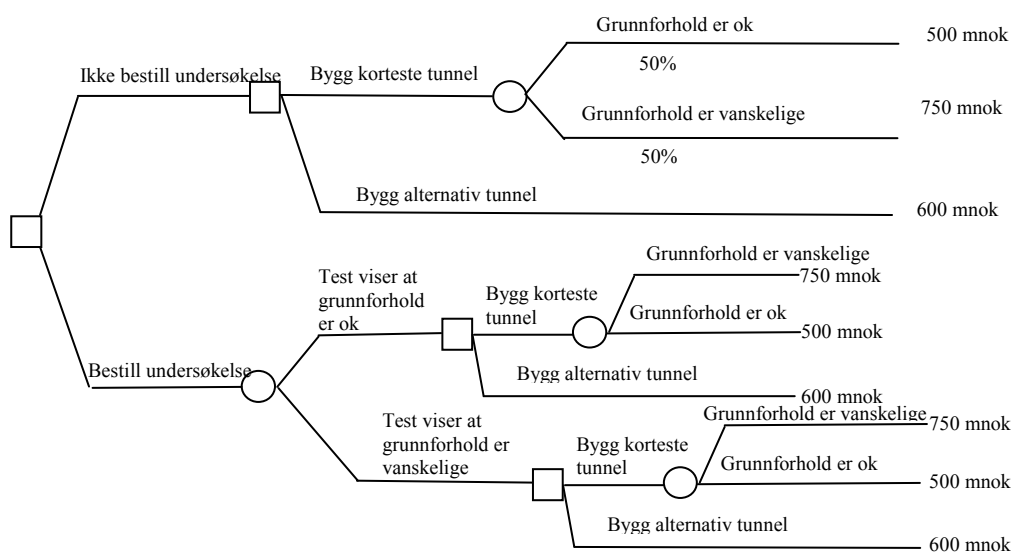
= Forventet verdi av perfekt informasjon = 75 mnok

Selv om vi ikke vet hva undersøkelsen vil koste, vet vi at det i hvert fall ikke er verdt å betale mer enn 75 mnok for å få den gjennomført - da under forutsetning av at den gir oss perfekt informasjon.

### Verdien av ikke perfekt informasjon

Spørsmålet er så hva verdien av ikke perfekt informasjon er. Vi vil fortsette å bygge på eksempelet over, men nå må vi trekke inn et bayesiansk rammeverk for å vurdere dette.

Beslutningsstrukturen i problemet blir noe annerledes enn i foregående eksempel, fordi det første vi må ta stilling til er om undersøkelsen skal gjennomføres eller ikke. Problemer er at undersøkelsen ikke kan avdekke om grunnforholdene er uproblematisk eller vanskelige med full sikkerhet – hvis det er ok grunnforhold, er det 90% sannsynlig at undersøkelsen vil avdekke dette. På den annen side, hvis grunnforholdene er vanskelige, er det 20% sannsynlig at undersøkelsen feilaktig likevel vil indikere at grunnforholdene er ok.



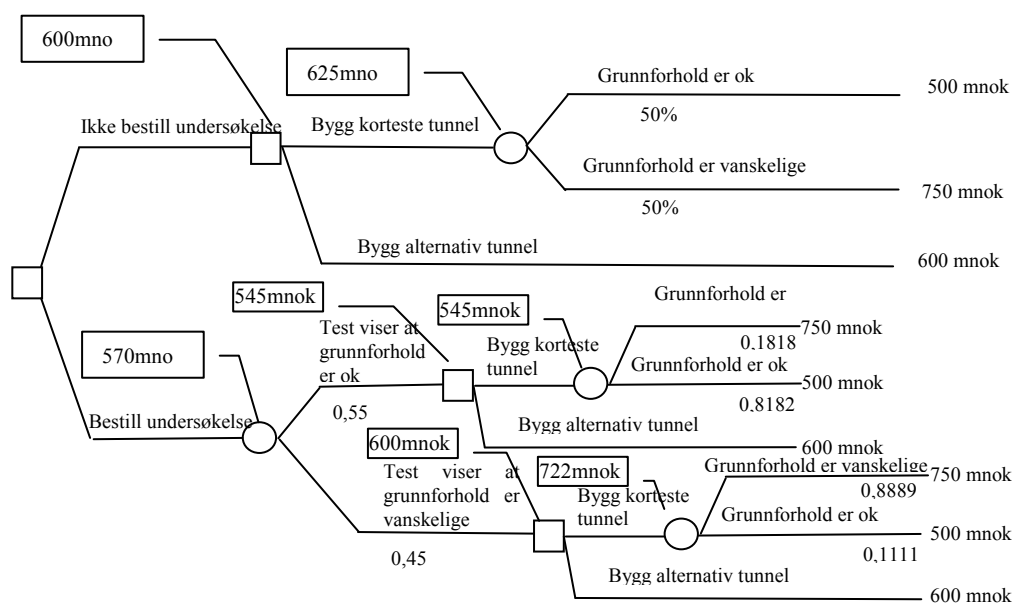
Figur 13-6 Beslutningstre for tunnelprosjekt når informasjon ikke er perfekt.

A priori sannsynligheten for at undersøkelsen viser at grunnforholdene er ok, og testen riktig indikerer dette, er  $0,5 \cdot 0,9 = 0,45$ . På den annen side er det en viss sannsynlighet for at grunnforholdene er vanskelige, men at undersøkelsen viser de er ok. Sannsynligheten for dette er  $0,5 \cdot 0,2 = 0,1$ . Dermed blir den samlede sannsynligheten for at undersøkelsen vil vise at grunnforholdene er ok på  $0,45 + 0,1 = 0,55$ . De respektive a posteriori sannsynligheter blir  $0,45/0,55 = 0,8182$  og  $0,1/0,55 = 0,1818$ , basert på Bayes' formel.<sup>42</sup>

Tilsvarende vil sannsynligheten for at grunnforholdene er ok, mens testen bommer og viser at de er vanskelige, være på  $0,5 \cdot 0,1 = 0,05$ . Videre vil sannsynligheten for at grunnforholdene er vanskelige, og at undersøkelsen viser dette være på  $0,5 \cdot 0,8 = 0,4$ . Den samlede sannsynligheten for at undersøkelsen viser at grunnforholdene er vanskelige vil være 0,45. Her blir de respektive a posteriori sannsynligheter  $0,05/0,45 = 0,1111$  og  $0,4/0,45 = 0,8889$ .

Vi kan sette sannsynlighetene og tilhørende forventningsverdier inn i beslutningstreet over.

<sup>42</sup> Bayes' formel:  $\Pr(A|B) = \Pr(B|A) \cdot \Pr(A) / \Pr(B)$ .



Figur 13-7 Beslutningstre med oppdaterte sannsynligheter basert på bayesiansk tilnærming

Ved de ulike beslutningspunkter har vi funnet hva som er optimal beslutning, gitt de ulike forventede verdier ved de ulike mulige beslutningsalternativer. Dette blir forventet verdi ved denne noden. Så har vi regnet oss tilbake i treet. Vi kan dermed sette opp regnestykket som viser verdien av ikke perfekt informasjon, og som blir sammenligningsgrunnlaget for hva vi er villig til å betale for dette.

Forventet kostnad uten undersøkelsen = 600 mnok

- Forventet kostnad med ikke perfekt informasjon

$$= 0,55 \cdot 545 + 0,45 \cdot 600 = \underline{570 \text{ mnok}}$$

= Forventet verdi av ikke perfekt informasjon = 30 mnok

Vi vil altså maksimalt være villig til å betale 30 millioner kroner for å få gjennomført en slik undersøkelse. Dette er verdien av ikke perfekt informasjon. Stegene i en analyse av den typen som er gjennomført over, kan oppsummeres i følgende punkter:

- Bestem hvilken handling som vil bli valgt med utgangspunkt i kun a priori sannsynligheter, og finn forventet avkastning av å velge gjennomføre denne handlingen
- Bestem mulige indikasjoner som ny informasjon kan gi

For hver mulig indikasjon

- Bestem sannsynligheten for at de ulike indikasjoner vil inntreffe

- Bruk Bayes' teorem for å revidere sannsynligheter i lys av disse indikasjonene
- Bestem optimal handling i lys av disse indikasjonene, dvs. ved å bruke a posteriori sannsynligheter og forventet verdi ved disse utfallene
- Multipliser sannsynligheten for at den enkelte indikasjon skal inntreffe med forventet verdi av den optimale handling gitt at indikasjonen inntreffer, og adder de resulterende verdier. Dette vil gi forventet avkastning ved ikke perfekt informasjon.
- Forventet verdi av ikke perfekt informasjon er lik forventet avkastning ved ikke perfekt informasjon minus forventet avkastning basert på den handling som er optimal gitt a priori sannsynlighetene.

Så lenge vi lenger til grunn at ny informasjon kan påvirke vår beslutning, har informasjon altså verdi. I neste del vil jeg se nærmere på hvordan ny informasjon, og påliteligheten av denne kan påvirke våre opprinnelige estimater.

## 13.5 Konsekvenser av ny informasjon før prosjektet iverksettes

### Bayesiansk tilnærming til ny informasjon

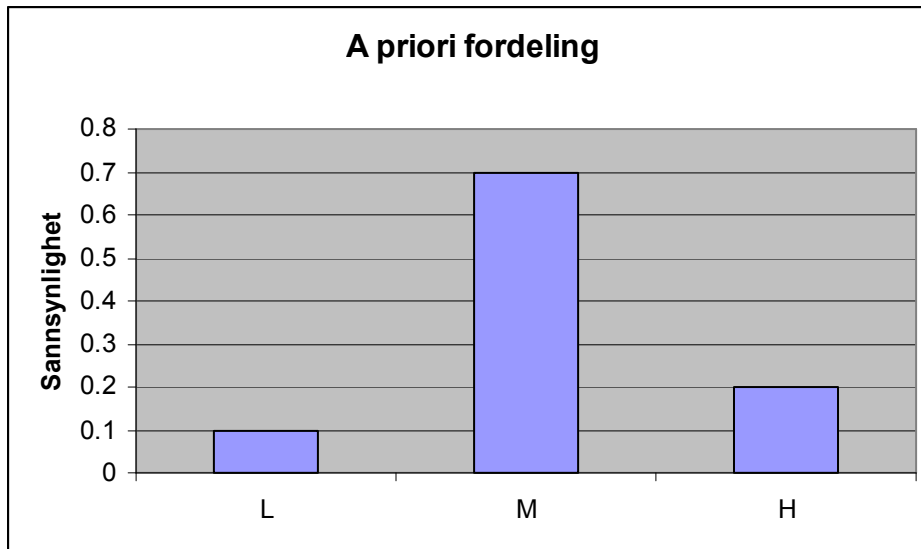
La oss ta utgangspunkt i et broprosjekt, der det er usikkerhet omkring trafikkgrunnlaget.<sup>43</sup> Anta at vi ut fra egne ekspertvurderinger og grundige analyser basert på tilsvarende prosjekter har kommet til at det er 70% sannsynlig at trafikken vil være på estimat M, 20% sannsynlig at trafikken vil være høyere (H), og 10% sannsynlig at den vil være lavere enn estimatet (L).

Så presenterer et eksternt rådgivningsfirma en prognose basert på oppdaterte trafikk tall som tilsier at trafikken vil ligge på det høye estimatet H, dvs. at prognosen i vårt opprinnelige estimat er for lav.

Vårt opprinnelige estimat var basert representert ved en apriorifordeling. Den nye fordelingen, basert på Bayes' teorem og oppdatert og ny informasjon, er aposteriori fordelingen. Apriori fordelingen er vist under:

---

<sup>43</sup> Også dette eksempelet er tilpasset fra Goodwin og Wright (2004).



Figur 13-8 A priori fordeling

Basert på en analyse av rådgivningsfirmaets tidligere treffsikkerhet i sine prognoser, kan vi sette opp følgende sannsynligheter:

$$p(\text{høy trafikkprognose gitt at trafikken vil bli høy}) = 0,9$$

$$p(\text{høy trafikkprognose gitt at trafikken vil være medium}) = 0,6$$

$$p(\text{høy trafikkprognose gitt at trafikken vil være lav}) = 0,3$$

Når rådgivningsfirmaet sier at trafikken blir høy, er det altså 90% sannsynlig at de har rett. Spørsmålet er så hva vår oppdaterte sannsynlighetsfordeling blir, gitt den nye informasjonen. De sammensatte sannsynlighetene blir følgende:

$$p(\text{Høy trafikk og høy trafikkprognose}) = 0,2 * 0,9 = 0,18$$

$$p(\text{Medium trafikk og høy trafikkprognose}) = 0,7 * 0,6 = 0,42$$

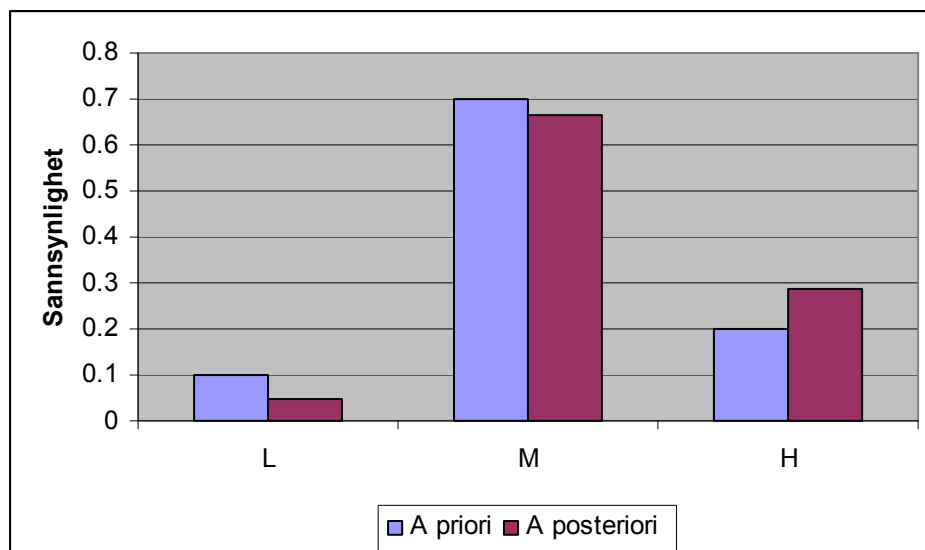
$$p(\text{Lav trafikk og høy trafikkprognose}) = 0,1 * 0,3 = 0,03$$

Summen av de sammensatte sannsynlighetene blir 0,63. Vi kan da finne *a posteriori* fordelingen basert på Bayes' teorem:

$$p(\text{Høy trafikk}) = 0,18 / 0,63 = 0,29$$

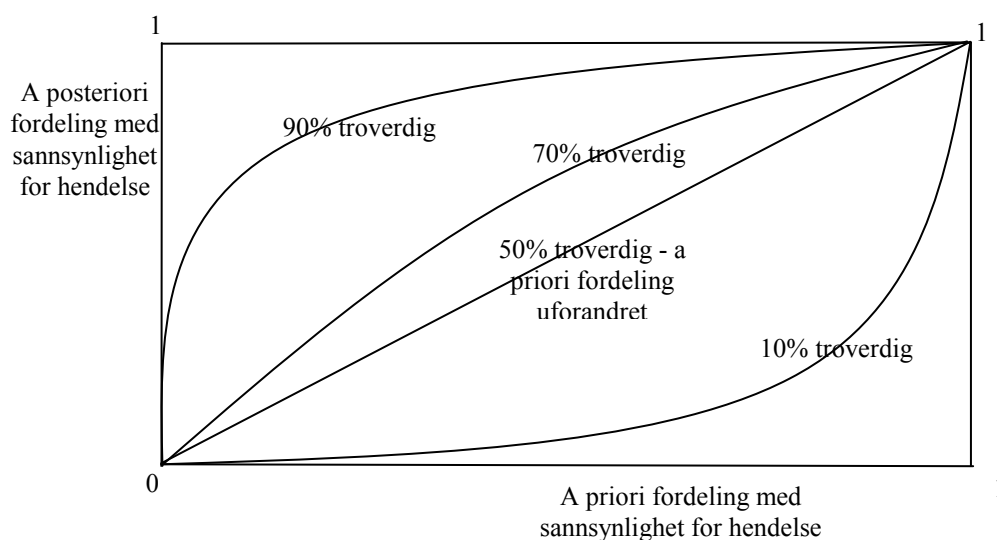
$$p(\text{Medium trafikk}) = 0,42 / 0,63 = 0,67$$

$$p(\text{Lav trafikk}) = 0,03 / 0,63 = 0,05$$



Figur 13-9 A priori fordeling sammenholdt med a posteriorifordeling

Vi ser altså at den nye trafikkprognosen, som pekte i retning av det høye trafikkestimatet, har medført at dette har fått økt sin sannsynlighet, mens de to andre har fått redusert sin. Dette avhenger selvsagt av hvor troverdig den nye informasjonen er. Dette kan illustreres ved følgende figur, der vi nå baserer oss på et enklere eksempel, der det kun er en sannsynlighet for en hendelse inntreffer eller ikke, mens den nye informasjonen gir ny informasjon om dette med ulik grad av troverdighet.



Figur 13-10 A posteriori-fordeling ved ulik grad av troverdig informasjon

Svært troverdig ny informasjon vil altså føre til en kraftig oppjustering av vårt opprinnelige estimat. Det er også interessant å merke seg betydningen av å aldri operere med sannsynligheter i hjørnepunktene. For eksempel: Hvis vi sier at det er 100% sikkert at en hendelse ikke vil forekomme, vil vi måtte opprettholde dette estimatet uansett hva slags ny informasjon som måtte komme til. Dette er helt tydelig irrasjonelt (men kanskje ikke uvanlig).

Imidlertid vil det også påløpe informasjon *etter* at prosjektet er igangsatt og ferdig. Spørsmålet er om dette vil kunne påvirke beslutningen vår. I så fall blir utfordringen hvordan vi kan ta hensyn til dette i et beslutningsrammeverk.

### 13.6 Konsekvenser av ny informasjon som først avdekkes *etter* at prosjektet er iverksatt

Svært ofte vil informasjon som påvirker lønnsomheten av prosjektet først bli avdekket *etter* at prosjektet er iverksatt. Hva er verdien av slik informasjon for beslutningen som må tas nå? Realopsjonsmetodikk er en analytisk tilnærming for prosjektvurdering som er særlig nyttig i slike sammenhenger der slik informasjon potensielt kan påvirke investeringsbeslutningen. Den vil være særlig nyttig i en tidlig fase, der det beslutningsutfordringen nettopp er å få et godt nok grunnlag for å velge mellom *alternative* konsepter.

Konsekvensen av at informasjon først avdekkes *etter* at prosjektet er iverksatt, er for eksempel en potensiell reversering av rangeringen av prosjektalternativer med og uten fleksibilitet i forhold til den rangering som følger av tradisjonelle tilnærminger. Vi kan også tenke oss at usikkerhetsavsløring kan påvirke tidspunktet for iverksetting av prosjektet. Men før vi går gjennom et enkelt regneeksempel for å illustrere dette, vil vi se på litt av den teoretiske bakgrunnen for at fleksibilitet har en tilleggsverdi. Realopsjoner er altså et sentralt begrep her, men vi vil først se på svakheter ved den tradisjonelle tilnærmingen i slike sammenhenger.

#### Problemer med tradisjonell investeringsvurdering

Ved tradisjonell investeringsanalyse utledes punkttestimater for hver periodes (forventede) kontantstrøm med utgangspunkt i ofte implisitte forutsetninger om for eksempel investeringsprofil, innteks- og kostnadsstruktur. Verdiberegninger skjer gjerne ved bruk av en vilkårlig risikojustert rentesats, som holdes konstant over tiden og er den samme for alle kontantstrømelementene. I sjeldnere tilfeller benyttes sikkerhetsekvivalente verdier. Det er flere problematiske sider ved den tradisjonelle, rett frem, tilnærmingen. Både selve modelleringen av den stokastiske kontantstrømmen og korrigeringen for risiko ved verdsettingen gjør at slike lønnsomhetsanalyser lett kan ha tvilsom verdi. Det er flere kjennetegn ved store offentlige investeringsprosjekter som bidrar til dette.

**Usikkerhet.** De fleste elementene som inngår i en lønnsomhetsanalyse av et stort offentlig investeringsprosjekt er usikre. I en tidlig fase vil det mulige utfallsrommet vært enda større enn rett før selve investeringsbeslutningen fattes. På analysetidspunktet kjenner en vanligvis ikke de fremtidige tallverdier kostnader, eventuelle inntekter eller nytteverdi. I tillegg kommer at heller ikke fremtidig tilgjengelig teknologi, konkurransestruktur og politiske rammebetingelser er kjent med sikkerhet.

**Sekvensiell informasjon.** Ny informasjon blir tilgjengelig ettersom tiden går. Dette kan påvirke vurderingen av prosjektet på flere måter. Selv ved uendret strategi justeres punkttestimatene etter konstaterte avvik fra implisitte forutsetninger som lå til grunn for prosjektbeskrivelsen. Vel så viktig er det om ny informasjon kan gi grunnlag for å endre beslutninger.

**Irreversible investeringer.** I motsetning til finansinvesteringer er gjerne realinvesteringer, ikke minst store offentlige investeringsprosjekter, i begrenset grad reverserbare. Selv om

eksterne forhold ikke skulle ha endret seg vesentlig, kan en ikke uten vesentlige kostnader komme tilbake til utgangssituasjonen. Investeringskostnaden har et betydelig innslag av ”sunk cost”, da alternativverdien av det fysiske anlegget er lavere enn investeringskostnaden. Det vil også være slik at investeringer i ekstra kapasitet i liten grad kan hentes inn igjen, om kapasitetsbehovet ikke skulle oppstå.

**Fleksibilitet.** Utnytting av ny informasjon og tilpassning til endrede rammevilkår er viktige forutsetninger for god økonomisk styring. Med fleksibilitet styrkes oppsidepotensialet kombinert med redusert fallhøyde i tilfelle ugunstig utvikling, i forhold til en passiv strategi basert på de initiale forventninger reflektert i opprinnelig prosjektbeskrivelse. Denne fleksibiliteten kan betraktes som innebygde opsjoner, og slike opsjoner har en verdi som lett overses ved tradisjonelle analyser. Praktikere fremhever da gjerne også fleksibilitet som et viktig positivt trekk ved vurdering av beslutningsalternativer. God utnytting av fleksibilitet kan gi betydelige økonomiske merverdier. Flexibilitet eksisterer i ulike faser av prosjektet, både før, under gjennomføring og i driftsfasen. Eksempler på nyttig fleksibilitet er mulighet for å utsette, endre skala eller eventuelt også forkaste et prosjekt. For et prosjekt i driftsfasen kan kontantstrømmer, eller nytteverdien også endres på mange måter, for eksempel ved å tilpasse kapasiteten i lys av ny informasjon. I forhold til politiske eller økonomiske rammebetingelser, for eksempel knyttet til næringsstruktur, miljøpolitikk og naturressurser kan nettopp fleksibilitetsaspekter være av meget stor betydning.

Verdien av slik fleksibilitet kommer imidlertid sjelden skikkelig frem i tradisjonell investeringsanalyse. Tradisjonelle nåverdianalyser fanger ikke opp den ekstraverdien som fleksibilitet skaper. Beslutningstrær fokuserer riktignok på sekvensielle beslutninger, men korrekt bruk er også her en utfordring, og tilnærmingen kan lett være misvisende.

Det er viktig å huske på at den tradisjonelle nåverdiregelen gjelder for uavhengige investeringer. Dersom utsetting er mulig, så kan ventealternativet oppfattes som et eget prosjekt, som er gjensidig utelukkende i forhold til straksalternativet. Gjennomføring straks må derfor ikke bare ha en positiv nåverdi, men denne må være tilstrekkelig stor til også å dekke alternativkostnaden ved at ventealternativet ikke lenger blir tilgjengelig ved et engangs, irreversibelt prosjekt. Videre vil fleksibilitet til å foreta beslutninger *etter* at ny informasjon er blitt tilgjengelig, være av verdi hvis den nye informasjonen kan endre beslutning som ellers ville blitt tatt. Usikkerhet kombinert med beslutningsfleksibilitet, gjør derfor et prosjekt mer verdifullt enn det ville være under sikkerhet eller ved mindre usikkerhet. Dette er velkjent fra finansielle opsjoner.

En utfordring blir dermed å benytte en beregningsmetode som på en riktig måte fanger opp verdien av fleksibilitet. Realopsjonstilnærmingen er et skreddersydd verktøy for analyser av komplekse situasjoner hvor usikkerhet, et, sekvensiell informasjon og fleksibilitet er fremtredende.

### **Realopsjoner – hva er det?**

Realopsjonstilnærmingen er et skreddersydd verktøy for analyser av komplekse situasjoner hvor usikkerhet, sekvensiell informasjon og fleksibilitet er fremtredende. Tilnærmingen hviler tungt på metode for verdsetting av finansielle opsjoner.

Opsjoner er en type verdipapirer, ofte kalt derivater, hvor verdien er knyttet til verdien av et underliggende finansobjekt, typisk en aksje. Det finnes flere typer finansielle opsjoner, og en



god oversikt finnes i Hull (2005). Vi skiller gjerne mellom en kjøpsopsjon og en salgsopsjon. En kjøpsopsjon gir eieren rett (men ikke en plikt) til å kjøpe det underliggende verdipapiret til en gitt kontraktspris. En salgsopsjon er tilsvarende en kontrakt som gir innehaver rett (men ikke plikt) til å selge en bestemt verdipapir til en avtalt kontraktspris. Finansielle opsjoner er begrenset til en kontraktsperiode; avtalen kan for eksempel være at en har en opsjon på å kjøpe aksjen til kontraktsprisen i løpet av de neste tre månedene (amerikansk opsjon), eller bare ved utløp på en bestemt dato og klokkeslett (europaisk opsjon).

De fem viktigste variablene som påvirker prisen på en finansiell opsjon kan også benyttes ved verdsetting av en realopsjon. Analogien er vist i tabellen under:

*Tabell 13-1. Analogi mellom finansiell opsjon og realopsjon.*

Finansiell kjøpsopsjon	Realopsjon
Aksjepris	Nåverdi av kontantstrøm
Utøvningspris	Investeringskostnad
Tid til utløp	Tid til investeringsmuligheten utløper
Usikkerhet i aksjepris	Usikkerhet i prosjektverdi
Risikofri rente	Risikofri rente

De vanligste realopsjonsaspekter som har blitt søkt fangetopp gjennom ulike analytiske tilnærminger er:

Opsjonen til å *vente* med å investere

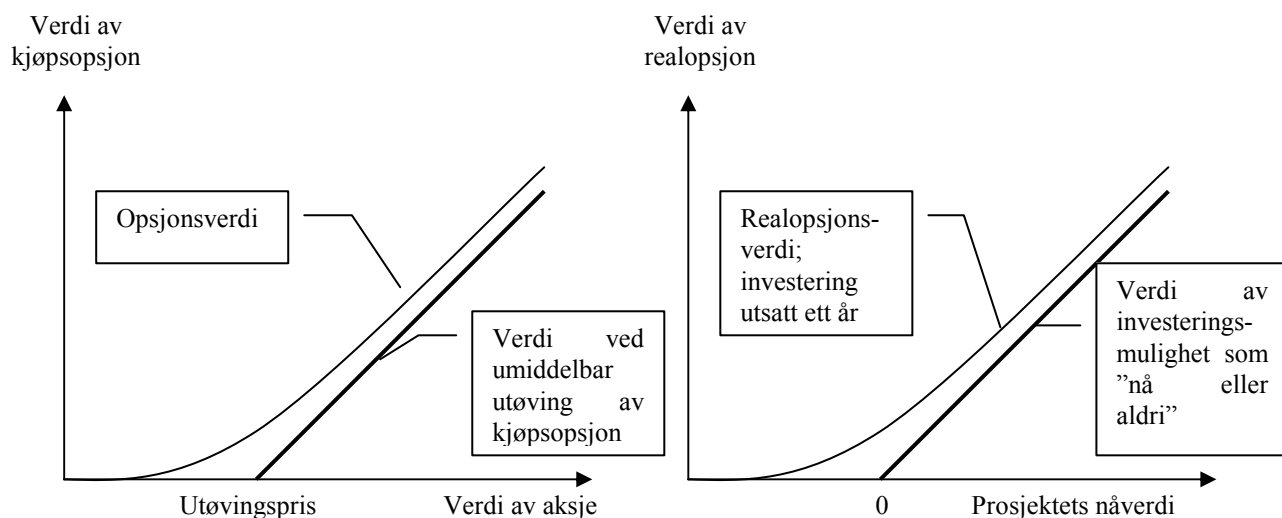
Opsjonen til å *utvide* hvis den initiale investeringen er vellykket

Opsjonen til å *redusere eller avvikle* prosjektet

Opsjonen til å *endre produksjonssammensetning* eller –metoder

De tre siste opsjonsaspektene kan vi kanskje slå sammen til verdien av fleksibilitet, selv om det selvsagt vil kreve separat beregnings- og beregningsmetodikk.

Analogien mellom en realopsjon og en finansiell opsjon kan illustreres ytterligere med utgangspunkt i det første realopsjonsaspektet. Dette er gjort i figuren under:



Figur 13-11 Analogi mellom finansiell kjøpsopsjon og en investeringsmulighet som kan utsettes i påvente av ny informasjon

Hvis vi antar at vi har en rett til å kjøpe en aksje til en bestemt pris på et bestemt tidspunkt, vil denne retten på utøvingstidspunktet være verdt noe dersom aksjen på det tidspunktet er verdt mer enn det vi har rett til å kjøpe den for. Den tykke linjen i figuren til venstre viser verdien av denne retten på utøvingstidspunktet, og vi ser at desto mer aksjen er verdt på utøvingstidspunktet, desto mer er opsjonen verdt. Er derimot aksjekursen lavere enn det vi har en rett til å kjøpe den for, er opsjonen verdiløs. En finansiell kjøpsopsjon blir imidlertid kjøpt og omsatt i perioden før opsjonen utløper. Utfordringen er å verdsette opsjonen gitt at aksjeprisen både kan gå opp og ned i tiden før utløp. Dette gjør at risiko eller aksjens variabilitet blir en viktig variabel i verdsettingen. Tidsaspektet gjør at også tid til utløp og rente blir en variabel i verdsettingen. Utøvningspris og aksjepris blir selvfølgelig også sentrale elementer i verdsetting, som vist i tabellen over. Selv om aksjens pris i dag ligger under utøvningsprisen er det alltid en mulighet for at den vil stige frem til utløp av opsjonen. Samlet sett innebærer dette at kjøpsopsjonsverdien frem til utløp alltid vil ha en positiv verdi, dvs. at den tynne buede linjen alltid vil ligge over verdien av opsjonen ved utløp.

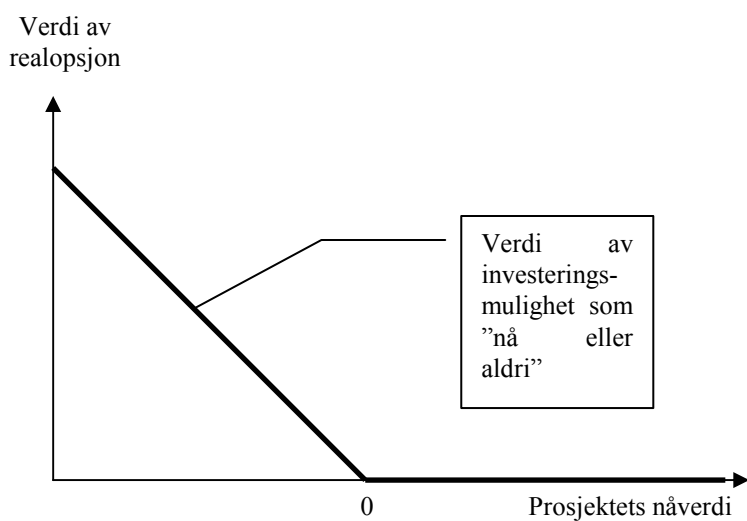
Tilsvarende vil verdien av investeringen betraktet som en opsjon, for eksempel at vi kan investere nå eller utsette investeringen til neste år for å avdekke usikkerhet i en variabel av betydning for investeringsbeslutningen, alltid ligge over verdien av å investere "nå eller aldri". Analogien er en finansiell kjøpsopsjon, der vi holder muligheten åpen for å investere. Usikkerheten preger investeringsbeslutningen, men ved å vente vil vi også utsette tidspunktet for når inntektene fra prosjektet påløper.

Investeringskostnaden (utøvningskostnaden) kan for eksempel være 1 mrd, mens nåverdien er like stor. Dette gir en netto nåverdi på null, som illustrert i figuren til høyre. Stiger nåverdien ut over 1 mrd. vil netto nåverdi stige tilsvarende. Dette er illustrert ved den tykke linjen i figuren til høyre. Hvis investeringen kan utsettes et år, er den en mulighet for at inntektene vil vise seg å være høyere, eventuelt kostnadene lavere enn det som lå til grunn til "nå eller aldri" verdianslaget. Dette vet vi først om et år, da denne usikkerheten blir avslørt. Da står vi overfor samme beslutning, å investere nå eller vente. På samme måte som en finansiell opsjon gir

denne muligheten til å avvete beslutningen prosjektet en ekstra verdi. Dette er illustrert ved den buede linjen i figuren til høyre. Prosjektet "invester nå" må sammenlignes med alternativprosjektet "vent og se". Dersom verdien av å vente er høyest, skal vi vente. Valget er basert på å velge det som gir størst verdi i:

$\text{Max}[\text{Verdi av investering nå; Verdi av å vente}]$ .

Utfordringen er selvsagt verdsettingsmetodikken. I eksempelet over, var realopsjonen sammenlignet med en finansiell "call opsjon", dvs. en rett til å kjøpe. Vi kan også tenke oss realopsjoner som best kan sammenlignes med en rett til å selge, dvs. en "put opsjon". Da blir avkastningsstrukturen som vist i figuren under.



Figur 13-12 En realopsjon med avkastningsstruktur som en put opsjon.

Videre kan vi, på samme måte som for finansielle opsjoner, tenke oss sammensatte realopsjoner. I tabellen under er det vist egenskaper ved de tre typene generiske realopsjoner.

Tabell 13-2. Egenskaper ved tre hovedtyper generiske realopsjoner.

Klasse	Egenskaper
Call type realopsjon	<p>Fanger opp oppsidepotensialet i prosjektverdi knyttet til underliggende usikker variabel</p> <p>Utøving innebærer typisk utlegg i form av investeringskostnad</p> <p>Utøves når forventningene om positiv avkastning øker</p> <p>Eksempel: ”invester nå eller vent”</p>
Put type realopsjon	<p>Forsikring mot tap knyttet til reduksjon i prosjektverdi på grunn av usikker variabel</p> <p>Utøving kan innebære utlegg i form av nedstengingskostnader eller lignende</p> <p>Utøves når forventningene om tap øker</p> <p>Eksempel: ”legg ned produksjonen nå eller vent”</p>
Sammensatt opsjon	<p>Et prosjekt kan innebære flere opsjoner</p> <p>Utøving basert på overordnet profittmaksimering</p> <p>Eksempel: ”Investering i oljeleting er en opsjon som gir en opsjon på feltutbygging som gir en opsjon på...”</p>

Når det gjelder fleksibilitet er det mest korrekt å betrakte prosjektet som en sammensatt opsjon, dvs. først en mulighet til å investere (realopsjon 1; verdi av å vente) så en mulighet til å tilpasse seg til ny informasjon (realopsjon 2; verdi av fleksibilitet). Det er viktig å huske på at den andre realopsjonen, verdien av fleksibilitet, også vil være til stede ved ”invester nå” alternativet, slik at vi foretar den rette sammenligningen mellom prosjektene. Det vil også være en kostnad knyttet til å utøve fleksibilitetsopsjonen, og valget er basert på å velge det som gir høyest verdi i uttrykket;

Max[status quo; ekspander prosjekt].

Vi kan tenke oss andre fleksibilitetsopsjoner, for eksempel steng produksjonen, der valget er Max[fortsett produksjonen; steng ned].

### Eksempel på verdsetting av fleksibilitet

For verdsettingsmetodikk er det etter hvert en omfattende litteratur. Brekke (2004) gir en god innføring med illustrerende eksempler. Det samme gir Hull (2005). Ellers er Dixit og Pindyck (1994) en standardreferanse med et sterkere teoretisk tilsnitt. Her vil vi forsøke å illustrere verdsettingsmetodikken med et sterkt forenklet ikke-teknisk eksempel. Den interesserte leser henvises for øvrig til nevnte referanser.

Som et eksempel for å illustrere realopsjonsmetodikk vil vi ta utgangspunkt i et FoU-prosjekt. Dette krever en grunnlagsinvestering på 6 mill. kr., og det er en 10% sjanse for at det lykkes. Dersom det lykkes kreves en tilleggsinvestering på 100 mill., men dette åpner for en sikker avkastning på 150 mill. kr. Lønnsomheten for dette prosjektet tradisjonelt vurdert er:

$$10\%*(-100+150) - 6 = -1 \text{ mill. kr.}$$

Dette FoU-prosjektet virker altså som en dårlig idé. Realopsjonstilnærmingen tar derimot utgangspunkt i at det dette prosjektet gir en *mulighet* til å investere i oppfølging – men det er *ingen forpliktelse* til å gjøre dette. Vi har en mulighet til å investere 100 mill. kr. hvis det viser seg at prosjekter en vellykket, men ingen forpliktelse.

La oss anta at markedet for det nye produktet kan vise seg å være på hele 250 mill. kr. dersom ingen konkurrenter lykkes med å utvikle konkurrerende alternativer. Men det kan også vise seg at de gjør det; i så fall er markedet bare verdt bare 50 mill. kr. Det anslås at det er en 50/50 sjanse for at konkurrentene lykkes, og forventet markedsverdi er på 150 mill. kr., det samme som den sikre markedsverdien i regneeksempelet over.

Dersom det ved avslutningen av FoU-prosjektet viser seg at konkurrentene har lykkes, har vi en beslutningssituasjon med et utlegg på 100 mill. kr. for en gevinst på bare 50 mill. kr., dvs. det som er verdien av produktet i markedet i dette tilfellet. Dette er et klart ulønnsomt foretagende, og prosjektet vil bli lagt ned, og man begrenser tapet til den initiale investeringen på 10 mill. kroner.

Dersom det viser seg at konkurrentene ikke har lykkes, har vi et marked på 250 mill. kr. som kan realiseres for en initiale investering på 100 mill. kr. Dette er klart lønnsomt. Siden det er 50% sjanse for at konkurrentene ikke lykkes, har vi en forventet verdi på 75 mill. kr. i dette tilfellet. Siden det er 10% sannsynlighet for at prosjektet lykkes, har vi en forventet verdi av FoU-prosjektet på:

$$10\%*(75) - 6 = 1,5 \text{ mill. kr.}$$

FoU-prosjektet er altså lønnsomt. Det paradoksale er at å introdusere usikkerhet altså har bidratt til at prosjektets lønnsomhet har økt. Ved tradisjonell nåverdianalyse vil økt usikkerhet bidra til redusert prosjektverdi, siden økt usikkerhet typisk reflekteres i avkastningskravet. Her er det motsatt. Usikkerhet er en del av virkeligheten, og realopsjonstilnærmingen gir en mer realistisk tilnærming til modellering av usikkerhet – som jo ofte avsløres over tid, og svært ofte etter at prosjektet er igangsatt eller også helt ferdig.

Selv om realopsjonstilnærmingen etter hvert er vel etablert på en rekke ulike områder og i ulike typer anvendelser, skal det ikke legges skjul på at det også er en rekke forskjellige teoretiske, empiriske og konseptuelle utfordringer som må løses for at den skal kunne benyttes for å gi kvantitative verdiesestimater i alle praktiske beslutningssituasjoner. Ikke desto mindre er

det et viktig redskap for å vise strukturen i beslutningsproblemet, og hva som tilfører verdi til de ulike prosjektoalternativer; fra å investere nå eller senere til å velge en litt dyrere løsning men som også er mer fleksibel med hensyn på fremtidige usikre utfall.

## 13.7 Referanser

Brekke, K. A. (2004): *Realopsjoner og fleksibilitet i store offentlige investeringsprosjekter*. Concept, Rapport: 8/2004.

Dixit, A. K., og R. S. Pindyck (1994): *Investment under Uncertainty*.: Princeton University Press.

Goodwin, P., og G. Wright (2004): *Decision analysis for management judgment*. Chishester: John Wiley & Sons.

Hull, J. C. (2005): *Options, futures and other derivatives*.: Prentice Hall.

## 14 FRA ANALYSE TIL BESLUTNING - SAMFUNNSØKONOMISK LØNNSOMHETSVURDERING

KÅRE P. HAGEN\*

### 14.1 Innledning

Lønnsomhet er et mål på effektiviteten i utnytting av knappe ressurser. Effektivitetsmålet avhenger av virksomhetens formål. Dersom virksomheten er en privat eid bedrift, vil som regel formålet med virksomheten være å maksimere dens bidrag til eierne inntekt og formue på hhv kort og lang sikt. Det periodiske overskuddet er gitt ved det som eierne kan ta ut av bedriften uten at det forringer nettoverdien av bedriftens markedsverdi. Markedsverdien vil være gitt ved nåverdien av den inntektsstrømmen som bedriften genererer for eierne. Positiv nåverdi betyr at virksomheten gir en gjennomsnittlig kapitalavkastning som er minst like stor som avkastningskravet, som vanligvis settes lik det eierne ville oppnådd ved å plassere de midler som er bundet i virksomheten i det eksterne marked. Lønnsomheten defineres med andre ord i forhold til en ekstern alternativavkastning for den kapitalen som bedriften legger beslag på. Alternativavkastningen blir dermed den benchmark som lønnsomheten i bedriften måles opp mot. Vanligvis vil avkastningen i finansmarkedet bli lagt til grunn, eventuelt korrigert for den risiko som eierne bærer gjennom bedriften. Alternativavkastningen blir da det som det koster eierne å binde kapital i virksomheten.

Alternativkostnadstankegang er et sentralt begrep i all lønnsomhetsanalyse. Det gjelder ikke bare i forhold til kapital, men for alle de faktorer som er nødvendige for produksjon i en virksomhet eller for gjennomføring av et bestemt prosjekt. Knapphetsbegrepet bygger på at bruk av ressurser til et bestemt formål fører til et økonomisk offer som representeres ved alternativkostnaden. For en privat bedrift er det økonomiske offeret det som bedriften må betale i markedet for arbeidskraft og nødvendig vareinnsats. Alternativkostnaden for produksjonsfaktorer som kjøpes i markedet er dermed gitt ved markedsprisene.

Inntektssiden er de kontantinntektene som bedriften genererer for eierne. Når differansen mellom inntekter og alternativkostnader er positiv, er bedriften lønnsom i bedriftsøkonomisk forstand. Alternativkostnadstankegang innebærer at selv om differansen mellom løpende inntekter og betalbare kostnader er positiv, så betyr ikke det nødvendigvis at bedriften er lønnsom i den forstand at det er den mest lønnsomme måten for eierne å bruke ressursene på. Det avhenger av hva som er relevant alternativavkastning. Dersom flytting av bedriften til et annet sted, f.eks. et annet land, fører til en høyere kapitalavkastning, er det denne benchmarken som lønnsomheten må måles opp mot. Bedriftens regnskap bygger på faktiske kostnader som kan være forskjellige fra alternativkostnader. I forhold til alternativkostnadene kan bedriften være ulønnsom selv om den går med et regnskapsmessig overskudd.

---

\* Kåre P. Hagen er professor ved Institutt for samfunnsøkonomi, NHH

Samfunnsøkonomisk lønnsomhet defineres i prinsippet på samme måte som bedriftsøkonomisk lønnsomhet. Forskjellen er likevel at det er flere "stakeholdere" knyttet til virksomhetens samfunnsøkonomiske overskudd. I prinsippet er det alle grupper som berøres av virksomheten i ulike roller og på forskjellige arenaer. Utover eierne er de viktigste stakeholderne konsumenter, arbeidere og skattebetalere. Litt spissformulert kan en si at kunder er viktige for den bedriftsøkonomiske lønnsomheten i den utstrekning de bidrar til bedriftens inntekter på kort eller lang sikt. Fra et samfunnsøkonomisk perspektiv er "kundernes" bidrag til lønnsomheten gitt ved den samfunnsøkonomiske verdien av konsumet av de varene som bedriften produserer. Slik sett kan en si at for bedrifter som produserer kun for eksport, er den bedrifts- og samfunnsøkonomiske vurderingen av inntektene sammenfallende da det som tilflyter norske borgere, er eksportinntektene siden konsumet foregår i andre land.

## **14.2 Samfunnsøkonomiske alternativkostnader: Forskjeller mellom det bedrifts- og samfunnsøkonomiske perspektivet.**

Den bedrifts- og samfunnsøkonomiske lønnsomheten vil generelt være ulik på grunn av at perspektivet for den samfunnsøkonomiske lønnsomhetsvurderingen vil være mer omfattende enn det som ligger til grunn for den bedriftsøkonomiske. Når perspektivet er forskjellig, kan også det økonomiske offeret være ulikt og dermed vil alternativkostnaden bli forskjellig. Det samfunnsøkonomiske offeret er det landet som nasjon ofrer ved å gjennomføre et prosjekt eller tiltak. Av flere grunner vil den samfunnsøkonomiske alternativkostnaden kunne være forskjellig fra den bedriftsøkonomiske markedsbaserte kostnaden.

Om vi tar utgangspunkt i bedriftsperspektivet, er det bedriftsøkonomiske offeret for vareinnsats som kjøpes i markedet åpenbart gitt ved de betalbare kostnadene. Normalt skaffes ressurser til veie også ved interne kryssleveranser mellom ulike avdelinger i bedriften. Spørsmålet er da hvordan produksjonsfaktorer og tjenester som er produsert internt i bedriften skal verdsettes i bedriftens lønnsomhetskalkyler. Dersom det f.eks. er en reparasjonsavdeling for bedriftens maskinpark, synes det naturlig å legge de betalbare kostnadene til grunn. Men om det eksisterer et eksternt marked for slike tjenester, er det en eksternt benchmark for verdsetting. Om bedriften står fritt når det gjelder valget mellom å ha en egen reparasjonsavdeling eller å kjøpe tjenestene i markedet, er alternativverdien gitt ved prisene i det eksterne markedet. Dersom markedsprisen er lavere enn kostnaden ved egenproduksjon, vil bedriften tjene å legge ned avdelingen og kjøpe tjenestene utenfra. I så fall er alternativkostnaden gitt ved markedsprisene. Om kostnadene ved egenproduksjon er lavere, vil en tjene å selge slike tjenester ut av bedriften. Kostnadene ved å bruke reparasjonsavdelingens tjenester internt blir da de inntektene en ofrer i det eksterne markedet. Offersynspunktet leder her entydig til at alternativkostnaden er den eksterne markedsprisen for sammenlignbare tjenester.

Den siste kategorien er tjenester som er så spesialiserte, at de ikke kan skaffes til veie eksternt. I dette tilfelle må en foreta en eksplisitt kalkyle for å anslå hva tjenesten faktisk koster. Offersynspunktet tilsier at en må ta utgangspunkt i alternativverdien som den nødvendige ressursinnsatsen har i alternativ intern virksomhet eller i det eksterne markedet. Spørsmålet om riktig kalkylepris oppstår dermed i det tilfellet at det ikke eksisterer noen eksternt markedsbestemt benchmark å forholde seg til.



Om vi legger den samme tankegang til grunn for en samfunnsøkonomisk verdsetting av ressursbruk for et land, er det klart at det internasjonale markedet blir det eksterne markedet som et land må forholde seg til. Det følger da uten videre at importprisen<sup>44</sup> blir alternativkostnaden for importert vareinnsats. Importkostnaden er den kjøpekraft som landet ofrer i det internasjonale markedet for vareimporten. På samme måten blir eksportinntektene ved vare- og tjenesteeksport den samfunnsøkonomiske inntekten ved at det er bidraget til landets kjøpekraft i det internasjonale markedet.

De fleste land har imidlertid også en importkonkurrerende sektor slik at alternativet til import er kjøp innenlands. Ut fra samme tankegang som ved bedriftens situasjon når kjøp i markedet er en mulighet, vil markedsprisen i det internasjonale markedet være det egentlig økonomiske offer for slik ressursinnsats for et bestemt innenlandsk formål. Dersom f.eks. forsvaret bygger båter i Norge til tross for at likeverdige båter kunne ha blitt bygget i utlandet til en pris som er lavere enn kostnadene hjemme, må det bety at produksjon hjemme må gi tilleggstjenester for landet i en eller annen form. Spørsmålet blir da om verdien av disse tilleggstjenestene forsvarer kostnadsdifferansen. Den potensielle importprisen blir da uansett det økonomiske offeret for anskaffelsen, mens tilleggskostnaden ved bygging innenlands blir kostnaden for tilleggstjenestene.

Det nasjonale scenariet som korresponderer med verdsetting av kostnader for spesialiserte bedriftsinterne tjenester, er det tilfellet hvor nasjonal virksomhet benytter ressurser fra skjermede innenlandske sektorer. Her er kjøp fra utlandet ikke et relevant alternativ, slik at den samfunnsøkonomiske alternativkostnaden må beregnes uten støtte i verdensmarkedspriser.

Det følger av dette at jo mer åpen økonomien er, desto enklere blir den samfunnsøkonomiske lønnsomhetskalkylen i den forstand at den kan baseres på markedspriser. Der er imidlertid en del andre forhold som kompliserer sammenhengen mellom markedspriser og samfunnsøkonomiske alternativverdier som vi diskuterer nærmere nedenfor. De knytter seg til tre hovedkategorier. Den første kategorien er knappe goder og ressurser som ikke blir handlet i markeder og som det derfor heller ikke fins markedspriser for. Den andre kategorien gjelder situasjoner der markedspriser eksisterer men gir feilaktige signaler om samfunnsøkonomiske verdier og alternativkostnader. Den tredje er skatter og avgifter som foruten å gi det offentlige en andel av overskuddet i all beskattet virksomhet, også fører til at markedsprisene blir forskjellige for ulike aktører i markedet slik at det ikke uten videre er klart hvilken pris som reflekterer den samfunnsøkonomiske alternativkostnaden.

### 14.3 Nærmere om forholdet mellom bedrifts- og samfunnsøkonomisk lønnsomhet

#### Manglende markeder

Dersom markedsbestemt lønnsomhet skal være retningsgivende for samfunnsøkonomisk lønnsomhet, fordrer det at alle knappe goder og ressurser blir prissatt i velfungerende markedet slik at samfunnsmessig nytte og kostnader blir ivaretatt i markedsbaserte lønnsomhetskalkyler. Den viktigste årsak til manglende markeder gjelder goder som det er

<sup>44</sup> Det forutsettes her at landets import ikke påvirker verdensmarkedsprisen.

vanskelig å håndheve eksklusiv eiendomsrett til, eller individualisere bruken av. Dersom håndheving av eksklusiv eiendomsrett er vanskelig, blir det vanskelig å innkassere betalingsvilligheten for slike goder, eller kreve kompensasjon for forhold som fører til forringelse av slike goder. Typiske eksempler på slike goder er de nasjonale fellesgodene forsvar, ordensvern og rettsvesen, der individualisering av tjenesteytingen er noe nær umulig. Andre eksempler er fyrlykter og oppmerkede seilingsleder. Selv om staten eier infrastrukturen, vil det være vanskelig å innkassere betaling for bruken. Et annet eksempel er opparbeidelse av friluftsområder og miljøtiltak som er fellestiltak rettet mot alle, og følgelig blir det vanskelig å individualisere bruk og verdsette nytte. Om vi ser utover nasjonale grenser, har vi globale fellesgoder som fiskeressurser i internasjonale farvann, ozonlaget, global oppvarming og klima. På disse områdene er håndheving av eksklusiv nasjonal eiendomsrett vanskelig og heller ikke ønskelig.

Når innkassering av brukernes nytte i form av markedsbasert inntekt er vanskelig, blir det også problematisk å anslå den samfunnsøkonomiske verdien av inntektssiden av slike tiltak da det ikke er markedspriser å støtte seg på. Verdsettingen må derfor anslås på annet vis. En vanlig metode er spørreundersøkelser blant dem som berøres av tiltaket, hvis de kan identifiseres og avgrenses. Det gjelder i særlig grad for miljøtiltak, som kan være investering i et bedre miljø, og for tiltak som forringer det eksisterende miljøet. Ved miljøforringelser tar spørreundersøkelsene sikte på å anslå de samfunnsøkonomiske kostnadene i form av i penger vurderte ulemper som slike virkninger har for alle dem som blir berørt.

Mange typer fellesgoder gir størst verdi dersom de brukes i sammenheng med private markedsgoder. Vi sier da at de er komplementære i bruken. Brukernes verdsetting kan i slike tilfelle i noen grad reflekteres i verdsettingen av komplementære markedsgoder. Fyr og oppmerkede seilingsleder gir størst nytte dersom det eksisterer sjøkart der de er avmerket. Betalingsvilligheten for de underliggende fellestjenestene vil da reflekteres i betalingsvilligheten for de markedsgodene som er nødvendig for å kunne nyttiggjøre seg fellesgodene. Tilsvarende vil verdien av investeringer i lokale miljøtiltak i boligstrøk bli reflektert i eiendomsverdiene i området. Investering i bedret lokal infrastruktur vil ha samme virkning og vil dessuten føre til bedre lønnsomhet og høyere markedsverdi for bedrifter som er avhengig av infrastrukturen. Her kunne en tenke seg at samfunnets betalingsvillighet for slike tiltak kan fanges opp ved en lokal eiendomsskatt.

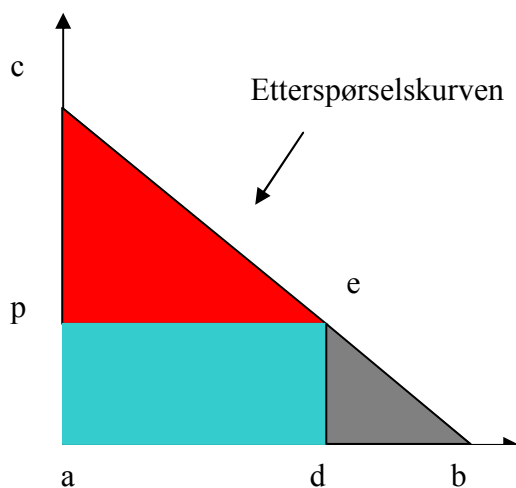
En annen grunn til at det ikke nødvendigvis vil være hensiktsmessig å innkassere brukernes betalingsvillighet selv om det er praktisk mulig, er at bruken av godet ikke påfører samfunnet nevneverdige kostnader. Det gjelder i særlig grad dersom kostnadene først og fremst knytter seg til etablering av den fasiliteten som ligger til grunn for tjenesten. Når den først er etablert, vil ofte brukskostnadene være neglisjerbare så lenge bruken ligger innenfor kapasitetsgrensen. Det er et sentralt prinsipp for samfunnsøkonomisk prissetting at den enkelte bør betale en pris som reflekterer den kostnad eller det økonomiske offer som bruken påfører samfunnet. I det forliggende tilfelle skulle det tilsi gratis bruk. Det gir også størst samfunnsøkonomisk verdi av de tjenestene som fasiliteten genererer. Eksempler på prosjekter som generer slike tjenester er samferdselsprosjekter som veier, broer og tunneler, kommunale tjenester som vannforsyning og avløp, og rekreasjonstjenester som museer og parker. Her er prising av tilgang til tjenesten mulig siden eksklusjon kan håndheves, men en pris høyere enn den marginale brukskostnaden vil isolert sett redusere den samfunnsøkonomiske lønnsomheten av prosjektet. Det skyldes at potensielle brukere med en betalingsvillighet for tjenesten som overstiger brukskostnaden, kan finne det ulønnsomt å benytte tjenesten når bruksavgiften overstiger brukskostnaden. Men når

prising av slike tjenester likevel griper om seg - spesielt innenfor samferdselssektoren - må det ses som en del av prosjektfinansieringen. Spørsmålet da blir hvilken finansieringsform som påfører samfunnet minst kostnader: skattefinansiering over offentlige budsjetter, eller finansiering ved beskatning av brukerne via bruksavgifter.

Oppsummeringsmessig kan vi si at fellesgoder skaper problemer for samfunnsøkonomisk verdsetting. Det er dels på grunn av at individuelle brukere ikke kan ekskluderes slik at en ikke har markedspriser å holde seg til, og dels er det på grunn av at prising av brukere ikke er samfunnsøkonomisk hensiktsmessig når bruken ikke medfører noe økonomisk offer for samfunnet. Sistnevnte tilfelle blir karakterisert ved at det ikke foreligger rivalisering i forbruket ettersom den enes bruk ikke berører andres muligheter for å bruke tjenesten.

Vi illustrerer dette ved et enkelt eksempel. Anta at prosjektet er bygging av en bro for biltrafikk og at en vurderer finansiering med bompenger. I nedenstående diagram viser den fallende kurven antall kryssinger over broen for alternative verdier på bomavgiften.

Kurven er å forstå som den marginale betalingsvilligheten for bruk av broen. Vi kan oppfatte den som etterspørselskurven. Jo høyere bomavgiften settes, desto færre vil bruke broen. Vi kan anta at de bruksavhengige kostnadene kan neglisjeres. Ved gratis bruk, vil den pengemessige nytten av prosjektet være gitt ved arealet  $abc$  under etterspørselskurven. Det er enklest å forstå dersom vi tenker oss at brukerne er ordnet langs etterspørselskurven etter fallende betalingsvillighet. Samlet betalingsvillighet må da være summen av betalingsvilligheten for alle dem som bruker broen. Bomavgiften viser da betalingsvilligheten til den marginale bruker. Om vi nå tenker oss en bomavgift lik  $p$ , vil bruken reduseres til volumet  $d$ . Dette medfører at de med betalingsvillighet lavere enn  $p$  ikke lenger finner det lønnsomt å bruke broen. Dette medfører en reduksjon i samlet betalingsvillighet med den gråfargede trekanten  $edb$  for dem bruker broen



Figur 14-1

Dette tapet - som blir kalt effektivitetstapet - er den samfunnsøkonomiske kostnaden ved brukerfinansiering i dette tilfellet. Effektivitetstapet fører til redusert lønnsomhet av

prosjektet. Alternativt kunne prosjektet ha vært finansiert ved generell beskatning. Denne finansieringsformen eliminerer effektivitetstapet ved brukerprising, men det fører på den annen side til et økonomisk tap ved at beskatning påvirker skattebetalernes økonomiske tilpasning. Fra et lønnsomhetssynspunkt vil den mest hensiktsmessige finansieringsformen være den som påfører økonomien minst effektivitetstap.

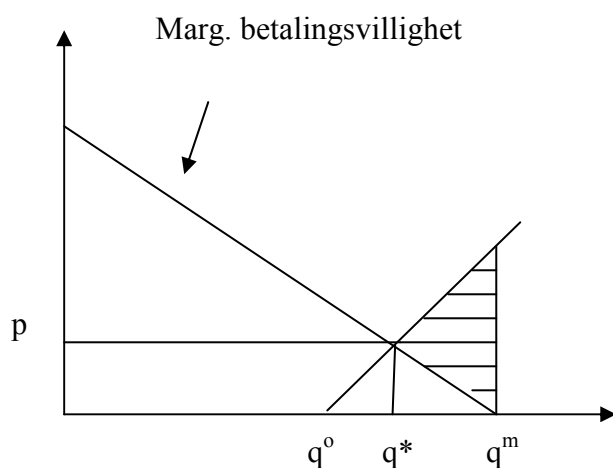
Vi ser fra figuren at effektivitetstapet avhenger av brattheten til etterspørselskurven. Jo mindre trafikken reagerer på bomavgiften, desto brattere vil etterspørselskurven være og følgelig vil effektivitetstapet bli mindre. Fra et effektivitetssynspunkt vil brukerfinansiering i slike tilfelle være gunstig. Det at samlet bruk reagerer lite på bruksavgiften, vil normalt skyldes fravær av tjenlige transportalternativer. I slike tilfelle vil brukeravgiften ha liten innvirkning på trafikkantenes valg av transportform, og den kommer derfor nært det som vi kaller for en nøytral skatt.

Bompengeinntektene er gitt ved det blå rektanget *adep* som er mindre enn den samfunnsøkonomisk nytten ettersom den bedriftsøkonomiske inntekten ved en lik bompengeavgift for alle ikke fanger opp konsumentoverskuddet som på figuren er gitt ved den røde trekanten *pec*. Prosjektet kan derfor være samfunnsøkonomisk lønnsomt selv om det er bedriftsøkonomisk ulønnsomt siden konsumentoverskuddet ikke fanges opp i den bedriftsøkonomiske kalkylen.

Noen typer fellesgoder har den egenskapen at bruken er å oppfatte som ikke-rivaliserende opp til en bestemt kapasitetsgrense, mens ytterligere bruk vil forringe andres mulighet til å nyttiggjøre seg godet på grunn av trengsel og kødannelse. Mange samferdselsprosjekter har denne egenskapen. Utover denne kapasitetsgrensen får godet karakter av rivaliserende ved at ytterligere bruk vil medføre en bruksrelatert kostnad i form av venting og annen form for ulempe. Denne eksterne kostnaden er et selvstendig argument for brukerprising, slik at den enkelte bruker må ta innover seg de kostnader som bruken påfører andre brukere. Vi ser på dette ved å modifisere eksemplet i tilknytning til figur 1 slik at vi for trafikk utover  $q^0$  har en stigende marginal køkostnad. Gratis bruk fører til en trafikk lik  $q^m$  og dermed til et samlet nettotap lik den skraverte trekanten ved at trafikk i intervallet  $q^*-q^m$  fører til en høyere marginal køkostnad enn marginal<sup>45</sup> betalingsvillighet. Innføring av en bomavgift lik  $p$  som avstemmer marginal køkostnad mot marginal betalingsvillighet for bruken av broen, eliminerer effektivitetstapet gitt ved den skraverte trekanten og gir samtid bompengeinntekter svarende lik  $pq^*$ . Køprising gir dermed både inntekter og en effektivitetsgevinst slik at det blir en form for såkalt ”dobbel dividende”.

---

<sup>45</sup> Nettotapet i dette intervallet er lik samlet betalingsvillighet minus samlet køkostnad der sistnevnte er gitt ved summen av de marginale køkostnader og angitt ved arealet under kurven for marginale køkostnader fra  $q^*$  til  $q^m$ .



Figur 14-2

Noen typer tjenester er blandete i den forstand at de har karakter både av fellesgoder der en ikke har eksklusjonsmuligheter og forbruket ikke medfører rivalisering, og av private rivaliserende markedsgoder der tjenestene kan omsettes i markedet. Eksempler på slike er goder som har en selvstendig eksistensverdi utover en ren bruksverdi. Selve muligheten for tilgang til godet, er et fellesgode ved at den gjelder for alle potensielle brukere. Dette kan f.eks. gjelde miljøgoder der eksistensen har en egenverdi utover den betalingsvilligheten som blir avslørt gjennom faktisk bruk. Bruksverdien kan avsløres ved at bruken kan prises ved tilgangs- eller bruksavgifter dersom eksklusjon er mulig. Eksistensverdien vil imidlertid ikke bli avslørt i avgiftene på faktisk bruk. Tilsvarende gjelder for beredskapsordninger som brannvesen, redningstjeneste og akuttmedisin. En påregnelig tilgang til slike tjenester har karakter av en forsikring, og betalingsvilligheten for denne forsikringstjenesten kunne i prinsippet tenkes innkassert gjennom en aksesspris utover selve prisen for tjenesten ved faktisk bruk. Dette vil imidlertid kunne by på gratispassasjerproblemer dersom eksklusjon ikke er mulig eller troverdig når der er ledig kapasitet. Den faktiske bruken vil imidlertid kunne prises på vanlig måte og vil da i prinsippet kunne innkassere bruksverdien for tjenesten. Men hvis en baserer lønnsomhetsvurderingen kun på inntektene fra faktisk bruk, vil eksistensverdien av tjenestetilbudet ikke bli fanget opp.

I noen sammenhenger kan eksistensverdien likestilles med verdien av leveringssikkerhet. Det kan f.eks. gjelde for nettverk som har karakter av kritisk infrastruktur<sup>46</sup>. Kritikaliteten går på at der ikke fins brukbare alternative teknologier som kan understøtte levering av tjenesten på produksjonssiden, og heller ikke tjenlige substitutter til tjenesten på etterspørselssiden. Eksempler på slik infrastruktur kan være elektrisitetsnett, vann- og avløpsnett og kommunikasjonsnett. Potensielle brukere kan ha betalingsvillighet for å opprettholde slik infrastruktur selv om de i dag ikke er aktuelle brukere. Der er imidlertid ikke noe marked hvor denne betalingsvilligheten kommer eksplisitt til syne.

Generelt er det to faktorer som driver slike eksistensverdier. Den ene er at dersom tjenesten forsvinner, vil det kunne ha irreversible konsekvenser enten ved at godet forsvinner for alltid (f.eks. regnskogen i Amazonas på grunn av rovhogst eller blåhvalen på grunn av rovfangst),

<sup>46</sup> Se NOU 2006:6 for en nærmere diskusjon av dette.

eller på grunn av leveringsavbrudd for samfunnskritiske tjenester (f.eks. at livsviktig strømforsyning uteblir på grunn av underinvesteringer i elektrisitetsnettet). Den andre faktoren er usikkerhet om fremtidige preferanser og etterspørsel. Selv om tjenesten er bedriftsøkonomisk ulønnsom ved at dagens bruk og etterspørsel for slike tjenester ikke dekker alle kostnadene ved å holde tjenesten oppe, kan tilbudet like fullt være samfunnsøkonomisk forsvarlig siden det kan eksistere usikkerhet om det fremtidige behovet. Det kan derfor være betalingsvillighet for en opsjon på å benytte tjenesten i fremtiden. Slik etterspørsel blir kalt opsjonsetterspørsel<sup>47</sup>, men der er normalt ikke markeder for slike opsjoner. Opprettholdelse av slike tjenester og nødvendig infrastruktur for understøttelse av slike tjenester blir derfor en fellesoppgave og løsningen av slike fellesoppgaver er et myndighetsansvar.

### **Markedsbestemte priser reflekterer ikke samfunnsøkonomisk riktige priser.**

I det ovenstående har vi redegjort for forhold som gjør det vanskelig å omsette visse goder gjennom markeder (mangel på eksklusjonsmuligheter og gratispassasjerproblemer), eller som gjør det uhensiktsmessig å finansiere tilbudet i markedet (små kostnader knyttet til bruk og underutnytting av kapasitet). Dette er viktige grunner til at et samfunnsøkonomisk tjenlig tilbud av slike goder generelt ikke kan skje på et bedriftsøkonomisk lønnsomt grunnlag. I dette kapitlet skal vi diskutere forhold som fører til at selv om goder omsettes gjennom veletablerte markeder, vil ikke markedsprisene nødvendigvis være en god rettesnor for samfunnsøkonomisk lønnsomhet. Spørsmålet blir da hvordan markedsprisene bør korrigeres for å kunne benyttes i en samfunnsøkonomisk lønnsomhetsanalyse.

Vi har tre hovedgrunner til at markedsprisene ikke samsvarer med samfunnsøkonomisk korrekte priser. Det er eksterne virkninger, naturlig monopol og skatter og avgifter.

#### ***Eksterne virkninger***

Eksterne virkning er virkninger som den enkelte aktør påfører andre aktører og som

- (i) ikke tas med i de kalkyler som ligger til grunn for aktørens beslutninger
- (ii) ikke går via markeder

Punkt (i) innebærer at den enkelte markedsaktør ikke tar hensyn til de samlede økonomiske kostnader og inntekter som kan tilskrives den enkeltes beslutninger. Privat- eller bedriftsøkonomisk lønnsomhet er i slike tilfelle ikke dekkende for samfunnsøkonomisk lønnsomhet. Beslutninger basert på bedriftsøkonomisk lønnsomhet vil derfor i situasjoner med eksterne virkninger føre til samfunnsøkonomisk ulønnsom ressursbruk. Punkt (ii) uttrykker at eksterne virkninger er en defekt ved markedssystemet. Det skyldes at markedsprisene ikke fanger opp alle velferdsrelevante sider ved konsum av forbrugsgoder eller bruk av produksjonsfaktorer. Det kan i mange tilfelle skyldes at det er knappe goder og ressurser som ikke blir omsatt i markeder, og som det derfor heller ikke fins markedspriser for. I så måte er problemer knyttet til samfunnsøkonomiske lønnsomhetsberegninger på områder med eksterne virkninger nært beslektet med de problemer som oppstår på grunn av manglende markeder.

---

<sup>47</sup> Se Arrow & Fisher (1974)

Eksterne virkninger kan være positive eller negative, og de kan forekomme mellom konsumenter, mellom produsenter og mellom produsenter og konsumenter. Positive eksterne virkninger betyr at en aktørs beslutninger har positive virkninger for andre aktører uten at den som genererer disse positive virkningene, blir kompensert for dette gjennom pris og belønningssystemet i markedet. Siden slike virkninger ikke blir reflektert i markedsbestemt inntekt og profitt, vil de derfor heller ikke bli tatt hensyn til ved profittmotivert tilpasning. Handlinger som har slike positive eksterne sidevirkninger, vil derfor bli utført i for liten skala vurdert ut fra samfunnsøkonomisk lønnsomhet. Omvendt er negative eksterne virkninger nytte- eller produksjonstap som andre aktører blir påført, uten at de som forårsaker slike kostnader blir holdt økonomisk ansvarlig for dette i markedet. Handlinger med slike virkninger vil derfor bli utført i for stort omfang.

Negative eksterne virkninger innebærer at markedspriser overvurderer netto samfunnsøkonomisk betalingsvillighet dersom et individs forbruk har negative eksterne virkninger for andre, eller at bedriftsøkonomiske kostnader undervurderer samlet samfunnsøkonomisk kostnad dersom en bedrifts produksjon eller faktorinnsats har negative virkninger for andre aktører. Positive eksterne virkninger innebærer tilsvarende at markedspriser undervurderer samfunnsøkonomisk nytte og bedriftsøkonomiske kostnader overvurderer samfunnsøkonomiske kostnader.

Eksempler på eksterne virkninger er mange. En viktig klasse av negative eksterne virkninger knytter seg til forurensninger. Når en bedrifts forurensende utslipp i et vassdrag har negative virkninger på andre nedstrøms brukere av vassdraget, er dette en samfunnsøkonomisk kostnad som ikke uten videre kommer til syne i bedriftens kostnader. Disse eksterne kostnadene kan f.eks. bestå i at friluftinteressenter får mindre nytte av vassdraget, eller at andre bedrifter som er avhengig av tilgang på rent vann får en ekstrakostnad på grunn av nødvendig rensing. Den forurensende bedriften vil derfor legge en for lav kostnad til grunn ved sin produksjonstilpasning. Forurensende virksomhet fører derfor til overproduksjon ut fra en samfunnsøkonomisk lønnsomhetsvurdering.

En viktig klasse av negative eksterne virkninger er gratis bruk av allmenninger - en problemstilling som går under betegnelsen allmenningens tragedie. Et eksempel på allmenningens tragedie er fritt fiske på knappe fiskeressurser. Når en fiskebåt fisker på en knapp fiskeressurs, vil dette redusere produktiviteten til andre som fisker på samme ressurs. Siden samme fisk ikke kan fiskes mer enn en gang, betyr dette at andre må bruke mer ressurser for å fiske opp samme kvantum som før. Dette er en ekstern kostnad som påføres andre fiskere, og som den enkelte fisker ikke vil ta hensyn til ved sin egen tilpasning. Fritt fiske vil derfor føre til overfiske ut fra en samfunnsøkonomisk nytte-kostnadsvurdering og ressursen kan forsvinne eller degraderes.

Som et tredje eksempel kan vi ta et med positive eksterne virkninger. Anta at en bedrift driver et opplæringsprogram for sine ansatte for å gjøre de mer produktive i bedriften. Verdien av dette programmet for bedriften er nåverdien av produktivitetsøkningen diskontert over den gjennomsnittlige tid arbeiderne blir i bedriften. Opplæringsprogrammet er lønnsomt så lenge nåverdien er positiv. Dersom produktivitetsøkningen avhenger av programmets skala, vil det lønne seg å investere i programmet så lenge økningen i netto nåverdi er positiv.

I den samfunnsøkonomiske lønnsomhetsanalysen må en imidlertid legge arbeidernes hele gjenværende yrkesaktive periode til grunn i lønnsomhetskalkylen. Den samfunnsøkonomiske lønnsomheten av programmet er dermed høyere enn den bedriftsøkonomiske, og det optimale

samfunnsøkonomiske omfanget av programmet vil også være større enn det som er bedriftsøkonomisk optimalt. Bedriften vil dermed kunne unnlate å investere i opplæringsprogrammer som er samfunnsøkonomisk lønnsomme, og vil om den investerer i opplæring av de ansatte, normalt investere for lite i forhold til hva som er optimalt ut fra en samfunnsøkonomisk lønnsomhetsanalyse.

Grunnen til avviket mellom den bedrifts- og samfunnsøkonomiske lønnsomheten av programmet er at bedriften ikke kan innkassere den samlede verdiskapingen som opplæringen genererer. Arbeiderne tar noe av verdiskapingen med seg når de forlater bedriften for å gå over i ny virksomhet, enten i form av høyere lønn i ny virksomhet, eller ved at verdien av den økte produktiviteten også tilflyter den nye arbeidsgiveren dersom den ikke er fullt reflektert i den nye lønnen. Bedriftsopplæringsprogrammet har generert en positiv ekstern virkning som gjør at den samfunnsøkonomiske lønnsomheten er større enn den bedriftsøkonomiske. Bedriften kan internalisere mer av produktivitetsgevinsten ved å investere i bedriftsspesifikk opplæring som har begrenset verdi utenfor bedriften. Dette reduserer gapet mellom bedrifts- og samfunnsøkonomisk lønnsomhet av bedriftsintern opplæring, og indikerer samtidig at generelt rettet utdanning bør være et offentlig ansvar. Problemstillingen er analog til forholdet mellom anvendt forskning og grunnforskning. Anvendt rettet forskning er kunnskap som har verdi bare i avgrensede anvendelser, og dens økonomiske verdi kan innkasseres ved å gi bedriften enerett på anvendelsen. Grunnforskning rettet mot frembringelse av ny basiskunnskap vil normalt ikke ha slike klart avgrensede anvendelsesområder og det vil derfor være vanskeligere eller mer uhensiktsmessig å innkassere verdien gjennom markedet.

Ut fra den innsikten vi har fått fra eksemplene ovenfor, ser vi at mange eksterne virkninger har sin rot i gratisproduksjon eller gratisforbruk av fellesressurser som ikke blir allokert gjennom markedssystemet. En mulig løsning på dette problemet er å subsidiere eller avgiftsbelegge aktiviteter med positive eller negative eksterne virkninger slik at brukerprisen svarer til de samfunnsmessige nytte- eller kostnadsvirkningene som bruken medfører. På denne måten kan disse virkningene bli internaliserte i markedsprisene slik at de som forårsaker disse virkningene, blir stilt overfor en privatøkonomisk kostnad eller inntekt som reflekterer den samfunnsøkonomiske verdien av de eksterne virkningene.

### ***Naturlig monopol***

Naturlig monopol brukes som betegnelse på virksomhet der det foreligger stor- og samdriftsfordeler på produksjonssiden. Stordriftsfordeler innebærer at enhetskostnadene faller med økende produksjonsvolum. Ved enwareproduksjon vil dette kunne være tilfelle dersom de variable produksjonskostnadene er små i forhold til de faste driftsuavhengige kostnadene knyttet til etablering og opprettholdelse av produksjonskapasitetet. Større produksjon betyr flere produkter å fordele de faste kostnadene på. De samlede kostnader pr enhet kan derfor bli lavere selv om de variable enhetskostnadene er stigende ved økt kvantum.

Naturlig monopoler har en ofte på områder der virksomheten krever store investeringer i infrastruktur. Det gjelder særlig innenfor nettverkbasert virksomhet som flyplasser, jernbane- og elektrisitetsnett, og vann- og avløpsnett. Problemet med naturlige monopoler er at uten prisregulering vil det kunne utnyttes monopolmakt ved å sette prisen høyere en grensekostnaden i produksjon for på denne måten å få hånd om en større del av konsumentoverskuddet. Siden dette vil være en ren overføring fra kjøperne av monopolets tjenester til monopolets eiere, vil ikke dette bidra til samfunnets verdiskaping. Monopolet vil derfor sette



for høye priser og produsere et lavere kvantum enn det som maksimerer den samfunnsøkonomiske verdiskapingen fra virksomheten. Monopoloverskuddet blir følgelig lavere enn det samfunnsøkonomiske overskuddet ved optimal prissetting som vil være gitt grensekostnaden som viser hva det koster å øke produksjonen marginalt.

Dersom monopolet utnytter monopolmakt ved leveranser av innsatsfaktorer til annen innenlandsk virksomhet, vil markedsprisen være høyere enn det marginale økonomiske offeret gitt ved merkostnaden. Dette gir ringvirkninger i form av for lite volum og samfunnsøkonomisk effektivitetstap i virksomheter som er avhengig av innsatsfaktorer fra innenlandske naturlige monopoler. I tillegg oppstår problemet hvilken pris som skal legges til grunn for samfunnsøkonomiske lønnsomhetsberegninger av prosjekter som er avhengig av vareinnsats fra mangelfullt regulerte naturlige monopoler. To kandidater peker seg ut. Det ene er det en må betale i form av markedspris. Det andre er det samfunnsøkonomiske offeret som er gitt ved den faktiske merkostnaden på produksjonssiden. Dersom monopolet i utgangspunktet er optimalt prisregulert, er det naturlig å legge markedsprisen til grunn slik at både privat og offentlig virksomhet som er avhengig av den monopoliserte innsatsfaktoren, legger samme kostnad til grunn i sine kalkyler.<sup>48</sup>

### ***Skatter og avgifter.***

Skatter og avgifter fører til at den prisen som produsenten mottar for salg av varer og tjenester er lavere enn den prisen som kjøperen betaler. Differansen utgjøres av skatten eller avgiften. For varer som er subsidierte vil den prisen som produsenten mottar være høyere enn den prisen som kjøperen betaler. Om vi tenker oss at en vare er belagt med en stykkskatt med et beløp på  $t$  pr enhet på produsentleddet, og markedsprisen eksklusive avgiften er  $p$ , så vil kjøperen betale prisen  $p+t$ , mens produsenten vil få en pris lik  $p$ . Dersom vi antar at dette er et frikonkurransemarked der produsenten vil produsere inntil pris eksklusive skatt er lik grensekostnaden, og kjøperne vil etterspørre tjenesten inntil pris inklusive skatt er lik marginal betalingsvillighet, vil marginal betalingsvillighet i markedslukevekten være høyere enn grensekostnaden. Skatten gjør at volumet blir mindre enn det som maksimerer det samlede samfunnsøkonomiske overskuddet og vi får et effektivitetstap i dette markedet. Effektivitetstapet skyldes at produsenten ser på stykkskatten som en kostnad mens den for samfunnet er en overføring til myndighetene. Den bedriftsøkonomiske lønnsomheten av økt produksjon blir dermed lavere enn den samfunnsøkonomiske og vi får for liten produksjon sett fra en samfunnsøkonomisk synsvinkel.

Også her vil det i en samfunnsøkonomisk lønnsomhetsanalyse oppstå spørsmål om hvilken pris om bør legges til grunn, eksklusive eller inklusive skatter og avgifter. Vi kan se på dette i tilknytning til beregning av samfunnsøkonomiske kostnader for bruk av innenlandsk arbeidskraft. Dersom arbeidslønn pr tidsenhet er  $w$  og lønsskatten er  $t$  vil arbeideren sitte igjen med  $w(1-t)$  av hver krone som utbetales i lønn. Hvis vi for argumentets skyld forutsetter at det er et frikonkurransemarked, vil arbeiderne tilby arbeidskraft inntil lønn etter skatt er lik det marginale offeret med mer arbeid. Bedriften vil på sin side etterspørre arbeid inntil verdsettingen av merproduksjonen (grenseproduktet) er lik lønn inklusive skatt. Spørsmålet er da hvilken lønn som bør brukes ved en samfunnsøkonomisk lønnsomhetsanalyse av

---

<sup>48</sup> Et mulig eksempel her kan være sement som er skjermet mot utenlandsk konkurranse på grunn av høye transportkostnader.

prosjektet, lønn eksklusive eller inklusive skatt. Det samme problemet oppstår ved behandling av arbeidsgiveravgiften som er skatt på bruk av arbeidskraft i bedriften.

Ved at skatter og avgifter fører til en kile mellom marginal betalingsvillighet og marginale kostnader, innebærer skattefinansiering av prosjekter et effektivitetstap, og dette er de egentlige samfunnsøkonomiske kostnadene ved skattefinansiering. I Norge er skattekostnadene anslått til 20 øre pr skattekroner<sup>49</sup>. Det betyr at ved skattefinansiering skal prosjektet belastes en finansieringskostnad på 20 pst av prosjektets netto budsjettvirkning for det offentlige. Alternativet til skattefinansiering er finansiering i markedet enten ved brukerbeskatning om det er mulig, eller via en positiv profittmargin på prosjektets tjenester. Dette fører også til et effektivitetstap ved at den prisen som brukerne møter er høyere enn det som økt produksjon faktisk koster. Fra et rent effektivitetssynspunkt vil ofte en deling mellom markedsfinansiering og skattefinansiering være det beste. Den optimale finansieringsmiksen vil da være bestemt ved at det marginale effektivitetstapet er likt ved skattefinansiering og finansiering i markedet. Det betyr at den siste krona koster samfunnet det samme i form av effektivitetstap enten den kommer som markedsinntekt eller i form av skatt.

## 14.4 Fra teori til analyse og beslutning.

En samfunnsøkonomisk lønnsomhetsanalyse av et konkret prosjekt krever en kartlegging av prosjektet både med hensyn til hva prosjektets genererer av ytelser, nødvendig innsats av produksjonsfaktorer og verdsetting av ytelsene og faktorinnsatsen. Verdsetting er basert på betalingsvillighet for output som retter seg mot innenlandske sluttbrukere og på alternativkostnader for faktorinnsats. Alternativkostnaden er basert på produksjonsfaktorenes verdi i beste alternative anvendelse, som kan være et alternativt prosjekt rettet mot samme formål. Det følger av dette at alternativkostnaden er kun positiv i de tilfelle produksjonsfaktoren har alternativ anvendelse på kort eller lang sikt. En slik kalkyle vil vise et positivt prosjektoverskudd bare i de tilfelle at prosjektet gir et større bidrag til landets verdiskaping enn hva ressursene ville ha gitt i den beste alternative anvendelse. I praksis knytter verdsettingsproblemen seg til i hvor stor grad en kan legge markedspriser til grunn. De prinsipielle problemene som dette reiser, ble behandlet i foregående kapittel. Her vil vi ta opp problemer knyttet til operasjonalisering og implementering. Vi vil også ta opp en del spørsmål som knytter seg til store prosjekter, behandling av informasjon og usikkerhet om sammenheng mellom innsatsfaktorer og produksjon og om priser og verdsetting, samt ønskeligheten av beslutningsfleksibilitet i prosjektgjennomføringen..

### Prosjektavgrensning og klargjøring av alternativer.

En lønnsomhetsanalyse er et verktøy for beslutningsstøtte når det gjelder ressursbruk, og er derfor bare relevant i de tilfelle at de ressurser som et prosjekt legger beslag på, har alternativ anvendelse. Dersom det ikke foreligger alternativer, er det heller ikke noe beslutningsproblem. Denne alternative ressursbruken kan være et alternativt prosjekt som da får karakter av et sammenligningsgrunnlag, eller det kan være den anvendelsen som ressursene har i utgangssituasjonen, som da gjerne blir betegnet nullalternativet. Det alternativet som er gjenstand for vurdering og nullalternativet blir da gjensidig utelukkende prosjekter. De fleste prosjekter har en tidsutstrekning både når det gjelder produksjon og kostnader.

<sup>49</sup> Se Finansdepartementet: NOU 1997: 27 og Finansdepartementet 2005

Prosjektdefinisjonen består da i å beskrive tidsutviklingen for tjenesteproduksjon, ressursbruk og kostnader over prosjektets levetid. Nullalternativet består i dette tilfellet i en fremskrivning av utgangssituasjon for hele prosjektets levetid, og dette blir da en referansebane som det aktuelle prosjektet blir sammenlignet med.

Noen prosjekter er av lokal karakter ved at de berører ressursbruk og tjenesteyting i et bestemt marked eller innenfor en enkeltstående sektor. Et eksempel på dette kan være bygging av et nytt rådhus for kommuneadministrasjonen i en kommune. Inntekten fra dette prosjektet er gevinsten ved at administrasjonen får bedre fysiske muligheter til å yte servise overfor innbyggerne i kommunen. Kostnadene består i differansen mellom kapital- og driftskostnadene for det nye rådhuset og kostnadene som vil falle bort når det gamle rådhuset som er nullalternativet her, legges ned, dvs merkostnadene. Dette beskriver de samlede økonomiske virkninger fra dette prosjektet.

Andre prosjekter kan være store med sektorovergrepene ringvirkninger. Bygging av bro over Hardangerfjorden vil være et slikt prosjekt. Formålet med prosjektet er dels å knytte Hardanger nærmere til Bergensregionen, og dels å ruste opp Rv7 over Hardangervidda til ferjefri hovedveiforbindelse mellom Oslo og Bergen. Dette prosjektet vil ha følgevirkninger i form av relaterte prosjekter. F.eks. vil dette kunne være prosjekter knyttet til rassikring i indre Hardanger. Når trafikken øker, vil de samfunnsøkonomiske kostnadene ved rasrisiko øke, slik at i forhold til den nye trafikksituasjonen vil det være lønnsomt å bruke økte ressurser på rassikring. Økt trafikk over Hardangervidda vil ha negative konsekvenser for den truede villreinstammen på Vidda. Dette kan kreve at deler av veien over fjellet legges i tunneler. Foruten kostnadene fører dette til at naturopplevelsen ved å kjøre over uberørte fjellområder blir redusert. Begge disse virkningene er følgevirkninger av prosjektet og bør behandles som en integrert del av prosjektet. Når det gjelder ringvirkninger for eksisterende virksomhet og trafikk på alternative ruter mellom disse to byene, er det mer uklart hva som er relevant å ta med i kalkylen. For eksempel vil turistbasert virksomhet langs Rv7 bli mer lønnsom noe som kan gi grobunn til etablering av ny virksomhet. Her kan det være vanskelig å skille mellom nyskapt virksomhet og hva som er overflytting av eksisterende virksomhet, f.eks fra Rv5 over Lærdal og Filefjell. Som en hovedregel kan en si at en bør være forsiktig med å inkludere slike ringvirkninger. En litt mer spesiell ringvirkning vil være at en får mindre trafikk gjennom Lærdalstunnelen slik at lønnsomheten av dette prosjektet blir mindre enn det som var forutsatt. Spørsmålet er om dette burde tilskrives Hardangerbroprosjektet som en kostnad. Dette er imidlertid en klar feilslutning. Kostnadene ved Lærdalstunnelen er historiske og irreversible og blir ikke påvirket av broprosjektet i Hardanger. Om derimot mindre trafikk over Lærdal fører til lavere kostnader til veivedlikehold, bør det tilskrives det nye prosjektet som en gevinst. Oppsummeringsmessig kan en si at dersom det beste alternativet til Hardangerbro er å legge til rette for Filefjell som hovedvei mellom Oslo og Bergen, vil lønnsomheten måtte anslås ved en differansebetragtning av samfunnsøkonomiske inntekter og kostnader ved de to alternativer fremskrevet over prosjektenes levetid.

## 14.5 Det samfunnsøkonomiske verdsettingsproblemet: Forholdet mellom markedspriser og korrekte samfunnsøkonomiske kalkylepriser

### Markedspriser og samfunnsøkonomisk inntekt.

Verdien av goder som retter seg mot innenlands privat forbruk, er basert på brukernes betalingsvillighet. Den marginale betalingsvilligheten faller normalt med økende kvantum slik at den gjennomsnittlige betalingsvilligheten for et gitt kvantum er høyere enn den marginale. Konsumentteorien bygger på at konsumentene etterspør varer inntil marginal betalingsvillighet er lik markedsprisen. Markedsprisen vil da reflektere den marginale betalingsvilligheten, og om vi antar at prosjektet ikke påvirker markedsprisen i nevneverdig grad vil korrekt kalkylepris være gitt ved den prisen som konsumentene møter i markedet. For varer som er belagte med fiskalt motiverte særavgifter eller merverdiavgift, innebærer det at kalkyleprisen bør være inklusive skatt og avgifter som blir den delen av betalingsvilligheten som det offentlige innkasserer. Om derimot særavgifter er motivert ut fra negative eksterne virkninger i forbruket, bør markedspris eksklusive slike avgifter legges til grunn ettersom de skal reflektere den eksterne samfunnsmessige kostnaden knyttet til privat bruk og derfor trekkes fra den private betalingsvilligheten. Et aktuelt eksempel her er bensinavgiften.

For varer og tjenester som eksporteres, vil den korrekte kalkyleprisen være netto eksportpris når vi antar den er upåvirket av prosjektet. Det betyr at den bør være eksklusive eventuelle eksportsubsidier eller -avgifter siden dette verken er en nasjonal inntekt eller kostnad.

Prinsipielt er verdsettingen av et fellesgode gitt ved summen av betalingsvilligheten til alle dem som vil ha nytte av godet. Når godet ikke kan individualiseres og konsumentene ikke ekskluderes, vil der ikke være noe marked for slike goder og derfor heller ikke priser som kan gi informasjon om den enkeltes betalingsvillighet. Betalingsvillighet og samfunnsøkonomisk nytte må derfor anslås. Vi kan skille mellom direkte og indirekte metoder for verdsetting. Direkte verdsetting er basert på intervjuundersøkelser om betalingsvillighet for tilbud av nærmere spesifiserte fellesgoder. Verdsettingen er hypotetisk i den forstand at den enkelte blir bedt om å angi betalingsvillighet for et fellesgode dersom det skulle bli tilbudt. Det at verdsettingssituasjonen er hypotetisk, er også den store svakheten ved metoden.

Indirekte metoder utnytter at det i mange tilfeller er perfekte komplementariteter i markedet til fellesgoder. Markedets verdsetting av perfekte komplementariteter vil da også kunne fange opp verdsettingen av fellesgodet. Slike perfekte komplementariteter kan være private goder eller aktiviteter som er nødvendige for å få tilgang til fellesgoder. Det gjelder særlig fellesgoder som er av lokal karakter. Tilgang til et godt bomiljø kan en f.eks. bare få ved å kjøpe en bolig i det aktuelle området. Verdsettingen av dette fellesgodet vil da kapitaliseres i markedsprisen for boligeiendom som da blir markedets verdsetting av dette fellesgodet, og på samme måten kan kostnaden for et lokalt fellesgode reflekteres i reduserte eiendomsverdier.

Markedsprisen reflekterer marginal betalingsvillighet, og dersom prosjektet påvirker markedsprisen negativt, vil markedsinntekten undervurdere den samfunnsøkonomiske inntekten. Som en approksimasjon kan en da legge den gjennomsnittlige betalingsvilligheten for prosjektets ytelser før og etter realisering til grunn.

### Markedspriser på innsatsfaktorer og samfunnsøkonomiske kostnader.

Her kan det være fruktbart å ta utgangspunkt i analogien med det bedriftsøkonomiske kalkyleprisproblemet for innsatsfaktorer som dels kjøpes i de eksterne markedet og dels skaffes til veie ved interne leveranser i bedriften. For nasjonen vil analogien til det eksterne marked være det internasjonale markedet. I utgangspunktet vil det økonomiske offeret for samfunnet være importpris eksklusive toll og andre importavgifter. Toll og importavgifter vil kun være overføringer fra importører til statskassen og landet blir verken rikere eller fattigere av at ”penger skifter hender”. Dersom prosjektet f.eks. består i å fornye politiets bilpark, vil det økonomiske offeret for landet være netto pris eksklusive engangsavgiften, og er det som bør legges til grunn i lønnsomhetsanalysen.

Imidlertid vil private bedrifter måtte anse toll som en bedriftsøkonomisk kostnad. Dersom offentlig virksomhet konkurrerer med privat virksomhet, vil dette gi det offentlige en fordel som skyldes behandlingen av importavgiften og ikke at det offentlige er mer effektiv enn private. Det at mer effektiv privat virksomhet blir utkonkurrert på grunn av forskjellsbehandling av importavgifter i lønnsomhetsanalysen, resulterer i en kostnad for landet som det bør tas hensyn til. En praktisk enkel måte å ta hensyn til dette problemet på, er først å avgjøre om det offentlige prosjektet har en åpenbar konkurranseflate med privat importavhengig virksomhet. Hvis svaret er ja, legges importpris inklusive importavgift til grunn, og om svaret er nei, benyttes pris eksklusive avgift.<sup>50</sup>

Tilsvarende resonnement gjelder for innenlandsk bruk av ressurser som alternativt kan eksporteres. Det mest aktuelle eksemplet her er verdsetting av gass i lønnsomhetsanalysen av gasskraftverk. Dersom dette ikke påvirker de priser som norsk gass selges for i utlandet, bør markedsprisen ute legges til grunn for kostnaden for innenlandsk bruk av gass. Dersom bruk av gass i gasskraftverk fører til negative eksterne virkninger for norske borgere, bør den pengemessige vurderingen av denne kostnaden inkluderes i kalkyleprisen. Når det gjelder avbrenning av naturgass, er det meste av forurensningene knyttet til utslipp av CO<sub>2</sub> der konsekvensene i hovedsak er av global karakter slik at de i liten grad bæres av norske borgere. Om det er velferdsvirkninger for norske borgere som legges til grunn, burde de holdes utenfor analysen. Norge har imidlertid påtatt seg å overholde en maksimumsgrense for sine CO<sub>2</sub>-utslipp. I den grad dette taket er bindende, vil de nasjonale utslippskostnadene ved et gasskraftverk bestå i den negative virkningen i form av redusert verdiskaping ved at annen virksomhet må redusere sine utslipp. Dette vil da være den nasjonale utslippskostnaden. Om det var et velfungerende innenlandsk marked for utslippskvoter, vil denne negative eksterne virkningen være reflektert i kvoteprisen. Dette er nok et eksempel på at samfunnsøkonomiske lønnsomhetsanalyser blir enklere i land med velutviklede markeder. Men samtidig blir også behovet for samfunnsøkonomisk overprøving av bedriftsøkonomiske kalkyler mindre.

Den gjenstående muligheten er verdsetting av innsatsfaktorer fra skjermet sektor, dvs ressurser der importalternativet ikke er til stede. I dagens situasjon vil det gjelde deler av arbeidsmarkedet. Problemet med verdsetting av arbeidskraft er at den på grunn av beskatningen blir forskjellig om en ser det fra tilbuds- eller etterspørselssiden i arbeidsmarkedet. Det økonomiske offeret ved å bruke arbeidskraft er dens verdi ved alternativt bruk. Dersom arbeidsmarkedet i utgangspunktet er i likevekt og import av arbeidskraft er uaktuelt, kan arbeidskraft til et bestemt formål skaffes tilveie enten ved økt tilbud av arbeid,

<sup>50</sup> Se Hagen (2005), kap. 5.2.7 for en grundigere diskusjon av dette spørsmålet.

eller ved redusert bruk av arbeidskraft i annen virksomhet, eller ved en kombinasjon av de to måtene. I det første tilfellet er offeret redusert fritid og i det andre tilfellet redusert verdiskaping ved alternativ bruk av arbeidskraften. Problemet er at på grunn av beskatningen vil verdsettingen være forskjellig i disse to anvendelsene. I et arbeidsmarked med effektiv konkurranse vil arbeiderne tilby arbeidskraft inntil den marginale verdien av tapt fritid er lik lønnsatsen etter skatt. Om lønnsatsen er  $w$  og skatten er 50% blir netto lønn  $0,5w$ . Dette blir da den marginale verdsettingen av fritid. Bedriftene etterspør arbeidskraft inntil brutto lønn inklusive arbeidsgiveravgift er lik verdien av arbeiderens grenseprodukt. Dersom vi setter denne avgiften til 15 %, blir den marginale verdsettingen av arbeidskraften på produksjonssiden  $1,15w$ . Vi ser da at om disse alternativkostnadene legges til grunn, blir prosjekter som har en betydelig sysselsettingseffekt mer lønnsomme enn prosjekter som konkurrerer arbeidskraften bort fra eksisterende produksjonsvirksomhet.

Dersom beskatningen er optimalt utformet, og der ikke er ufrivillig arbeidsledighet i utgangspunktet, vil det likevel være unaturlig å anse økt sysselsetting som et ekstra pluss ved prosjektets lønnsomhet. Om det virkelig var slik, kunne ikke skattesystemet være optimalt i utgangspunktet og det ville være mest hensiktsmessig å oppnå denne sysselsettingseffekten ved endret beskatning. En bør derfor kalkulere som om et offentlig prosjekt trekker arbeidskraft bort fra privat sektor, slik at en i offentlige kalkyler bruker brutto arbeidslønn inklusive faktorskatter på arbeidsgiversiden. Ved dette oppnår en også konkurransemessig nøytralitet mellom offentlig og privat bruk av arbeidskraft.

## 14.6 Avkastningskrav og det offentliges diskonteringsrente

Mange offentlige prosjekter er kapitalkrevende prosjekter. Det gjelder f.eks. innenfor samferdsel, kraftforsyning og vann og avløp. Kalkyleprisen på innsatsfaktoren kapital blir dermed en viktig faktor for prosjektene lønnsomhet. Rent prinsipielt bør kalkyleprisen på binding av kapital i et prosjekt reflektere alternativkostnaden som er den samfunnsøkonomiske avkastning som kapitalen ville gi i beste alternative anvendelse. Om vi holder oss til en lukket økonomi betraktning og tar utgangspunkt i at det innenlandske kapitalmarkedet er i likevekt, kan kapital for et bestemt formål skaffes til veie på to måter: enten ved økt sparing i økonomien eller ved redusert kapitalbruk i alternativ virksomhet eller ved en kombinasjon av de to mulighetene. Kostnadene ved å fremskaffe kapital ved økt sparing er den kompensasjon i form av avkastning som sparerne må ha for å være villig til å utsette konsum. Dersom konsumentene er optimalt tilpasset i kapitalmarkedet, vil det være gitt ved realrenten etter skatt. Alternativt kan kapital frigjøres ved at private bedrifter reduserer sine investeringer. Kostnaden ved dette er redusert verdiskaping i alternativ privat kapitalbruk. Dersom beskatningen virker nøytralt på bedriftenes kapitalbruk, er denne alternativavkastningen gitt ved realavkastningen før skatt. Men for at kapital kan skaffes til veie ved økt sparing eller reduserte private investeringer, må prosjektet føre til en høyere realrente. Om prosjektet er stort nok og kapitalen er knapp, trenger ikke dette være helt urealistisk. I dagens situasjon har vi imidlertid et åpent kapitalmarked med full kapitalmobilitet der det innenlandske rentenivået er knyttet til renten i det internasjonale kapitalmarkedet. I et slikt scenario er det ganske usannsynlig at selv store prosjektet vil ha noen renteeffekt på innenlandsk kapitalbruk slik at utenlandsrenten blir den relevante alternativavkastningen for innenlandsk kapitalbruk. Det gjelder også for offentlige prosjekter.

## 14.7 Behandling av risiko i prosjektanalysen.

Kapitalkrevende prosjekter har ofte lang økonomisk levetid. Det innebærer at prosjektenes lønnsomhet er avhengig av faktorer som ligger langt frem i tid og som det kan herske betydelig usikkerhet om på beslutningstidspunktet. Denne usikkerheten kan dels skyldes forhold som ligger utenfor beslutningstakerens kontroll. Det kan være den fremtidige nedbørsituasjonens betydning for lønnsomheten av et kraftverk, trafikkutviklingens betydning for lønnsomheten av en ny flyplass, o.l. Usikkerheten kan også skyldes forhold som i noen grad kan påvirkes av beslutningstakeren enten ved prosjektets utforming og gjennomføring, eller ved at en kan bruke ressurser på å skaffe til veie informasjon om viktige lønnsomhetsfaktorer som det hersker usikkerhet om på beslutningstidspunktet.

Det er vanlig å anta at ved viktige økonomiske beslutninger har de aller fleste aversjon mot risiko. Det innebærer at i forhold til den matematiske forventningen til det usikre resultatet av en beslutning, legges det større vekt på negative avvik enn på positive avvik. Når resultatet måles i penger, innebærer det at det sikkerhetsekvivalente resultatet er lavere enn den forventede verdien og differansen er den kompensasjon som en må ha for å være villig til å bære den risikoen som beslutningen medfører. Risikobæring fører med andre ord til en kostnad og denne kostnaden må tas med i lønnsomhetsanalysen av prosjektet. For risiko som ligger utenfor beslutningstakerens kontroll, blir spørsmålet om hvordan denne risikokostnaden kan kvantifiseres og hvordan den bør komme inn i lønnsomhetskalkylen. For risiko som kan påvirkes enten ved informasjonsinnhenting om kritiske lønnsomhetsfaktorer før prosjektet igangsettes, eller gjennom mer fleksible prosjektkonsepter, blir oppgaven å beregne kostnadene og gevinstene ved ytterligere informasjon eller ved mer fleksible prosjektløsninger for å være tilpasningsdyktig i forhold til endringer i underliggende lønnsomhetsfaktorer. Den første typen av risiko kan vi anse som eksogen risiko som beslutningstakeren må forholde seg til. Den andre typen er endogen i den forstand at den avhenger i noen grad av prosjektets utforming og gjennomføring.

### Eksogen risiko og prosjektlønnsomhet.

Det er en grunnleggende innsikt fra porteføljetilnærmingen til investeringer under usikkerhet<sup>51</sup> at en kan ikke se på et prosjekts risiko i isolasjon, men prosjektets bidrag til den samlede økonomiske risiko som beslutningstakeren bærer. Den er gitt ved risikoen knyttet til avkastningen på beslutningstakerens totalportefølje. Risikoen for et prosjekt avhenger med andre ord av hvilken portefølje prosjektet inngår i og hva beslutningstakeren kan gjøre for å sikre seg mot slik risiko. En viktig sontring i den sammenheng er mellom usystematisk og systematisk risiko.

Usystematisk risiko er risiko som kan reduseres ved at prosjektet inngår i en større portefølje der negative og positive variasjoner i resultatene i varierende grad opphever hverandre. Mulighetene for å redusere risikoen på investeringsporteføljen ved å spre kapitalen på flere prosjekter avhenger derfor av hvordan avkastningen på enkeltprosjektene er korrelerte. Om et prosjekts avkastning er perfekt negativt korrelert med avkastningen på porteføljen bestående av de øvrige prosjektene, vil prosjektet kunne eliminere porteføljerisikoen, og prosjektet kan ses som en sikring av den øvrige porteføljen. Dersom det er negativ men ikke perfekt negativ korrelasjon, vil porteføljerisikoen bli redusert, men ikke eliminert. Bare for prosjekter der

---

<sup>51</sup> Markowitz (1959)

avkastningen er ukorrelert med porteføljeavkastningen er prosjektets bidrag til porteføljerisikoen gitt ved dets egenrisiko. Prosjekter med avkastning som er positivt korrelert med porteføljeavkastningen, vil ha et bidrag til porteføljerisikoen som overstiger egenrisikoen.

Systematisk risiko karakteriserer prosjekter der lønnsomheten er avhengig av en eller flere usikre faktorer som er felles for prosjektene. Det kan f.eks. være lønnsnivå, importpriser, den internasjonale markedsrenten, o.l. Dette gjør at prosjektenes avkastning blir positivt korrelerte. Systematisk risiko kan ikke reduseres eller elimineres ved diversifisering som består i å spre investeringene på flere prosjekter. Dersom en har effektive risikomarkeder for spredning av risiko, er det derfor bare prosjektets bidrag til den systematiske risikoen som en trenger å håndtere i lønnsomhetskalkylen.

For offentlig sektor som investerer i en lang rekke prosjekter, vil risikoen til det offentliges investeringsportefølje være gitt ved den systematiske risikoen for porteføljen. Men når det offentlige anlegger en samfunnsøkonomisk lønnsomhetsbetraktning for sin investeringer, vil den relevante risikoen knyttet til et offentlig prosjekt være gitt ved prosjektets virkninger for variabiliteten i landets samlede verdiskaping, som er gitt ved nasjonalinntekten for den verdiskapingen som skjer gjennom markedet. I et samfunnsøkonomisk perspektiv vil trolig det aller meste av usystematisk risiko vaskes bort i landets totale portefølje av investeringsprosjekter

### **Diskontering og risikojustering**

Risiko kan bli ivarettatt på to ulike måter i en samfunnsøkonomisk lønnsomhetskalkyle. En kan erstatte usikre fremtidige prosjektoverskudd med sikkerhetsekvivalente overskudd, og så diskontere disse ned med den risikofrie diskonteringsrenten. Det betyr i praksis at anslag for fremtidige inntekter nedjusteres og fremtidige kostnader oppjusteres for å ta høyde for risiko. Fordelen med denne fremgangsmåten er at en får skilt mellom tidspreferanse og risikovurdering når det gjelder verdsetting av framtidige inntekter og kostnader.

Alternativt kan en å ta hensyn til risiko ved å justere opp diskonteringsrenten, og så diskontere ned forventningsverdien av de fremtidige prosjektoverskudd med en risikojustert diskonteringsrente. Risiko håndteres på denne måten ved et økt avkastningskrav, som gjennom økt diskonteringsrente fører til en nedjustering av forventet nåverdi.

Prinsipielt er det mer tilfredsstillende å håndtere risiko gjennom risikojustering av anslagene for fremtidige inntekter og kostnader enn ved risikojusterte diskonteringsrenter. Ved justering av risiko i størrelsene "over brøkstreken" i investeringskalkylen får en et klarere bilde av hvor stor del av nedjusteringen av fremtidige verdier som skyldes preferanse for fremskyndet konsum, og hvor mye som kan tilskrives risikojustering. Dermed blir det enklere å forholde seg til hva som er en fornuftig risikojustering og riktig risikofri diskonteringsrente. Om en velger å justere diskonteringsrenten for risiko, vil en ofte ut fra praktiske hensyn benytte en konstant diskonteringsrente over prosjektets levetid. Men for at dette skal være teoretisk tilfredsstillende, forutsetter det en geometrisk utvikling av risikoen knyttet til periodeoverskuddene over tid via rentes-rente effekten av risikopåslaget. Dersom prosjekt-risikoen er milepælsrisiko som oppløses tidlig i prosjektets levetid, vil en slik framgangsmåte kunne føre til en betydelig overvurdering av prosjektrisikoen etter at milepælen er passert

Mange prosjekter kan medføre store miljøkostnader i lang tid etter at virksomheten er nedlagt. Eksempler på dette er deponering av radioaktivt avfall fra kjernekraftverk og fjerning av faste



oljeutvinningsplattformer til havs. Med risikoaversjon taler usikkerhet på prosjektanalysetidspunktet om størrelsen på slike kostnader for at de blir oppjusterte i forhold til forventet kostnad. En risikojustert diskonteringsrente vil imidlertid ha motsatt effekt av det en ønsker å oppnå, og det viktig at en i lønnsomhetsanalysen er seg bevisst slike utslag.

### Markedsprisen på risiko og det samfunnsøkonomiske avkastningskravet

Prinsipielt kommer det ut på det samme om vi neddiskonterer forventede prosjektoverskudd med et risikojustert avkastningskrav, eller om vi neddiskonterer sikkerhetsekvivalent overskudd med den risikofrie renten siden det ene kan avledes fra det andre. Komponenten for risikojustering i diskonteringsrenten kan vi se på som en kalkulert pris for prosjektets risiko. Generelt vil det være en fordel å knytte samfunnsøkonomiske kalkylepriser til markedspriser slik at vi så langt som mulig kan benytte objektiv prisinformasjon i markedet i den samfunnsøkonomiske lønnsomhetsanalysen. I aksjemarkedene blir risikoen priset i form av et risikojustert avkastningskrav. Når markedsinformasjon om risikoprising er gitt i form av en risikokorrigert avkastning, kan en lettere utnytte markedsprisen på risiko dersom en korrigerer for prosjektets risiko med en risikojustert diskonteringsrente i stedet for ved risikojusterte anslag for prosjektoverskudd

Risikoen knyttet til en enkeltstående investering i porteføljen kan kvantifiseres ved kovariansen mellom investeringens avkastning og porteføljeavkastningen. Når vi bruker variansen til porteføljeavkastningen som risikomål på totalporteføljen, vil forholdet mellom denne kovariansen og variansen til totalavkastningen uttrykke prosjektrisikoen som andel av porteføljerisikoen. Dette blir kalt for investeringens beta ( $\beta$ ). Dersom beta er lik null, betyr det at investeringens bidrag til porteføljerisikoen er lik null. Investeringen er dermed å anse som risikofri.  $\beta = 1$  betyr at investeringen har presis samme risikoprofil som totalporteføljen. Dersom vi forestiller oss at hver enkelt investering er en aksje og totalporteføljen er alle aksjer som handles i aksjemarkedet - den såkalte børsporteføljen - vil en aksjes betaverdi vise hvordan den usikre aksjeavkastningen samvarierer med avkastningen til markedsporteføljen. Vi kan videre anta at den risikofrie avkastningen i markedet er lik  $r$  og at den forventede avkastning på børsporteføljen er lik  $E(R)$ . Den forventede meravkastningen  $E(R) - r$  er dermed den meravkastningen i forhold til den risikofrie markedsavkastningen som markedet må ha for å være villig til å holde børsporteføljen. Når vi multipliserer denne meravkastningen med aksjens betaverdi, får vi den meravkastning som markedet krever for å være villig til å holde den aktuelle aksjen. Vi benevner markedets risikojusterte avkastningskrav til denne aksjeinvesteringen  $k$ . Vi har da

$$k = r + \beta[E(R) - r]^{52}$$

For  $\beta = 0$  blir markedets avkastningskrav lik den risikofrie renten, mens det for  $\beta = 1$  blir lik  $E(R)$  som er avkastningskravet til markedsporteføljen.

---

<sup>52</sup> Teorien bak denne formelen for risikojusterte avkastningskrav for aksjeinvesteringer med usikker avkastning går under navnet kapitalverdimodellen (Capital Market Pricing Model - CAPM). Se f.eks. Mossin (1982) for en presisering av modellens forutsetninger og utledning av dette investeringskriteriet)

Denne tilnærmingen kan legges til grunn for anslag på det risikojusterte samfunnsøkonomiske avkastningskravet til en investering. Analogien til markedsporteføljen vil nå være landets samlede investeringsportefølje og avkastningen på denne representeres ved nasjonalinntekten. Om vi gjør den forenklede antagelsen at børsporteføljens risikoprofil er representativ for risikoprofilen til landets samlede investeringsportefølje, kan vi legge aksjemarkedets prising av risiko til grunn for den samfunnsøkonomiske prisingen av risiko. Det samfunnsøkonomiske avkastningskravet kan da bestemmes ved å finne den beta-verdien som er representativ for investeringens risikoprofil.

En viktig forutsetning for å bygge den samfunnsøkonomiske kostnaden knyttet til risikobæring på prising av risiko i aksjemarkedet, er at risikoprofilen til avkastning på børsporteføljen er noenlunde representativ for risikoprofilen til nasjonalinntekten og at risikoaversjonen til dem som handler i børsnoterte aktiva er representativ for risikoaversjonen til dem som bærer risikoen knyttet til realinvesteringer generelt, og spesielt i offentlig sektor. Rent praktisk vil det å basere seg på børsdata innebære at en til en gitt offentlig investering finner en børsnotert risikokopi - dvs med samme risikoprofil - og at en så legger beta-verdien til denne risikokopien til grunn. Det forutsetter at avkastningskravet til børsporteføljen, dvs investeringer med beta-verdi lik 1, også reflekterer det samfunnsøkonomiske avkastningskravet for investeringer med børsporteføljens risikoprofil.

Det er imidlertid en viktig forskjell mellom den avkastningen som prises på børsen og den avkastningen som er relevant for samfunnet. Børsverdien og den avledete prisen på risiko vedrører den avkastning som tilfaller aksjonærene i de underliggende selskapene. Den samfunnsøkonomiske avkastningen omfatter imidlertid også den delen av avkastningen som tilfaller dem som har stilt lånekapital til rådighet for selskapet, og i tillegg den delen som tilfaller myndighetene som skatt. For å gjøre den samfunnsøkonomiske avkastningen på en offentlig investering risikomessig sammenlignbar med den avkastningen som prises på børsen, bør en ta utgangspunkt i det børsnoterte selskapets totalavkastning før selskapsskatt. Det underliggende prinsipp blir da å finne et børselskap med en risikoprofil på totalavkastningen som er representativ for risikoprofilen for den offentlige investeringen. Markedets krav til totalavkastning før skatt for dette selskapet blir da et markedsbestemt avkastningskrav for den offentlige investeringen.

Siden aksjonærene i hovedsak bærer den økonomiske risikoen i aksjeselskaper, sier det seg selv at beta-verdien for totalavkastningen er lavere enn for egenkapitalavkastningen. Om vi antar at selskapets gjeld er fullstendig sikker (konkursrisiko lik null), vil beta-verdien for avkastningen på totalkapitalen være  $\alpha\beta_E$  der  $\alpha$  er egenkapitalandelen og  $\beta_E$  er beta-verdien til egenkapitalavkastningen. En må imidlertid også ta hensyn til at meravkastningskravet  $E(R)-r$  for å holde børsporteføljen og som kan utledes fra børsdata, er etter selskapsskatt og personskatt på aksjeinntekt. Meravkastningskravet blir ofte kalt for markedets risikopremie og vi lar  $MP^s$  stå for denne risikopremien etter person- og selskapsskatt. Når selskapsskattesatsen er lik den personlige kapitalinntektsskattesatsen - som i Norge - har vi at  $MP^s = E(R^s) - r(1-t)$  der toppskriften  $s$  står for etter skatt og  $t$  er den felles skattesatsen. Egenkapitalavkastningskravet etter skatt,  $k_E^s$ , blir da  $k_E^s = (1-t)r + \beta_E MP^s$

Dersom vi ser bort fra at långivere også kan bære risiko<sup>53</sup>, blir det risikojusterte avkastningskravet på totalkapitalen før skatt lik

$$k_{Tot} = \frac{k_E^s}{1-t} = r + \frac{\alpha\beta_E}{1-t} MP^s$$

Dersom vi antar en risikofri realrente lik 2,5%, gjennomsnittlig egenkapitalandel lik 0,4 for børsnoterte selskaper, en risikopremie etter skatt i aksjemarkedet lik 4,5% og skattesats lik 28%, får vi fra ovenstående formel et risikojustert realavkastningskrav på 5% for en offentlig investering som har en risikoprofil for totalavkastningen før skatt lik den for et representativt børsnotert selskap ( $\beta_E = 1$  og  $\alpha = 0,4$ ).

## 14.8 Prosjektevaluering i tidligfase.

I en tidligfase i planleggingen og lønnsomhetsevalueringen av et prosjekt vil en typisk stå overfor et uttrykt behov eller etterspørsel etter en nærmere bestemt type tjenester, f.eks. et transportbehov, og problemstillingen er å finne frem til et prosjektkonsept som tilfredsstiller dette behovet på en mest mulig effektiv måte. I noen tilfelle kan dette ta form av en kostnadseffektivitetsanalyse som går på å finne frem til en prosjektutforming som realiserer et bestemt mål til lavest mulig kostnad. En slik effektivitetsanalyse vil imidlertid kompliseres av at den fremtidige nytte- og kostnadssiden til prosjektet kan variere på en ikke fullt ut predikerbar måte sett fra analysetidspunktet. Om forventet kostnad legges til grunn, kan det skjule det forholdet at prosjektkonseptet kan føre til svært høye kostnader i noen scenarier.

### Diskonteringsrentens betydning for lønnsomhetsanalysen i prosjektets tidligfase.

Det kan her være viktig å slå fast at problemet med å evaluere lønnsomheten av et prosjekt i planleggingsfasen, skyldes i hovedsak manglende informasjon om de viktigste faktorene som er bestemmende for lønnsomheten. Dette må først og fremst løses gjennom innhenting av ytterligere informasjon og et valg av prosjektkonsept som er tilstrekkelig fleksibelt til å kunne dra nytte av ny informasjon. Å kompensere for informasjonsmangelen gjennom valg av diskonteringsrente er lite målrettet. Avkastningskravet vil likevel være en benchmark som avkastningen på et lønnsomt prosjekt må matche. Slik sett fører et gitt avkastningskrav til at potensielle prosjekter blir sortert i to grupper: en lønnsom og en ulønnsom. Når en har nokså omtrentlig informasjon om lønnsomheten i prosjektet, kan det imidlertid være en betydelig risiko for feilsortering. Slike sorteringsfeil kan være av to typer. Den ene er at en aksepterer prosjekter eller prosjektkonsepter som ikke burde ha vært akseptert. I analogi med statistisk inferensteori kan vi kalle dette for feil av type II. Den andre er at prosjekter som burde ha vært akseptert, blir forkastet som ulønnsomme i planleggingsfasen. Dette blir da feil av type I. Gitt at informasjonen i utgangspunktet er mangelfull, vil begge typer beslutningsfeil kunne forekomme. Dersom akseptkriteriet er at forventet nåverdi skal være positiv, blir den relative forekomsten av disse to typene beslutningsfeil avhengig av avkastningskravet. Med et høyt avkastningskrav blir det færre prosjekter som passerer hinderet og en får færre feil av type II, men samtidig flere av type I. Omvendt vil et lavt avkastningskrav føre til at flere prosjekter

<sup>53</sup> Vi antar at lånerentene tilfaller innenlandske långivere slik at renteinntektene er å anse som nasjonal inntekt.

passerer lønnsomhetshindret, og en får relativt flere feil av type II, men færre av type I. I et slikt perspektiv vil avkastningskravet som preliminært sorteringskriterium avhenge av hvilken type feil som anses som mest alvorlig. For spesielt irreversible prosjekter kan det være viktig å unngå feil av type II, dvs å unngå å akseptere prosjekter som er ulønnsomme gitt korrekt informasjon. Dette kan tale for et høyt avkastningskrav i den innledende sorteringen mellom lønnsomme og ikke lønnsomme prosjekter.<sup>54</sup>

Generelt gjelder det at jo mindre informasjon en har om lønnsomheten, desto mer risikofyllt vil prosjektet være. Risikotillegget i avkastningskravet er avhengig av prosjektets systematiske risiko. I tidligfasen vil den systematiske risikoen kunne påvirkes gjennom prosjektets utforming. Alt annet likt gjelder det da å velge et prosjektkonsept som minimerer den systematiske risikoen. Generelt vil det si å prøve å finne frem til utforminger som gjør prosjektets lønnsomhet mindre konjunkturbetinget. Det kan f.eks.være grunn til å hevde at norske borgere er overeksponerte for oljerelatert risiko. Slik sett vil det fra et samfunnsmessig risikosynspunkt kunne være et viktig moment for valg av konsept at prosjektets lønnsomhet i minst mulig grad blir positivt avhengig av oljeprisen. På denne måten kan prosjektets risiko og avkastningskrav påvirkes gjennom valg av prosjektutforming og blir på denne måten i noen grad endogent bestemt.

### **Endogen risiko: Innhenting av informasjon og mulighet for å dra nytte av ny informasjon ved utforming og timing av prosjektet.**

De fleste store investeringsprosjekter er irreversible. Det gjelder f.eks. samferdselsprosjekter som veier, tunneler og flyplasser. Det betyr at størstedelen av investeringskostnaden ikke kan gjenvinnes når investeringen først er foretatt. Dernest vil det herske usikkerhet om viktige lønnsomhetsdrivende faktorer på investeringstidspunktet. Dersom risikoen er av typen milepælsrisiko, vil noe av risikoen bli oppløst etter hvert som tiden går. For det tredje vil en ofte ha et valg med hensyn til timing og måten som prosjektet gjennomføres på. Det kan f.eks. være fornuftig å vente på ny informasjon eller på at milepælsrisiko blir oppløst. Optimal timing vil da være bestemt ved en vurdering av kostnaden ved å vente mot verdien av å holde investeringsbeslutningen åpen (holde investeringsopsjonen åpen). Vente-og-se-beslutningen krever derfor en metode for å beregne hva investeringsopsjonen er verdt, som viser hvor mye en kan være villig til å ofre for å beholde beslutningsfleksibilitet, dvs holde muligheten åpen for å ikke gjennomføre prosjektet. Tilsvarende kan en velge en trinnvis gjennomføring av prosjektet dersom en kan regne med å få ny informasjon underveis i prosjektperioden. Når det gjelder bygging av en ny flyplass og den viktigste risikofaktoren er usikkerhet om den fremtidige trafikkutviklingen, kan det være fornuftig med en trinnvis utbygging. Det kan skje ved at en starter med én rullebane, og lar utvidelsen til to rullebaner være avhengig av trafikkutviklingen. En trinnvis utbygging vil normalt bli dyrere enn om en hadde bygget den i full skala umiddelbart. Denne merkostnaden er imidlertid å anse som en form for forsikringspremie mot det uheldige utfallet å bli sittende med overkapasitet i tilfelle av en

---

<sup>54</sup> Det kan også argumenteres for at den relative forekomsten av disse to typene av beslutningsfeil avhenger av måten prosjektutredningsarbeidet organiseres på. En hierarkisk organisasjonsstruktur med mange kontrollposter (decision gates) vil minimere forekomsten av type II feil, mens en flatere struktur med delegert beslutningsmyndighet vil minimere forekomsten av type I feil. Se Sah & Stiglitz (1986) for en nærmere diskusjon av dette.

ugunstig trafikkutvikling. Lønnsomheten av en fleksibel løsning med trinnvis utbygging vil avhenge av hvor mye tilleggsinformasjon en får ved å vente, og den risikoaversjonen en har.

### **Informasjonsinnhenting før beslutning fattes.**

Forventningsverdiene til fremtidige prosjektoverskudd er basert på den informasjon som er tilgjengelig på beslutningstidspunktet. Det er da en underforstått forutsetning at informasjonsinnhenting forut for beslutningen er optimalisert. Prosjektrisikoen reduseres ved innhenting av supplerende informasjon før beslutningen fattes. Generelt gjelder det at det er lønnsomt å innhente prosjektrelevant informasjon så lenge som forventet verdi av ytterligere informasjon er større enn kostnadene. Verdien av ytterligere informasjon er gitt ved sannsynligheten for at ny informasjon skal endre optimal beslutning multiplisert med nyttegevinsten ved endret beslutning betinget av den nye informasjonen. Operasjonalisering av denne beslutningsregelen krever at en er i stand til å oppdatere sannsynlighetene for lønnsomhetsrelevante utfall i lys av ny informasjon. Det er imidlertid viktig å merke seg at for at supplerende informasjon skal ha verdi, må det være en sannsynlighet forskjellig fra null for at den skal lede til endret beslutning. Dersom optimal beslutning i forhold til eksisterende informasjon om lønnsomhetskritiske faktorer ikke vil kunne påvirkes av ytterligere informasjon, er tilleggsinformasjon verdiløs.

Vi antar at prosjektet dreier seg om en veitunnel i fjell fra A til B, og kostnadene er svært avhengig av hvor godt fjellet egner seg for tunneldriving. Dersom valget står mellom en slik tunnel og en alternativ veiforbindelse som ikke forutsetter tunnel, vil informasjon om fjellets beskaffenhet kunne være viktig for trasévalget. I en slik situasjon kan det være lønnsomt å bruke ressurser på prøveboringer for å avdekke de faktiske forhold for tunnelarbeid. Om det derimot ikke fins brukbare alternativer til tunnelen og veiforbindelsen mellom A og B er absolutt nødvendig, vil denne informasjonen ikke ha noen betydning for beslutningen (men den kan selvsagt ha betydning for planleggingen av tunnelarbeidet).

Det følger da av dette at ytterligere informasjon vil ha størst verdi i planleggingsfasen av et prosjekt der mulighetene for tilpasning til ny informasjon er størst. Etter hvert som prosjektet gjennomføres og ulike irreversible valg blir tatt, reduseres mulighetsområdet for alternative prosjekttilpasninger og ny informasjon blir mindre verdt. En irreversibelt prosjekt som sett fra beslutningstidspunktet er ulønnsomt vurdert på grunnlag av den informasjon en har i etterhånd, vil derfor kunne være lønnsomt å drive videre når det først er ferdigstilt<sup>55</sup>.

### **Opsjonstilnærming til investeringer under usikkerhet.<sup>56</sup>**

Et mulig beslutningsalternativ er å utsette beslutningen om gjennomføringen av prosjektet i påvente av ytterligere informasjon. Dersom prosjektet er lønnsomt i forhold til den informasjonen som en har i utgangspunktet, betyr dette at prosjektets forventede verdiskaping utsettes. Dette fører til en ventekostnad når beslutningstakeren har en positiv tidspreferanse. Poenget med å vente med iverksettelsen må være at en får ytterligere lønnsomhetsinformasjon

---

<sup>55</sup> Her kan flytoget ved Oslo Lufthavn Gardermoen være et godt eksempel.

<sup>56</sup>Opsjonstilnærming til realinvesteringer under usikkerhet går under betegnelsen realopsjonsteori. En god innføring i realopsjonsteori vil en finne i Dixit & Pindyck (1993)

etter som tiden går, eller at det blir mindre kostbart å innhente slik informasjon på et senere tidspunkt. Dette vil spesielt gjelde milepælsrisiko, som helt eller delvis oppløses på et bestemt fremtidig tidspunkt. Dette ville imidlertid ikke i seg selv være noe argument for å vente med prosjektet dersom en uten kostnader kunne komme tilbake til utgangssituasjonen, dvs situasjonen før prosjektet ble iverksatt, dersom informasjonstilgangen på et senere tidspunkt skulle vise at en hadde satset på et ulønnsomt prosjekt. Hvis derimot prosjektet er irreversibelt, for eksempel på grunn av at det nødvendiggjør investeringer av typen ”sunk cost” som ikke kan gjenvinnes selv om prosjektet legges ned, vil gjennomføring straks medføre en kostnad i form av tapt beslutningsfleksibilitet i en fremtidig situasjon med ny informasjon. Utsettelse av iverksettelse gir en *opsjonsverdi*, som er forventet verdi av å ha muligheten til å kansellere prosjektet dersom den oppdaterte lønnsomhetsinformasjonen gjør prosjektet ulønnsomt. Dette kan for eksempel skje ved at det realiseres en svært ugunstig verdi for den usikre faktoren som ligger til grunn for milepælsrisikoen. Dersom opsjonsverdien er positiv, vil en positiv risikojustert forventet nåverdi ikke være en tilstrekkelig betingelse for at det er lønnsomt å gjennomføre et prosjekt straks. Den forventede nåverdien ved gjennomføring i dag må i tillegg være større enn opsjonsverdien ved å vente med beslutningen om gjennomføring av prosjektet. Opsjonsverdien blir dermed en alternativkostnad for beslutningen om umiddelbar gjennomføring.

I en viss grunnleggende forstand er det tre kilder til usikkerhet i en investeringsanalyse. Den ene er usikkerheten om fremtidige prosjektoverskudd. Den andre er usikkerhet omkring investeringskostnaden og den tredje er usikkerheten omkring fremtidig alternativkastning på den kapitalen som bindes i investeringen. Vi skal se på opsjonsaspektet ved de tre kildene til usikkerhet i tilknytning til tre enkle eksempler.

### Usikkerhet om fremtidige prosjektoverskudd.

To utbyggingsalternativer for et prosjekt er aktuelle:  $A_0$  som står for utbygging straks, og  $A_1$  som står for utbygging om ett år. Investeringskostnaden er 540 mill kroner. De årlige prosjektoverskudd er avhengig av hvilken av to mulige tilstander som inntreffer når prosjektet er ferdigstilt om ett år. I den ene tilstanden er prosjektoverskuddene 45 mill. pr år i all fremtid og i den andre tilstanden er de 15 mill. kroner pr år. Sannsynligheten for hvilken tilstand som vil inntreffe er 50:50. Diskonteringsrenten (avkastningskravet) er 5% og beslutningstakeren antas å maksimere forventet nåverdi.

Vi lar  $E[N(A_i)]$  stå for forventet nåverdi av alternativ  $A_i$ ,  $i = 0, 1$ .

$$E[N(A_0)] = \frac{0,5 \cdot 15 + 0,5 \cdot 45}{0,05} - 540 = 60^{57}$$

Dvs utbygging straks gir en forventet nåverdi på 60 mill. kroner og er dermed lønnsom i følge det konvensjonelle nåverdikriteriet.

Om prosjektet utsettes i ett år, vil nåverdien i den ugunstige tilstanden være  $(15/0,05) - 540 = -240$  og følgelig er prosjektet ulønnsomt. Dersom den andre tilstanden inntreffer, blir

<sup>57</sup> Vi har her benyttet oss av at nåverdien av en uendelig betalingsrekke med en krone pr år og med diskonteringsrente  $r$  er lik  $1/r$ . Dvs  $NV = \frac{1}{1+r} + \frac{1}{(1+r)^2} + \dots + \frac{1}{(1+r)^n} + \dots = \frac{1}{r}$

nåverdien  $(45/0,05) - 540 = 360$ . Ved utsettelse blir det dermed bare aktuelt å realisere prosjektet dersom den gunstige tilstanden inntreffer. Vurdert ut fra tidspunkt 0 blir forventet nåverdi ved å utsette prosjektet dermed lik

$$E[N(A_1)] = 0,5 \left( \frac{(45/0,05) - 540}{1,05} \right) = 171$$

og det er nesten tre ganger så høyt som i straksalternativet. Grunnen til dette er at ved å utsette prosjektet kan en velge bort tapstilstanden siden usikkerheten blir oppløst etter ett år. Differansen mellom forventet nåverdi ved ventealternativet og straksalternativet er opsjonsverdien ved å beholde beslutningsfleksibilitet til usikkerheten er oppløst, som i dette tilfellet blir lik 111 mill. Dersom den tapte opsjonsverdien regnes som en kostnad i straksalternativet, ser vi at straksalternativet er ulønnsomt. Vi kan merke oss at dersom valget står mellom investering i dag eller aldri, er straksalternativet lønnsomt da det gir en høyere avkastning enn avkastningskravet. Men dersom alternativet er å vente i ett år med beslutning om gjennomføringen, er straksalternativet ulønnsomt.

Nå kan det tenkes at fleksibiliteten ved å ha mulighet til å utsette prosjektet i ett år fører til høyere investeringskostnader. En kan da regne ut hva den maksimale økningen i investeringskostnadene kan være for at vente-å-se-alternativet ikke skal bli ulønnsomt. Dersom vi lar denne investeringskostnaden være  $I_{Max}$ , er den gitt ved

$$60 = (0,5) \left[ \frac{45/0,05 - I_{Max}}{1,05} \right]$$

som gir  $I_{Max} = 786$ . Det betyr at så lenge økningen i investeringskostnadene holder seg under 246 mill, kan kostnadsøkningen sees på som en lønnsom investering i beslutningsfleksibilitet som gjør det mulig å vente ett år med beslutningen om iverksettelse.

Tilsvarende kunne en tenke seg at en trinnvis gjennomføring av prosjektet i stedet for umiddelbar gjennomføring i full skala, gir en opsjonsverdi ved at kapasiteten kan tilpasses ny informasjon om behov og etterspørsel. En kan da regne på hvor stor kostnadsøkning en kan tåle ved trinnvis gjennomføring uten at dette blir ulønnsomt sammenlignet med umiddelbar gjennomføring i full skala.

### Usikkerhet om investeringskostnaden

Usikkerhet om investeringskostnaden kan i særlig grad gjelde kapitalintensive prosjekter som det tar lang tid å ferdigstille. Det gjelder f.eks. innenfor energisektoren som store vannkraftutbygginger og for gasskraftverk der investeringsbehov i tilknytning til renseteknologier for CO<sub>2</sub> er usikre. Vi ser på problemstillingen ved hjelp av eksemplet på side 292 som vi modifierer ved å anta at fremtidige prosjektoverskudd er helt sikre og lik 45 mill. pr år, og at investeringskostnaden om ett år er usikker, men at investeringskostnaden for investering i dag er kjent. Vi antar at iverksetting straks gir en investeringskostnad på 800 mill slik at straksalternativet gir en nåverdi på 100 mill. Om ett år antas investeringskostnaden å være 1200 med sannsynlighet = 0,5 eller 400 med samme sannsynlighet 0,5 som gir en

forventet investeringskostnad på 800 lik kostnaden i straksalternativet. Vi ser imidlertid at ved å vente blir prosjektet bare realisert hvis den gunstige tilstanden inntreffer. Forventet nåverdi blir da  $0,5 \cdot 500$  og neddiskontert til i dag blir det  $250/1,05 = 238$ , slik at vente-å-se-alternativet er klart bedre enn iverksetting straks.

### Usikkerhet om diskonteringsrenten.

Vi antar som et siste eksempel at investeringskostnaden og fremtidige prosjektoverskudd er kjent, men at diskonteringsrenten (den fremtidige alternativavkastningen) fluktuerer på en upredikerbar måte. Renteusikkerhet vil ha to effekter på en investering. For det første vil det ha en gunstig effekt på forventet nåverdi. Vi antar for eksemplets skyld at vi har uendelig tidshorison. Anta videre at den usikre renten kan ha to utfall: 5% eller 15% med sannsynlighet 50:50 slik at forventet rente er 10%. Et prosjekt som gir konstante prosjektoverskudd på 1 krone for all fremtid vil med en rente på 10% ha en nåverdi på 10 kroner. Forventet nåverdi når renten kan være enten 5% eller 15% med lik sannsynlighet, blir imidlertid  $0,5 \cdot 1/0,05 + 0,5 \cdot 1/0,15 = 13,33 > 10$ . En usikker diskonteringsrente gjør derfor investeringen mer attraktiv dersom forventet nåverdi er beslutningskriteriet<sup>58</sup>.

Men selv om renteusikkerhet er gunstig for forventet nåverdi, vil det like fullt kunne være lønnsomt å vente med beslutningen om iverksettelse til renteusikkerheten er eliminert. Det skyldes at vente-å-se-alternativet gjør det mulig å unngå en mulig tapssituasjon dersom avkastningskravet blir for høyt. Slik sett vil usikker diskonteringsrente ha samme virkning for utsettelsesbeslutningen som usikkerhet om fremtidige prosjektoverskudd. Vi anskueliggjør dette ved på nytt å modifisere eksemplet på side 292. Vi antar nå at prosjektoverskuddene er kjente og lik 45 pr år og at investeringskostnaden er 540. Når prosjektet kommer i drift om ett år, vil diskonteringsrenten være kjent og vurdert ut fra i dag antas den da å bli hhv. 5% og 15% med lik sannsynlighet.

Vi har da at nåverdien av prosjektoverskuddene er

$$45/0,05 = 900 \text{ med sannsynlighet } 0,5$$

$$45/0,15 = 300 \text{ med sannsynlighet } 0,5.$$

Følgelig blir forventet nåverdi av prosjektoverskuddene 600 som er høyere enn hva den ville ha vært hvis renten var sikker og lik 10%, da  $45/0,1 = 450$ . Men samtidig ser vi at dersom tilstanden med høy rente inntreffer, så er ikke prosjektet lønnsomt. Dersom vi investerer straks, får vi en forventet nåverdi på  $600-540 = 60$ . Om vi derimot utsetter beslutningen til renten er kjent om ett år, får vi forventet nåverdi lik

$$0,5 \left( \frac{45/0,05 - 540}{1,1} \right) = 180/1,1 = 163$$

<sup>58</sup> Teknisk sett er dette en konsekvens av den såkalte Jensens ulikhet. Den sier at forventningsverdien av en konveks funksjon av en tilfeldig variabel er større enn funksjonsverdien evaluert for forventningsverdien til variabelen. Resultatet følger da av at nåverdifunksjonen  $NV = 1/r$  er en konveks funksjon av  $r$ .



som er betydelig høyere enn nåverdien i straksalternativet.

Som oppsummering kan vi si at økt variabilitet i diskonteringsrenten, men slik at forventet rente forblir uendret (såkalt mean-preserving spread), vil øke forventet nåverdi av prosjektet. Videre blir det også mer lønnsomt å vente med beslutningen om iverksettelse av prosjektet. Grunnen til det er at ventealternativet gjør det mulig å unngå tapstilstander på grunn av høyt fremtidig avkastningskrav og sannsynligheten for slike tilstander øker med økt variabilitet i diskonteringsrenten. Generelt vil opsjonsverdien øke med økt variabilitet i de underliggende faktorer som bestemmer lønnsomheten i prosjektet.

## 14.9 Referanser

Arrow, K. J. and A. Fisher, (1974), "Environmental preservation, uncertainty and irreversibility, *Quarterly Journal of Economics*, 89, 312-319.

Dixit, A. K. and R. P. Pindyck, (1994), *Investment under Uncertainty*, Princeton University Press

Finansdepartementet, (1997), *Prinsipper for lønnsomhetsvurderinger i offentlig sektor*, NOU 1997:27

Finansdepartementet, (2005), *Veileder i samfunnsøkonomiske analyser*.

Hagen, K. P., (2005), *Økonomisk politikk og samfunnsøkonomisk lønnsomhet*. Oslo: Cappelen Akademisk Forlag

Justisdepartementet, (2006), *Når sikkerhet er viktigst - Beskyttelse av landets kritiske infrastruktur og kritiske samfunnsfunksjoner*. NOU 2006:6.

Markowitz, H. M., (1959), *Portfolio Selection*, Blackwell.

Mossin, J., (1982), *Investering og kapitalkostnad*. Tanum-Norli A/S

Sah, R.K. and J.E. Stiglitz (1986), "The architecture of economic systems: Hierarchies and Polyarchies", *American Economic Review*, 76, 716-727.